2009

Design Strategies for Adaptive Social Composition: Collaborative Sound Environments

Livingstone, Dan

http://hdl.handle.net/10026.1/1252

http://dx.doi.org/10.24382/1471
University of Plymouth

DESIGN STRATEGIES FOR ADAPTIVE SOCIAL COMPOSITION: COLLABORATIVE SOUND ENVIRONMENTS

By

DAN LIVINGSTONE

A thesis submitted to the University of Plymouth
In partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Computing, Communications and Electronics

June 2009

Abstract

In order to develop successful collaborative music systems a variety of subtle interactions need to be identified and integrated. Gesture capture, motion tracking, real-time synthesis, environmental parameters and ubiquitous technologies can each be effectively used for developing innovative approaches to instrument design, sound installations, interactive music and generative systems. Current solutions tend to prioritise one or more of these approaches, refining a particular interface technology, software design or compositional approach developed for a specific composition, performer or installation environment. Within this diverse field a group of novel controllers, described as 'Tangible Interfaces' have been developed. These are intended for use by novices and in many cases follow a simple model of interaction controlling synthesis parameters through simple user actions. Other approaches offer sophisticated compositional frameworks, but many of these are idiosyncratic and highly personalised. As such they are difficult to engage with and ineffective for groups of novices. The objective of this research is to develop effective design strategies for implementing collaborative sound environments using key terms and vocabulary drawn from the available literature. This is articulated by combining an empathic design process with controlled sound perception and interaction experiments. The identified design strategies have been applied to the development of a new collaborative digital instrument. A range of technical and compositional approaches was considered to define this process, which can be described as Adaptive Social Composition.

Dan Livingstone

CONTENTS

CHAPTER SUMMARIES

Chapter 1 describes current approaches to *collaborative sound environments* referencing significant developments and identifying specialist terms from the wider field of Computer Music. A glossary of these terms is presented, these are categorised as Features, Qualities, Models and Behaviours. This establishes a framework for distinguishing between different design approaches to new digital instruments.

Chapter 2 presents the notion of *adaptive social composition,* identifying principles and core features. Field specific vocabulary is applied in context to develop a new design approach to collaborative musical interfaces. Compositional approaches are discussed and specialist terms are further unpacked. Principles for structuring turn based or collaborative interactions are established. The core elements for a new form of composition, Adaptive Social Composition, are established.

Chapter 3 articulates *design strategies* for interface, environment and software design. The importance of *experiment design* is identified and results from a controlled test using *analysis of variance* are presented for developing interaction and behaviours. Design strategies for Adaptive Social Composition are developed considering *interface design*, *collaborative environment,* and *mediating software* as an integrated model.

Chapter 4 presents the Orb interface design as a *proof of concept*. Interaction design and classes of gesture are explained to demonstrate how the design strategies developed in this thesis have been applied to implement a novel controller that manifests the principle *'easy to learn, difficult to master'*.

Chapter 5 *Conclusion* identifies contributions to knowledge, and addresses the core research questions posed in this thesis:

Firstly, can a musical interface be designed that engages the novice but has the expressive qualities and personalisation of a traditional instrument?

Secondly, can the principles found in turn based board games be used to develop social composition frameworks that are intuitive to use in a collaborative musical context?

And thirdly, how would one describe such a system, what qualities and characteristics would it manifest and how would one design it?

LIST OF TABLES AND ILLUSTRATIONS

All photographs, diagrams and charts are original.

LIST OF PUBLICATIONS

A number of peer reviewed international conference papers were published, reporting the various stages of this research. These are referred to in the main text and where appropriate extracts are provided. Published papers are also included in appendix 2.

LIVINGSTONE, D. 2001 The Space Between the Assumed Real and The Digital Virtual *In:* Ascott, R. *Reframing Consciousness Art, Mind and Technology* Intellect Books 2001 ISBN: 1841500135 pp.138-143

LIVINGSTONE, D. 2003 Emergent Behaviour in the context of Reactive Compositional Environments *In: Proceedings of the IX Brazilian Symposium on Computer Music (IX SBCM) Music as Emergent Behaviour, 2/8 August 2003, University of Campinas,* Brazil. pp.235-240

LIVINGSTONE, D. MIRANDA, E. 2004 Composition for Ubiquitous Responsive Environments *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami, Florida: International Computer Music Association.*
pp.321-325

LIVINGSTONE, D. MIRANDA, E. 2005 Orb3 - Adaptive Interface Design for Real time Sound Synthesis and Diffusion within Socially Mediated Spaces *In: proceedings of the International Conference on New Instruments for Musical Expression, 26/28 May 2005, Vancouver,* Canada. pp.65-69

LIVINGSTONE, D. O'SHEA C. 2005 Tactile Composition Systems for Collaborative Free Sound *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami, Florida: International Computer Music Association.*
pp.687-690

LIVINGSTONE, D. MIRANDA, E. 2005 Orb3 - Musical Robots within an Adaptive Social Composition System *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami,* Florida: International Computer Music Association.
pp.543-546

ACKNOWLEDGEMENTS

Firstly, I thank my supervisors Professor Eduardo Miranda and Professor Angelo Cangelosi for clarity of guidance and support, this has been invaluable for establishing both a theoretical framework for Adaptive Social Composition Systems and resolving experiment design for extending and evaluating interaction within such systems. Colleagues and fellow researchers within the Interdisciplinary Centre for Computer Music Research have also contributed through research presentations and seminar discussion, while peers in the field have commented objectively on my conference publications for ICMC and NIME. I also extend my thanks to John Eastment and Nathan Poate who have provided invaluable technical support. My colleague Peter Anders was also a source of inspiration with many lively discussions on the nature of interaction within Cybrid environments in the initial stages of this research. Finally, my immediate family Ursula, Isaac and Eva have been patient and supportive while I have prioritised this research.

AUTHORS DECLARATION

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Committee.

Relevant interdisciplinary seminars and conferences were regularly attended at which work was presented. Several peer reviewed papers appeared in proceedings of International Conferences specific to the field of study. Presentations were given at research seminars organised by the Interdisciplinary Centre for Computer Music Research (ICCMR), previously a number of presentations were also given at related interdisciplinary conferences as a member of Centre for Advanced Enquiry into Interactive Arts (CAIIA) 1999 to 2002.

Conference Presentation and Publication.

- International Computer Music Conference, Barcelona 5th to 9th September 2005
- New Interfaces for Musical Expression Conference 26th to 28th May 2005
- International Computer Music Conference, Miami, 1st to 6th November 2004
- IX Brazilian Symposium on Computer Music, Campinas 2nd to 8th August 2003

- Consciousness Reframed 2000 3^RD International Research Conference Centre for Advanced Inquiry in the Interactive Arts University of Wales College, Newport.23 - 30 August 2000

Other presentations:

- Invited Speaker Europrix Instructors Network Conference, Vienna 2005
- Invited Panel Member 10th International Symposium on Electronic Art) 2000, Paris, France, December 7-10, 2000.

Word count of main body of thesis: 70,750

Signed

Date 17th March 2008

Introduction

This thesis identifies a range of established fields and related musical approaches to inform the design of interactive music systems intended for social or collaborative interaction. A collection of 'Design Strategies' have been identified and developed by investigating representative works from these fields, evaluating core features and observing participant behaviour and direct interactions with novel musical interfaces. A 'Design Strategy', in this context, can be summarised as a conceptual tool to resolve a practical problem, while using field specific terms to encapsulate an interaction mode or gestural process. The principle, 'easy to learn, difficult to master' is a useful phrase to summarise a system that features an evolving rule-set. But to apply this principle within a collaborative musical environment the context and mode of interaction needs to be considered. A Design Strategy goes beyond reiterating a generic principle, and can encompass interaction mode, gesture capture method and compositional context using field specific terms. This approach is intended to facilitate knowledge transfer to inform the design of new digital instruments. The phrase 'Adaptive Social Composition' is used in the title of this work and in the main text to describe a musical environment that is software or interface mediated. The term adaptive is used to express that this software or interface mediation is context sensitive. Interactions or gestures of participants are monitored and parameters remapped to provide a dynamic compositional context. In certain fields of computer music, the term 'adaptive' is applied specifically to a system that is based on a neural network. In a collaborative context, the term can also be used to describe interactions between people and musical processes that are mediated by software. This form of software mediation is context sensitive. For example, musical parameters are remapped according to monitored user actions or events within a compositional framework, new gestures can be learnt, and previous exchanges can be recalled and transformed during performance. This can be achieved by tracking and comparing gestures, using sound analysis, or by building and exchanging a shared repertoire. Simple pattern recognition can be used to evaluate actions or sound material in context. When certain gestures or sound combinations are identified or repeated, alternate feedback or parameter mappings can be introduced. By contextualising musical interactions new gestures can be learnt by software to evolve new material. Social Composition refers to the principle of group interactions influencing sound material though direct and indirect participant actions. This phrase is proposed for works which combine novel 'Tangible interfaces' and software mediation with the notion of an 'open work'. An 'open work' can be understood as a compositional framework where sound material and interactions are provided and a set of relationships are identified but performers or participants are free to reinterpret this compositional framework and generate new relationships. So 'Adaptive Social Composition' describes a real–time collaborative process that is exploratory, based on a conversational model of listening and exchange to influence a shared composition. The phrase 'Collaborative Sound Environment' is used to clarify the form of interaction within such a compositional framework. A collection of interactions, sound material and sound processing are combined or generated in response to participants presence or direct interaction with an emphasis on co-creation between people and a musical system comprising software and novel interfaces.

The literature for the field of computer music is extensive, and yet although considerable documentation of instruments, systems and methods are available, the development of novel or tangible interfaces tends to be technology driven or focussed around a particular interaction model such as 'control' or 'expressive gesture'. This limitation can be overcome by applying an empathic design process, working with end users of these technologies to identify effective solutions collaboratively. The author attended a number of international conferences, conference workshops, improvisation sessions and performances to develop this research. A wide range of novel controllers, digital instruments and sound installations were investigated, a small selection of these works are critiqued to articulate a range of interaction modes, system models and participant behaviours.

Recent published research papers, exhibited works or digital instruments are considered in the context of Adaptive Social Composition, which employs the concept of 'plasticity' in the way different elements of such a system can be integrated through software design and structured interaction modalities. In contrast, many contemporary musical interfaces are designed for use by groups of novices. These often follow a simple percussion metaphor with direct control over sampling or synthesis processes. These 'Tangible' interfaces are intended to be intuitive to use, in fact, many designs follow a turn-based model of object placement and pre determined structural relationships. These approaches echo traditional board games in terms of a controlled environment with hand placed playing pieces with simple rules and groups of 2 to 4 players. However, many lack the sophisticated interplay that the combinations of moves and context sensitive rules that are inherent in a well designed game. As such, many of these systems remain 'novel', with little credibility as effective shared musical instruments. They simply do not have the expressive range of either a traditional instrument or a custom built digital instrument. This thesis argues that this is because the context and nature of this form of social interaction is not developed effectively in a musical context. Without the invisible infrastructure of a well-designed set of possible actions and responses, these systems are reduced to the collaborative control of basic synthesis or effects parameters. In some cases, the objects themselves are generic, a simple Perspex puck or counter with no indication of its associated function. I propose a new compositional form combining moveable objects with a tracking system and mediating software for collaborative co-creation between people and a mediating system. At the core of such a system, the interface or digital instrument is a highly significant element, and yet there is little information beyond the field specific technical integration of components and software to motivate effective development. This integrated approach has considerable potential if a compositional framework can be developed that is responsive to users and not simply limited to the placement and orientation of generic passive objects. To enable the development of this compositional form, Adaptive Social Composition, a series of Design

Strategies are presented. To develop these strategies it is first is necessary to deconstruct the problem.

This thesis poses the following research question, presented in three parts:

Firstly, can a musical interface be designed that engages the novice but has the expressive qualities and personalisation of a traditional instrument?

Many 'novel' controllers have been designed to extend 'expressive' control or support collaborative interaction with computer music, but few go beyond a literal sensing of gestures and fixed parameter mappings. Others are customised variations of conventional instruments, highly personalised with intangible controls. Many of the sounds these interfaces produce are based on signal processing chains; a similar effect can be achieved with even the earliest synthesiser by randomly plugging and unplugging jack cables to re-route a signal.

Secondly, can the principles found in turn based board games be used to develop social composition frameworks that are intuitive to use in a collaborative musical context?

These long established principles include the notion that a board game should be easy to learn but difficult to master. That evolving rule-sets can be used to teach novices a new game, whilst retaining interest for experienced players. That in addition, known moves or structural devices can motivate individual strategies and responses that are context sensitive. The concept of perfect and imperfect information may also be useful for developing context sensitive compositional frameworks using different classes of gesture. The principles of video games interaction may also be used to develop more engaging novel interfaces that have adaptive features; special moves, combos, context based goals and challenges with adapting skill levels are well established features of commercial game titles. Simple pattern recognition is used to provide dynamic context and focus user actions through a simple tangible control system, the basic game-pad.

And thirdly, how would one describe such a system, what qualities and characteristics would it manifest and how would one design it?

There are numerous related works from diverse fields, which have distinctive features that are highly significant. How are these systems evaluated and what language is used? The literature in the field of Computer Music is extensive and yet there are no satisfactory design strategies that combine interaction model, system model or participant behaviours. To effectively design collaborative musical interfaces the intended context for interaction needs to be understood and ultimately it is the nature of engagement, the relation between *gesture and mind* (Paine, 2002) that the designer must address if these interfaces are to develop beyond existing models of musical interaction.

The established design approaches of these disparate fields (Machover and Chung, 1989), documented through dedicated journals and international conferences are also considered in terms of interaction methods and compositional approaches in order to differentiate between them and identify complimentary design strategies.

A summary analysis of selected works representative of each identified field is considered to identify new compositional techniques or potential new musical forms. The established models of player and instrument, composer and performer, artist, audience and installation have already been thoroughly discussed by researchers in these fields with the well-documented development of numerous new musical practices. Key papers and current development within these fields are referenced throughout, identifying key terms that begin to establish a new vocabulary for evaluating and extending new interactive and adaptive systems within the broader field of Computer Music. Selected examples of novel interfaces, compositional processes and collaborative live practices have also been documented on digital video during the process of this research.

Diverse groups of practitioners are developing new forms of organised sound, new interfaces for musical expression, new gestural and tactile interfaces for control, composition and performance with interactive music systems. All based on the pioneering work of previous generations of performers, composers, artists, designers, broadcasters, instrument designers, engineers, technicians and individual scholars. There are several significant international conferences where these diverse individuals present their research and gather to perform and demonstrate new works, novel instruments or systems. The International Computer Music Conference ICMC and New Interfaces for Musical Expression (NIME) provide an annual focus for new developments in these fields, while international festivals such as Ars Electronica provide an interface with the Arts community and broader public.

Paradiso in his role as a conference organiser, addressed the current generation of this innovative group of practitioners and researchers at the 2002 edition of the New Interfaces for Musical Expression Conference, hosted by MIT Medialab, Dublin 2002.

"The vocabulary in this field is likewise in its infancy – there's still no common set of standards with which to evaluate designs, and as goals are so varied in different applications, its unclear whether this can ever be effectively accomplished. Indeed, the practitioners in this field spring from many walks of life; academic researchers, musical performers and composers, dancers and choreographers, artistic designers, video game developers, interactive and media installation artists, teachers (from university to grammar school), and therapists (special needs, exercise, and relaxation), to name a few."
(Paradiso, 2002 pp.2)

This is still very much the case today. What all these developments have in common whether it is intentional or as a by-product of a creative process or a scientific method, are the conceptual tools to develop a deeper understanding of interaction between people, systems and musical structures. This is an extremely complex process of perception, interaction, reaction, adaptation and creation. The significant factor in this context is not the mastery of an instrument, a virtuoso performance, or design of a sophisticated controller. It is not the broadcast and dissemination of new musical forms through networks, online databases and streaming technologies. It is the potential contribution to knowledge based on observation of the emergent behaviours and interaction models that arise from these disparate fields that give us the framework or 'vocabulary' for understanding the evolution of interaction with the medium of collaborative sound works.

The first step of this thesis is to identify, develop and contextualise new interaction modes, system models and participant behaviours that can inform the design of collaborative music systems. This addresses the first research question. This is achieved by identifying this emerging vocabulary from the available literature. The principle 'accessible to novices yet engaging to musicians or composers' is considered as highly desirable in the field of tangible interfaces, however new design strategies are required for the development of such integrated systems. Indeed this continues to be a core question within this field as expressed by Tod Machover in an eloquent keynote speech for the same conference (NIME-02)

"How do we create interactive situations that stimulate rather than placate, leading the participant beyond the surface and into

thoughtful consideration of rich, expressive, meaningful experiences?"
(Machover, 2002 pp.3)

The second step is to consolidate this emerging vocabulary and apply it to articulate the core features of Adaptive Social Composition. This addresses the second research question. An analysis of recent published works in related fields and an overview of the system design and prototyping of the Orb3 collaborative sound environment, a system of my own design, are presented. Principles established in board games are unpacked. Recent published papers documenting the progress of this research to the international community, presenting initial findings, system designs and conceptual approaches to Adaptive Social Composition are also included (see appendix 3), where relevant extracts from these papers are included in the main text.

The third step is to extend the core features of Adaptive Social Composition using a controlled experiment to identify new participant behaviours. A compositional approach is presented that combines a range of interaction models with an evolving rule set to design of a new digital instrument. The third research question is addressed. This demonstrates the application of the new design strategies, of this thesis, to a novel interface prototype that encapsulates the principle 'easy to learn, difficult to master'.

Chapter 1

Collaborative Sound Environments

The essence of collaboration between individuals and a mediated environment is co-creation, whether this explicit through the manipulation of controllers such as in the early work of Toshio Iwai, or discrete such as in the conceptual scores of John Cage intended for groups of musicians where instructions shape and motivate new interactions dependent on performance context. Collaboration within a musical context can also imply sharing, mutual agreement or tacit referencing of actions and events to generate new material in real-time. This ‘live’ or ‘on the fly’ activity is rewarding for experienced musicians who learn to adapt prior experience and musical exchanges to new situations. These shared activities evolve from a framework of structures and conventions embedded in musicians past experience, and are partly constrained by the choice of instrument or interface. With a novel interface or new musical context, it is harder to identify the underpinning elements that make for a successful or productive collaborative exchange. The notion of a collaborative sound environment is useful to encompass the elements that are combined to make a collaborative, live process possible. Whether this is based on instruments, novel controllers or a mediated environment there is usually some implicit structure that mediates the resulting sound material. This structure could be the acoustics of a room, the qualities of instruments used, the processing chains or events that manipulate them or the behaviour of participants. A useful definition of a collaborative sound environment can be drawn from the fields of Interactive Music, Machine Musicianship, Live Performance Practice or Sound Installation. Of course, within each of these fields, there are numerous examples that are hard to categorise or have elements that overlap or draw on other compositional approaches. We will begin by identifying some of these related research fields which have initiated new approaches, both technical and compositional, in order to establish a satisfactory rationale for our definition. Once we have established this framework, we will further investigate the technical methods and interaction models evident in each approach.

"...While many have also been successful in designing controllers and interactions that "hook" a novice user, even in distracting, high-powered public spaces, few have been able to make such systems "nourishing" as well, capable of encouraging deeper exploration and continued discovery and creativity."
(Machover, 2002 pp.2)

Machovers' keynote address at the international conference 'New Interfaces for Musical Expression' can be understood as a call to the expressive control community to consider the wider compositional framework and human factors inherent in integrated mediated Computer Music systems. One way to achieve this is to investigate diverse examples from the established fields of Computer Music, interaction design and entertainment systems in relation to interactivity and immersion. This notion of 'nourishing' implies a sophisticated form of human computer interaction, missing from many 'interactive' systems, and identifies a need to develop and design compositional strategies to refine and extend the interactive process within social composition systems. Interactive music has developed considerably since Rowe introduced the concept of machine musicianship (Rowe, 1993) whereby a system can manifest musician like characteristics and responses. Within collaborative systems designed for novices, an established musical framework is not necessarily explicit, but the mediating system should still have the capacity to recognise gestures, actions and behaviours within a compositional framework that is accessible and tangible. In a previous paper (Livingstone, 2001 - see appendix 2), one such strategy is articulated within a conceptual framework. In this paper the intention was to establish a conceptual model of interaction that is adaptive, outlining a strategy to motivate this collaborative process and establishing a term to underpin this interaction model, *perceptual construct*, which encapsulates a series of relationships or parameters to define the point at which human- machine collaboration can be identified. This took the form of a simple cognitive task, where participants were asked to visualise eight points in space, relative to their current position. They were then instructed to move these points or coordinates so that they formed a cube. Each participant was then invited to describe their cube, to share what they had created with others, and to note differences in the orientation and cubes of others. The collaborative process is described as a three-stage strategy: *perceptual shifts*, *translation* and *integration.* The model is established as a *tri-part system.* Emphasis was placed on the nature of this *process-driven collaboration* as a mediated experience, where the system manifests agency and the potential for learning. A prototype system

'SoundSpace' was implemented (Livingstone and Swain, 1997) which allowed a user to manipulate sound objects created by the system as *points in space* (Varela, 1991), changing their parameters by repositioning them in relation to other sound objects. The system was able to record the positioning process and map this interaction to relocate other objects within the soundscape, so a simple composition based on capture and sonification of these spatial relationships could be created. This was effective as a proof-of-concept prototype, establishing a means for exploring collaborative interaction and implementing the *perceptual construct* model. In essence, this early development provided the rationale for an interactive compositional framework that could translate a concept into a gesture through a focussed task that resulted in a tangible shared representation. The next stage in this work was to draw on expertise from related disciplines to establish a field vocabulary to contextualise the emergent behaviours and potential collaborative forms from related practices. By exploring the implementation of related systems and the underpinning technologies, design approaches and evaluative language used, numerous fields of practice-based research were revealed.

1.1 Related Research Fields

The range of critical, creative and technical areas than can be drawn upon to inform the development of adaptive social composition systems is broad, traversing both Arts and Science disciplines. The forms of interaction evident in many social processes can be considered in order to refine communications and sound relationships within mediated social composition systems. The fields of Human Computer Interaction (HCI), Psychoacoustics, Ambisonics and Sonic Arts in general all demonstrate approaches that potentially offer both established and new *compositional models* (Lippe, 1996) that help to identify a range of behaviours to refine our perceptual process while interacting through the medium of sound. Within the extensive field of Computer Music, the innovative approaches developed over the past forty years or so have led to numerous compositional and technological innovations, or *new ways to play* (Paradiso, 1997). Technological development has enabled research groups and individuals to establish approaches inspired by biological systems, data transformation and physical sound manipulation that previously would have required bespoke systems, complex orchestration and even specific architectural provision for these works to evolve. This potential for even an individual to achieve a level of complexity in interface, systems or software design which was previously unattainable has in itself lead to the evolution new collaborative practices or *Poly-Media* (Alsop, 2003) that draw on even more disparate fields, many of which are inspired by techniques and simulation processes developed in the sciences.

There have been significant developments and innovation in novel controllers and hybrid instrument design (Tanaka, 2000). These interfaces are intended for different user groups and are used in different contexts. For clarity, various strands of research that are generally considered to be subject domains within the wider field of Computer Music are discussed. These domains overlap; techniques, technologies and activities evident in each domain can be found across the full spectrum of Computer Music and within mediated performance practices. However, it is useful to consider the compositional processes typically manifested within each domain in order to establish core principles for adaptive social composition.

In the following paragraphs theses subsets of Computer Music research are discussed:

Interactive Music
Machine Musicianship
Performer Machine Interaction
Algorithmic & Evolutionary approaches
Emergent Behaviour
Sound Diffusion

1.1.1 Interactive Music

In its simplest form, interactive music (Winkler, 1998) can be described as an *interpretative* process, an approach whereby some human musical activity is identified and initiates a response within the software. This response could be as simple as modifying parameters to vary the performers approach in hearing the modified output, or it could be skilfully programmed to identify nuance of emphasis in a live performance and restructure either the live material itself or generate complimentary sound material itself. Interactive music processes are often an extension of an individual performer's musical practice, in which software is designed to extend the musical domain in some way by reshaping the performance. This emphasis on ‘interpretative’ gives an indication of method. A shared language is essential for the interpretation to be musically significant. In many examples, MIDI and methods for audio analysis have been used to facilitate this exchange allowing a continuous flow of musical information between performer and system with subsequent interventions by both software and performer establishing a tangible musical dialogue. Other approaches include the modification of instruments or the invention of new methods of control. More recently with the development of the Open Sound Control (OSC) (Wright et al, 2003) protocol, it is equally possible to communicate additional data from the performance environment for broadcast and exchange across both physical and virtual locations. But the essential ingredient of interactive music is *co-creation* where a number of processes are combined to generate new material and sustain the compositional process as a live activity.

“Like good conversations, interactive compositions succeed by encouraging spontaneity while residing within the dynamic context that seems whole and engaging”
(Winkler, 1998 pp. 4)

Whether this is a structured performance or an improvisational event, interactive music can model many of the relationships between audience, composers and performers to develop new approaches. Numerous works are developed from the established models found in interactive composition: *performance model, instrument model* and *composition model*. Within the *performance* model, it is often the case that the composer has created a clearly defined role for a performer/instrument and computer and yet it is the dynamic interplay between these 'players' that makes the interactive process within the music tangible to an audience. Live Performance of these interactive compositions is often titled explicitly: 'composition for viola and computer' and so on. In many cases a *reactive model* is evident, the software, typically developed in programming systems such as Max (Website reference 1), LISP (Website reference 2) or PD (Pure Data) (Website reference 3). These programs are designed to identify musical events though standard MIDI messages and audio analysis; beat detection, pitch, velocity in relation to scored material, using sequences or algorithms embedded in the software which are articulated through synthesis of new compositional material. When these works are performed, it is often a virtuoso instrumentalist who is positioned on stage alongside either a laptop or desktop computer, so the audience can experience the musical exchange as the composition unfolds. This process is both visible and audible as the performer's virtuosity is challenged by new compositional elements generated through software interpretation of both the scored material and the live interpretation of it. Although this music is described as interactive, the software itself is often a mediator of the composed material through performer interpretation of a score and software response through carefully crafted synthesis and score-following processes developed by the composer. The computer requires input from the performer to process, to generate a response. As this compositional practice evolves many new approaches have emerged. Some place emphasis on more sophisticated tracking of performer interaction (Merlier,

2003) encouraging new musical gestures (Cammurri et al., 2004) through modified instruments and custom controllers for musical expression (Wessel et al., 2002). What these approaches emphasis is the capture or translation of an event, action or gesture into a performance, the software behind these systems is designed to translate direct gestures encoded as data and does not apply any compositional significance to performer behaviour outside the immediate moment of interaction, following an instrument model of direct control. The potential for co-creation of musical material between performer and system is not fully explored. Others have created ways to emphasise the virtuosity of the software as virtual performer, through live electronics and even robotic instrumentation, as in the case of 'Guitar Bot' (Singer et al., 2003). Further well-established approaches implement virtual instrument/performer communities (Whalley, 2005) or adaptive musical agents (Beyls, 2005) developed from complex agent based modelling (Whally, 2004) and evolutionary or algorithmic composition (Spicer, 2004).

The notion of virtual instruments and even virtual performers is well established within mainstream commercial music software. Ranging from simulated software versions of existing instruments, or entirely new software-only instruments to virtual instrument/performer packages such as 'virtual guitarist', 'virtual bassist' and more recently 'virtual vocalists' 'Leon', 'Lola' and 'Miriam' from Zero G (Website reference 4). These commercial software developments may well provide some additional tools for the individual composer of conventional mainstream or material that is more esoteric but the significant element within this research, is the development of autonomous or adaptive collaborative musical structures within the interactive performance process that enable more complex musical interplay to unfold. Any instrument, be it a traditional string, wood, brass or a virtual software instrument (there are literally thousands of 'VST' plug-ins and instruments available) needs to be played! Whether the performer is an accomplished musician or a software program, practice, rehearsal, improvisation or reprogramming is a requirement to refine the performance. Even commercial sequencing and performance software such as Ableton 'Live' (Website reference 5) are including elements that provide rudimentary algorithmic and generative processes to mediate and reprocess material. What the Computer Music research music community are developing goes far beyond this by implementing systems that can mimic, learn and adapt to human performance, software agents and distributed performance technologies for collaboration (Jorda, 2005). This is significant as it brings us closer to an understanding of ourselves as communicating organisms, exploring our perceptual and cognitive processes in relation to other phenomena. At the same time it is equally important that the tools and structures we create to interact with technology do not limit or define our interaction with each other.

From the perspective of social composition within a mediated system, it is the spontaneous nature of musical exchange that provides a significant design challenge. In mediated systems where collaborative interchange is a key element, the design of rule-based processes to structure the context of a *performance*, the operation of a modified *instrument* and the structure of a *composition* form the foundation of many interactive music works. Recent manifestations of the form include live musicians performing with computer controlled musical robots (Singer et al., 2005). The development of such systems are well documented, with regular performances appearing on the International circuit at conferences such as New Interfaces for Musical Expression (NIME) and the International Computer Music Conference, (ICMC).

The domain of interactive music and the associated software, techniques, interfaces and compositional approaches have had a significant impact on many interactive works that integrate people, spaces and computers to create new musical forms. Many of these works are *interpretative* that is to say a common language or protocol, often MIDI, is used to map and transform interaction, or dynamically control musical data transformation between people and systems. Established platforms such as *MaxMSP/Jitter* (Website reference 1), *Pure Data* (Website reference 2), *Big Eye* (Website reference 6), *Eyesweb* (Website reference 7), *Eyecon* (Website reference 8), *Processing* (Website reference 9), as well as mainstream media applications enable more complex software mediated approaches. These approaches can incorporate video tracking, gesture capture, collaborative networks (Weinberg, 2002) and shared databases (De Jong, 2005) and can use a range of additional development tools with integration through the Open Sound Control (OSC) Protocol (Wright et al. 2003) for example. It is at this point in a system design that the process of communication can become subjective; a deeper understanding of the relationship between these elements needs to be established.

"Physical parameters can be appropriately mapped to musical parameters, such as weight to density or register, tension to dissonance, or physical space to simulated acoustical space, although such simple one to one correspondences are not always musically successful. The composers job then is to not only map movement data to musical parameters, but to interpret these numbers with software to produce musically satisfying results"
(Winkler, 1998 pp.320)

These approaches are significant in their own right but fall short in terms of social or collaborative live musical exchange because they cannot adapt to the behaviour or actions of novices beyond a control model. Generally, the method of video tracking or gesture capture encodes an action in context of a prescribed set of musical relationships, and places the onus on the composer to provide a meaningful framework

Key terms; characteristics of interactive music and core models.
Reactive, Interpretive, Conversational, Co-creation.
Instrument model, Performance model, Compositional model.

1.1.2 Machine Musicianship

Robert Rowe's seminal text provides an excellent introduction to this core aspect of Computer Music composition (Rowe, 1993). Rowe's classification articulates a player paradigm for interactive music systems, that is, the system should have player-like skills and responses, it should be able to both analyse and respond to both encoded musical material via MIDI through *score-following, audio analysis* and *transformation* of performance elements. Such systems should also have the capacity to interpret *gestural language* in a musically relevant context. Rowe details a number of interactive music systems *Cypher* by Rowe himself (Rowe 2003), *Natural Selection* (Campion 2002), for large scale electronics and MIDI piano *Izquierda e Derecha* (Kimura M. 1998), for MIDI violin and MIDI piano. These works primarily fall into the 'pitch-to-MIDI' domain for gesture analysis, and have contributed hugely to the development of a tangible communication process between players and systems. Recent technologies have enabled a higher resolution of gesture capture through audio and complex sensor data analysis and video tracking in addition to the MIDI performance and audio processing used in these earlier systems, but the core concept of embedding player characteristics and behaviours with a degree of autonomy and musicianship remains. Recent research developing works extending the player model within interactive music systems shows that detailed audio analysis continues to be a core strategy in mediating performance. William Hsu (2005) has developed such a system with saxophonist John Butcher and provides a detailed account of their strategy for capturing detailed gestural control through timbral gestures, post processing and categorization relative to the timbral characteristics of the acoustic instrument; *Noisiness*, *Prominence, Presence of sharp attacks* and so on. This system extends the player paradigm through the inclusion of software agents with the capacity to improvise independently of each other but in relation to the collaborative material performed.

Machine musicianship offers significant methods for applying higher-level musical interpretation by software or software agents for collaboration between accomplished musicians and software based musical entities. These approaches tend to be focused to an individual musician and specified instrument. Usually these system qualities are developed over a significant period of time with extended collaboration between a performer/musician and systems creator/composer. Machine musicianship is not normally a characteristic of tangible interfaces intended for novices or collaborative groups. This is primarily because these approaches follow simple control, sampling or sequencing paradigms. However, musicianship is a highly desirable characteristic to include, if the objective is to deliver engaging musical contexts. Within a tangible interface, an interpretive software process can be used to extend the expressive potential of an interface beyond conventional parameter mappings.

Key terms: machine musicianship methods and features.
Score-following, Audio Analysis, *Transformation, Gestural language.*

1.1.3 Performer Machine interaction

There are at least two distinct approaches in the interface design of performer-machine interaction. One can be described as the extended or 'hyper–instrument' approach, the other is often an unencumbered interface using motion capture or a gesture-based system. In the extended or hyper- instrument paradigm, the design is usually based around a conventional *instrument model*, adding sensors or idiosyncratic controllers to extend the instrument. This allows the performer to extend the range of sound material that the instrument can produce and the tangible control by extending the sensor measurement and benchmarking for interaction through *gestural control* (Wanderley and Battier, 2000). The most significant examples in this area of hybrid instruments also articulate performer mannerism and behaviour to extend *performance expressivity*. The Overtone Violin, (Overholt, 2005) takes this field one stage further. Rather than modifying an existing violin, the emphasis is placed on understanding the acoustic principles of the original instrument and applying the same level of detail and refinement to the design and integration of sensors for live performance and expressive control. This in turn led to the design of a highly resolved six string instrument with an integrated system for signal processing for individual strings and gestural control that extends the *gestural vocabulary* (Mathews, 1984) of accomplished players of the original instrument. Playing skills and learnt behaviours for manipulating strings are mapped to the layout and position of sensors on the headstock of the instrument so players can learn and explore new sound-manipulation processes intuitively. Similarly a sensing glove is used while bowing to enable the subtlety and control exercised over the analogue sound to be harnessed for signal processing. Of course, this approach encourages new performance behaviour through the dynamic expressivity and versatile yet familiar interface that the *Overtone* instrument provides. This is considered in terms of instrumental gesture, later in this chapter. A perceived limitation of this approach is that although highly evolved technically the musician

(fig 1. *Overtone* Violin by Overholt)

"Trained violinists are able to pick up the Overtone Violin and play the strings fluently. However, there is another gesture vocabulary beyond that of acoustic violins in dealing with the extra sensors that requires the development of new skills to master. While this necessitates new playing techniques, the process of learning is facilitated by similarities to the older technique."
(Overholt, 2005 pp.35)

is constrained by prescribed relationships that can only be adjusted during performance by stopping playing. For example ultrasound sensors are used to map the distance between the players bowing hand and the receiver embedded in the instrument, the sensor is worn in a glove, while the adjustment of synthesis parameters is controlled by rotary pots embedded in the headstock. This means that while the action of the bow on the strings can be varied in the traditional way, the additional layers of synthesis and signal processing cannot be adjusted without interrupting the performance. This example digital instrument is a significant feat of engineering, which embeds features that could be housed in conventional rack equipment with foot pedals into the instrument itself, with bespoke onboard processing so no external computer is required. It features sophisticated gesture acquisition but lacks any mediating software that exhibits machine musicianship. It may be intuitive for violinists to learn in terms of instrument gestures but the extended features and resulting sound material generated are idiosyncratic and obscure.

The unencumbered approach is more often found in collaborative performance or installation situations (Cammuri, 1995). Typically these systems use forms of collision detection, pattern recognition or customizable video analysis methods via a live video feed in order to map a hand gesture or full body movement to an event such as sound synthesis, video control or lighting effect. There are numerous research centres and performance collectives developing and extending existing software for implementing such systems; Steim; *Big Eye* (Website reference 6), Palindrome; *Eyecon* (Website reference 8), Cycling74 *Jitter* (Website reference 1) and of course the Open Source *Eyesweb* platform (Website reference 7). There are also numerous externals listed in the 'max objects database' (Website reference 10) for the *Max/MSP/Jitter* environment such as *Taptools* by Timothy Place (Website reference 11) and *CV.jit* by Jean-Mark Pelletier (Website reference 12) or extras for mainstream commercial Multimedia Authoring environments such as *Trackthemcolors* for Director MX (Website reference 13) and the

Java multimedia API. Each system is customizable but there are fundamental principles that determine the forms of interaction they are best suited to respond to. For example, Steim's original *BigEye* software provided a number of tools for mapping rectangles onto the video stream, enabling different events to be triggered via MIDI output when certain conditions were met. Palindrome's *Eyecon* is more live-dancer oriented, using a fixed camera. The two main methods used allow for the outside or 'bounding box' space inhabited by a moving individual to be mapped to an event or series of event triggers, usually via MIDI messages. Similarly, lines and other shapes can be drawn over the video stream so that when a moving element intersects with a drawn element further actions can be triggered by the software. Jitter allows customized bespoke software to be designed. From the simplest blob or colour tracking to complex video analysis methods such as optical flow, using a wealth of externals and third party packages is available. *Eyesweb* offers a sophisticated toolset for performers and researchers *Gesture processing*: *Space and Motion Analysis* libraries, which are in continual development.

"The extracted measures can be used as input for clustering algorithms in order to group trajectories having similar features. In the real space, this approach can be used to identify points moving in a similar way (e.g., points associated to the same limb in the case of the Lucas-Kanade feature tracker). In a semantic space, it could allow grouping similar gestures, eg; gestures communicating the same expressive intention."
(Cammuri et al. 2004 pp.6)

Significantly these systems are designed to go well beyond the standard high tech motion-capture systems by providing predictive and analytical tools that can be used to develop semantic and conceptual relationships between seen and predicted events. This approach also has the potential to form new live practices that motivate and identify emergent behaviour as a collaborative process (Moroni et al. 2000). We have already outlined the principles used in interactive music and considered the nature and strategies of machine musicianship. Simple multiple camera systems working at relatively low resolutions have also contributed significantly to these approaches, for example, a more abstracted application of these methods can be seen in Rokeby's *Very Nervous System* or *VNS* (Winkler 1998) where two 6 bit video feeds of 128 x 240 pixels are mapped against user-specified grids and light levels can be compared.

"The analysis and interpretation of movement data holds great promise for interactive composition, dance, and creating responsive music in virtual reality systems."
(Winkler, 1998 pp.314)

Current developments in the performer-machine gestural interaction field address the potential system limitations and the balance of aesthetic issues of non-tangible capture systems. These approaches allow unencumbered control of virtual instruments with many degrees of freedom through new strategies for mapping a *multiplicity of parameters* (Dobrian, Bevilacqua, 2003)

A significant element of the performer machine approach is that sophisticated software and software extensions have been developed to provide numerous ways of analysing physical actions or performer movement dynamics. A limitation is that the linkage between actions and musical events or parameter mappings or tailored to a given performer, instrument or environment. In many cases, the freedom to map a multiplicity of parameters does not lead to lead to musicianship and has to be carefully orchestrated. This subset of the field produces powerful software tools for gesture acquisition that can be incorporated with the principles of musicianship to build mediated compositional systems

Key terms: Performer machine interaction.
Gestural control, Gestural vocabulary, Performance expressivity.

1.1.4 Algorithmic and Generative Approaches

These approaches are highly significant within adaptive composition systems that include software mediation of human interaction and a degree of autonomy within the compositional process. Gestural interaction design can extend the player-instrument model by applying principles of co-creation to facilitate social interaction. Musical information is easier to manipulate with algebra than procedural code (Goggins, 1996) allowing an initial sequence of notes or sounds to be the seed for evolving new material, as is evident within a wide range of compositions and compositional tools that utilize generative granular synthesis, genetic algorithms and cellular automata.

"We are particularly interested in those cellular automata that display cyclic behaviour, self-organisation and/or pattern propagation properties."
(Miranda, 2001 pp.137)

Researchers in this domain are developing new interactions between virtual performers/players using Artificial Life methodologies to implement agents and agency through modelling biological systems. These systems in turn allow for the evolution of social and cultural phenomenon, closing the gap between human perception, cognition and musical behaviours by modelling social interactions within a shared musical context and observing the resulting evolution.

"The motivation of the agents is to form a repertoire of tunes in their memories and foster social bonding. In order to be sociable, an agent must form a repertoire that is similar to the repertoire of its peers. Sociability is therefore assessed in terms of the similarity of the agents' repertoires.
(Miranda, 2001 pp.150)
Sociability in these cases applies to peer groups of virtual agents but the underpinning model of establishing a shared process or repertoire, as a motivator for interaction is equally applicable to human/interface/software interaction. Other approaches in this field explore autonomous generative systems that can also be interacted with or influenced through sensors and real-time control via

audience, performer or composer gestures (Moroni et al. 2000).

Current research emphasizes the significance of *hybrid approaches* to non-linear generative improvisation (Whalley 2005) that utilizes a *conversational model* (Paine, 2002) to human-computer interaction.

"What is lacking in the application of agent technology to music/sound is a means to balance the interests of a conversational model of human computer interaction with a model of music/sound as a language that communicates affectively; and a common platform for the distribution of works, such as interactive gaming that may allow ideas to be tested and integrated wider communities outside the academy or media art circles."
(Whalley 2005 p693)

Within a compositional context, these developments can contribute significantly to the development of collaborative music that is mediated by software agents. Indeed the study of musical language through the development of agent-based musical communities is at the forefront of this field (Miranda, 2005). Other significant contributions to this field include the work of Blackwell (Blackwell, 2004) who has investigated free improvisation through *indirect* interaction between Swarms and humans. This body of work establishes a real-time interactive music-making approach with computers, as opposed to the many systems that provide accompaniment or adjust and modify existing material. Blackwell emphasizes the significance of *free improvisation* and the *self-organising* nature within these systems. Of significance is that this approach provides a more open interaction framework. This enables sophisticated interaction between humans and virtual entities, through collaborative musical behaviour.

"Since improvised music operates without any prior agreement on tonality, pulse or dynamics, the question arises as to what the musicians actually respond to. One answer is that the musicians are attentive to the expressive qualities of their musical environment. Expressive qualities are high-level descriptions of the music, including attributes such as event density, average loudness and pitch. The participants interact by either trying to match the expressive quality of the musical environment, or by attempting to change it in some way."
(Blackwell, 2004 pp.137)

With an adaptive social composition approach, it is clear that any design strategy should consider the Human Computer Interface, if higher-level expressive interchange is desired. A series of design strategies should be resolved that effectively motivate a collaborative process between humans and a real-time compositional process. Genetic or algorithmic approaches are ideally suited to providing adaptive elements within Computer Music software. Can these processes be applied to develop novel interfaces with dynamic parameter mapping? For effective free improvisation between humans and virtual agents, a tangible method of interacting within a real-time process is essential. If the compositional process is collaborative, a social musical environment with sensors for exchange of expressive data between humans and virtual agents is a valid design approach. A collaborative approach of co-creation is significant if we want to experience the musical process and contribute to it in a reciprocal context. Allowing audiences to engage directly with these processes has the potential to motivate new forms of creativity without the constraints of previous models. Similarly, mediated systems cannot learn from abstract data without a shared language to transform events into meaningful dialogue.

Key terms: algorithmic and evolutionary characteristics.
Self-Organising, Sociable, Reciprocal.

1.1.5 Emergent Behaviour

Across the broad spectrum of algorithmic approaches, the term *emergent behaviour* is used to describe interactions between cells or agents that are not predetermined. The term, *emergent behaviour* has been adopted by a wide range of practices to describe the unscripted interactions that arise between people, interfaces and software environments. A number of fields of practice consider *emergent behaviour* to be a key element in realising new creative forms. Creators across these fields manifest compositional processes, immersive environments and novel interface design. Technical approaches include video tracking systems, human computer interaction, interactive and generative music and collaborative soundscapes. The potential for new interdisciplinary forms integrating gesture capture, motion tracking, sound synthesis and collaborative forms between people/performers/composers and their environments is a developing field of research that investigates process driven collaboration to inform the design of reactive compositional spaces.

Process driven collaboration (Livingstone, 2000) can be described as an embedded strategy that instigates a shared goal to stimulate interaction or participation. This can be between performers and their instruments, composers and sound material or participants and technologically mediated experiences. Increasingly these embedded strategies can be found at the software layer of interactive or compositional systems, for example an Algorithmic approach extending the potential for both the generation of new sound relationships where the dynamics of the environment or performance are directly affected by participants of the system and the system is perceived to be responsive, indeed across a range of fields of practice this 'responsiveness' has been identified and extended, leading to a number of ways of describing emergent behavior. Where once we would have described interaction between users and systems with a clear hierarchy implicit in the language used, we now find these relationships have evolved, in part due to the increased use of embedded strategies to facilitate real time compositional processes in response to interaction. These forms of collaboration between participants and systems in many cases lead to new forms of behavior being identified as an extension of the creative potential of both parties, this language of behavior is playing a key role in the development of new interdisciplinary collaborative processes.

As practitioners and researchers from differing fields exchange expertise and approaches, new possibilities come into focus and a deeper understanding of the language of interdisciplinary work is reached. Technologically mediated relationships can be very effective across a range of resolutions, for example an interaction as simple as moving or clicking a mouse forms the primary act of interaction most people have with computer technology, but clearly the resolution of this act is determined by the sophistication of the interpretation of the act. Both on the part of the user, and in terms of 'expectation' or 'anticipation' of the system or software design; the resolution of the act is multiplied by the understanding of the range of anticipated or implied behaviour, so any system that multiplies the resolution to extend the language of reciprocal engagement with a context or process embedded within the work has the potential to manifest emergent behaviour. Koert van Mensvoort of the Eindhoven University of Technology has developed an 'active cursor' method for simulating haptic feedback:

"The position of the cursor channel is normally used for input only. We developed a cursor interface in which the system manipulates the cursor position to give feedback to the user. The user still has main control over the cursor movements, but the system is allowed to apply tiny displacements to the cursor position. This system has a lot in common with existing force-feedback systems, except for the fact that in force-feedback systems the location of the cursor is manipulated as a result of the force sent to the haptic display, whereas in our system the cursor location is directly manipulated."
(Mensvoort, K. 2002)

The key point here from the perspective of interdisciplinary practice is the increase in resolution of information possible from one human computer exchange by identifying the wider context of a process driven act. As researchers in the field of Human Computer Interaction (HCI) increase the possible range of reciprocal interaction with feedback processes simulating tactile sensations through visual stimulus, these methods can be added to the possible language of behaviors that can drawn on in the design of interactive environments and novel controllers.

There is a clear potential here when we begin to describe mouse movement as gestures, as nonverbal language but there are also significant implications on how a system is programmed to react or 'learn' from this, establishing a process of collaboration or dialogue. In many areas of practice direct manipulation of the media or processes inherent in a system is not a key requirement, the system or piece has been resolved and an increase in resolution of the reciprocal cycle is achieved by a number of means. A low-tech but nevertheless engaging approach can be seen in the early collaboration between Woolf and Beck whose approach integrates sound sculpture within 'reactive' robots that appear to display a range of autonomous behaviour. Simple analogue sensors and control circuits are used to extend the interface of a system to allow intuitive interaction to take place.

"Despite its simplicity, Echidna exhibited a large range of interesting sonic behaviours. This behaviour reflects not the sophistication of the underlying electronics, but the complexity of the environment in

which the sculpture is situated."
(Woolf, S. and Beck, T. 2002)

During their research the observed that audiences attributed behaviours to the robotic sculptures

"Despite the simplicity of its control circuitry, Boundless appears to display complex autonomous behavior. If approached by an observer it will attempt to withdraw, as if trying to flee from a perceived predator. If approached by several people from more than one side, Boundless jitters indecisively, as if unsure of which way to turn.
(Woolf, S. and Beck, T. 2002)

It is perfectly reasonable to counter this observation by suggesting that participants ascribe interpretation to perceived actions and react accordingly but if these non-explicit modes of interaction are more clearly understood then the potential for sophisticated compositional and collaborative processes within reactive environments becomes a realistic proposition. Just as the designers of screen based interaction are developing subtle but sophisticated visual feedback systems to enhance immersion through representation of tactile, physical properties within a software environment, creators of collaborative sound environments have used gesture capture and motion tracking to enhance a systems reactive properties, to both participants and environmental parameters.

These two examples show how the concept of *resolution* is highly significant to the design of novel controllers and mediated musical environment. A simple interaction can lead to emergent behaviour, characteristics of emergence can be described in interdisciplinary works for public exhibition. Audience response can be stimulated by interfaces with autonomous behaviours, extending engagement.

Key terms: Audience engagement and response.
Emergent behaviour, Process driven collaboration, Resolution.

1.1.6 Sound Diffusion

There is a wide range of standards for the recording, reproduction and spatialisation of sound, and these formats are well established and thoroughly documented elsewhere. Likewise, specialist ambisonics software (Website reference 16) and bespoke environments for performance are similarly well established. The electroacoustic performance paradigm is also well documented with a composition being diffused in real-time from prepared material through the use of standard or customized control surfaces using faders and rotary pots for volume, panning and EQ control for live playback or mixing as performance. A recent development in this field is to develop software that learns the composition or performance behaviour of the composer. (Melo et al. 2005). Recently, specialist control surfaces such as the LEMUR (Website reference 14) offer a highly flexible dynamic touch-screen control-surface for real-time performance and control with sophisticated graphical representation of object properties/physics that can be mapped to any parameter to extend gestural control. These new touch-screen controllers have the potential to engage non-musicians in musical behaviour through tactile control and visualization of performance parameters. Dynamic Acoustic feedback is extremely significant both in sustaining collaboration and extending engagement between human listeners and a mediated compositional process. More complex or real-time sound diffusion allows for higher resolution of feedback, for example using positional sound to indicate a relationship or motivate an interaction. Similarly, the complex spatialisation of sound can change the way we perceive it. Current graphical interfaces emphasise the potential of direct trajectory control within spatial works for complex parameter mapping that extends musical language.

"Depending on the strategy, trajectory can be a parameter generator for a synthesis approach or for a transformational approach. This leads us to think about the importance of the interface in electronic music and of the role of the mapping of abstract data toward aesthetical intentions..."
(Thiebaut, 2004 p23)

The use of haptic control of sonification models (Hermann et al., 2002) has been documented within exploratory data analysis to extend the way we way interact with data through new perceptual representations forming a tight kinaesthetic linkage between actions and sounds for dynamic visualization. This approach to sound diffusion can also be embedded in compositional systems as a strategy for motivating social interactions between human performers within an agent-mediated soundscape.

Many public sound artworks use sound diffusion to immerse visitors effectively, and a number of these directly encourage visitors to *play intercept, change or make contributions* (Bandt, 2004) and facilitate collaborative interaction through sonification of movement or gestural interaction. A logical development of these works is to generate strategies for real-time diffusion that mediate sound relative to the participant action or location. Within ubiquitous technology networks, new models of *proximal interaction* are being explored to enhance musical experience (Tanaka et al., 2004).

To establish adaptive social composition as a shared framework of novel interfaces, sound diffusion and relative positioning of participants can be combined. By tracking performer or player locations and associating their contribution within a shared composition, it is possible to orient learners and interpret individual actions. These actions can be interpreted in different ways, motivating a range of musical behaviours. This provides a compositional context for mapping the multiplicity of parameters more typical of performer machine systems. Actions or gestures can be sonified within a shared context, introducing musical relationships and revealing alternate combinations for novices to explore.

Key terms: sound diffusion features for participant engagement. *Proximal interaction.* Participant behaviours; *play intercept, change, contribute.*

1.1.7 Summary (section 1.1)

This section presented an overview of a subset of areas within the diverse field of Computer Music, identifying established research in various complimentary areas related to novel interfaces for musical expression and collaborative sound environments. The notion of expressivity is introduced with examples provided to show different approaches for engaging a performer or participant within a musical context. The motivational role of sound diffusion in providing feedback for actions between participants and interactive works or collaborative sound environments is also considered.

The next section provides a review of technologies used across these complimentary fields with attention to classes of gesture and control that form an underpinning framework for adaptive social composition. Terms introduced in this section are expanded upon using selected examples.

1.2 Interaction Models and Technologies

Interaction Design has a very significant role to play in the creation of new gestural interfaces for technology-based sound practices and software mediated collaborative forms. Performers and designers of these works re-invent and develop the way we interact with technology both physically and perceptually. This, in turn, leads to new forms of interactivity that are content and process-based, i.e. not limited to a specific controller, function or interface navigation metaphor. Interactivity in the context of contemporary Computer Music extends far beyond the established frameworks of interface, software or emulated instrument design. Other significant factors are evident to varying degrees across the full spectrum of current works and *hybrid* or *polymedia* (Alsop, 2003) practices in this field. New mediated or collaborative compositional approaches engage a variety of strategies to create and orchestrate musical material (Impett, 1999). Investigating the underpinning interaction models (Livingstone and O'Shea 2005 – see appendix 2) and their associated technologies can give us a better understanding of the cognitive processes (Cook 2001) and individual or group behaviours (Timmermans et al., 2004) that are revealed by these approaches. A representative selection of tangible or novel interfaces was selected for critique, to identify the constraints placed on the intended users of these systems. The common feature of each of these systems is that they provide an iterative musical process, inviting users to place or change the relationship between objects to modify an existing soundscape. In some cases, the objects had explicit functions that could be learnt by a participant. For example *Soundgarten* (Woolf 2002) offers a toy like collection floor based building process, combining elements of a single object, enabling younger children to record, modify and arrange samples through exploration and live sampling of their own sounds by placing or plugging elements together to build their own logical relationships. This work prioritises exploration, new functions are revealed through trial and error, *exploratory* interaction is rewarded. *ISS Cube* (Quarta, 2003) functions as a collaborative table top spatial mixing surface for up to four participants using simple movement of tactile objects to diffuse sounds. This interactive surround sound environment builds on a more strategic model of play and exchange with intuitive interaction based on relative movement and location of generic small discs, reminiscent of many board games and intuitive to use with collaborative 'positioning' of predefined sound samples. In this example the interaction mode is primarily *organisational,* participants can effectively remix ambient environment samples to reconfigure a diffused soundscape but the source material is fixed, with minor adjustments to sample parameters. *Audiopad* (Patten and Brecht, 2002) is a well-documented work in the audiovisual tactile mixer field offering real-time visual feedback in addition to a tangible control interface with tactile elements. This tangible interface comprises a system of colour-coded pucks with back projected graphics to illustrate active functions on an opaque tabletop surface. Again despite the informative display method the range of actions and processes are those achievable with a conventional music sequencer with prepared material, initially engaging for novices and intuitive to use but the fundamental interaction is to reorganise this prepared material collaboratively. *RGB Player* (Barter, 2004) allows manipulation of sequence and pattern, either collaborative or turn based placing/removal of coloured objects on a rotating table with overhead video tracking. Participants are invited to select and place a variety of different objects, software tracking allocates a sound sample to each placed object and its relative location determines the sequence of playing. *Block Jam* (Newton-Dunn et al, 2003) combines an element of building or assembly to construct a sequence, pattern variation and control of audio using highly engineered interconnecting blocks, reminiscent of dominoes. Both these examples prioritise sequential interaction. A paper describing the core features of each system is included in appendix 2 (Livingstone and O'shea 2005). Each work is considered in terms of the dominant interaction model and the social context that mediates participant behaviours. A more advanced approach can be seen in the design of *ReacTable* (Jorda et al, 2003). Originally, this system used tangible objects with differing textural qualities and shapes illustrative of the synthesis functions they controlled. However, during development these tactile user identifiable elements were found to be difficult to track so topological markers on the underside of generic Perspex pucks, similar to those used in *Audiopad*, were adopted. The improved tracking method enabled more complex actions: rotation, speed of movement and orientation to be combined to trigger different types of synthesis. The associated function is back-projected around the controlling puck with dynamic visual effects as feedback to the user. *Reactable* is described as a Tangible interface for collaborative performance, indeed several *Reactables* can be networked so that groups of players can interact remotely. A notable real-time performance of this system was included at ICMC 2005, using two *Reactables* one located in Barcelona and one in Linz at the ARS Electronica festival. This example exhibits a *relational* model of interaction. Predefined functions are associated with a collection of hand placed and manipulated pucks, proximity of differing functions reveals new relationships between objects. These proximal relationships generate new sound material, primarily through simple synthesis and processing chains. Although providing an engaging spectacle the effect of the placement of pucks and their associated function is not known until they are placed, and the resulting sounds or visual feedback revealed. So, although the system supports real time collaborative exploration there is no mediating framework to compliment the process. Furthermore, during the scored performance featured at ICMC 2005 the player number was reduced to two on each table. In demonstrations of *Reactable,* it was notable that when more than two users placed and manipulated interface elements it quickly became very complex to inter-relate the multiple visual representations and the resultant sonification, motivating *relational* and *exploratory* interactions. This has been complimented by inviting composers to score pieces for performers to recreate or interpret, proving the versatility of such tangible systems, effectively the composer takes a mediating role, establishing a framework for musically significant interactions.

Each example discussed above is intended for a different social context, for example; group discovery, individual or turn based interaction, collective play or collaboration performance. By observing participant behaviour, it was possible to determine the main interaction model each system reinforced, these are summarised as *control*, *exploratory*, *sequential*, *organizational* and *relational.* These simple models can be used to signify groups of related system events and participant behaviours; as such, the terms are useful to focus technical development within a compositional context. These terms are proposed and introduced as part of the field vocabulary identified in this thesis and are a significant factor in establishing effective design strategies for Adaptive Social Composition. These examples, like many Tangible interfaces, are intended for novices, however as the more successful examples are more widely used it is only reasonable to expect a higher level of expressivity if players are to be more than pawns in a scored performance. The compositional process of such systems should be embedded, either in software or at least in the developmental process. Design strategies that consider potential interactions, gestures and responses are required to develop these compositional processes.

Key terms; identified interaction models within tangible interfaces:
Control, exploratory, *sequential*, *organizational*, *relational, transformative.*

1.2.1 Gesture – *Direct* and *Indirect* control

Let us reflect on one of the most accessible and easy to learn forms of communication, that of a simple gesture. A series of subtle or direct movements that reinforce or substitute language, or as described by Benjamin the 'wilful expression of thoughts and feeling through visible bodily action' (Benjamin, 2002 p1). Gesture can be considered as a physical embodiment of thought if the context for the gesture is understood. If an intended gesture can be understood as a physical mapping of a cognitive process, then it follows that capturing that gesture in context enables effective communication. Of course, any musician will confirm that this is obvious. If classically trained they will have no doubt experienced gestures from a conductor or co-performer to indicate or share changes in the tempo, timing and dynamics of a performance, but they may not have considered that every conscious psychomotor control decision they make to control their instrument can also be considered as a learnt or adapted gesture. Furthermore, many musicians when fully engrossed in playing their instrument often manifest a whole new body language of performance mannerisms that an audience or co-performer can respond to, but that a conventional instrument cannot. For example, consider a jazz vocalist moving expressively and singing across a broad range of frequencies and volumes to reinforce or emphasise emotive elements within a song while also controlling her body to best perform her vocal technique. A fixed microphone would not be capable of adapting to the continual shifts in volume alone without compromising the resulting amplified or recorded vocal. Under these conditions, the humble microphone becomes a simple instrument controlled by gesture, with the performer mediating the distance and orientation of the microphone through arm and hand movements. Hewitt and Stevenson (2003) have developed such a system, modelling this behaviour to create new compositional and performance software. The same applies to groups of performers improvising, adapting their playing style to create space for each other to collaborate under the call and response model (Lippe, 2000). Designing software that recognizes conscious gestures or combinations of gestures and remaps these movements to parameters that modify and transform a compositional or performance process (Modler et al., 2003) is as valid a form of instrument design as constructing a physical instrument. The key advantage of capturing gestures with software is that the program can be created to learn and adapt to these movements, whereas with a physical instrument it is the player who learns and adapts his gestures to create new musical material. With a traditional instrument, the instrument is more than passive interface. The performer is in control, and potentially limited to the inherent range of physical 'interventions' that the instrument is designed to respond to, plucking, blowing, hitting, bowing etc. The context of performance and the nature of an instrument varies the level of control a musician has, a keyed instrument has more constraints than a stringed one, a bowed instrument introduces further complexity and subtlety of expression. Musicians adapt their playing to the performance context. Mastery of an instrument of this nature enables extremely sophisticated and complex musical material to be performed, and there are well-established musical forms that modify and extend traditional instruments to create new musical forms. Interactive Music where software processes are triggered by or respond to these instruments, and tracked through MIDI or audio analysis, are normally functions within the structural context of a musical performance. Using gesture capture provides a further level of expression (Wanderley and Orio, 2001), an additional layer of language that provides the potential to capture both conscious and subconscious interaction, or learnt and emergent behaviours.

"Since it seems clear that many forms of gestural expression are elaborated as modulations of practical actions, but since, nevertheless, such gestural expressions may be part and parcel of

utterances that express even the most abstract ideas, this seems to give support to the view that the 'tools of thought' — conceptual metaphor — are grounded in bodily action in the physical world." (Benjamin, 2001 pp.3)

Gestural expression can be seen as a direct action linking the cognitive process of a performer with the physical or musical world. If this linkage can be integrated effectively in a form that a computer can interpret and respond to then a more efficient interface between thought and technology can be developed.
Gesture can be captured using a range of methods, from worn electronics such as data gloves with sensors for orientation, acceleration and bend sensing to un-encumbered video-based tracking, using pattern recognition methods or spatio-temporal appearance modeling (Zhu et al., 2000). Sophisticated models have been created to identify gestures using a range of technologies, these 'gesture libraries' can be added to software designs effectively to capture physical actions, but the nuance of a performers body language or the sounds produced and perceived by the performer go beyond what current solutions can usefully communicate. A useful classification and terminology for these methods developed by Depalle and colleagues (Depalle et al., 1997) at IRCAM (Institut de Recherche et Coordination Acoustique/Musique), Pompidou Centre, Paris, provides a framework for deciding on appropriate methods for composers seeking to apply these techniques. The key terms for gesture acquisition are introduced in the remainder of this section. These terms identify and explain alternate capture methods. These differences are significant when developing design strategies that

combine a number of these methods. Combining a number of methods is desirable when more than one interaction model is used in a mediated system. This enables different classes of gesture to be identified by software in different contexts that are co-created during a performance. In more advanced systems it is usual for combinations of these approaches to be combined so it is important to differentiate between them if one is to fully understand the context or potential application of a given gesture. The following terms are introduced by describing well-documented examples that use higher-level combinations of these approaches:

Instrumental gesture
Non-instrumental gesture
Haptic sensing
Non-haptic sensing
Direct gesture acquisition
In-direct gesture acquisition

Instrumental Gesture: this approach uses a sensor or collection of sensors attached to the instrument or embedded in a custom-built instrument to generate discrete or continuous values in response to the sensed variable. The sensors used can be categorized as *Haptic* such as a simple switch or potentiometer such as is found on an electric guitar, bend sensors as used in many novel controllers or pressure/touch sensors such as the beautifully crafted touch-based interface instruments crafted by Don Buchla; the *Thunder*, *Lightning* and *Mirimba Lumina* (Buchla, 2005). In this instance, sensors categorized as *Haptic* are typically variable resistors used to mediate parameters through the sense of touch providing a tactile interface.

(fig. 2 Thunder Instrument/Interface by Buchla)

Sensors categorised as *Non-haptic* are typically used to capture the relative position, orientation or motion of the instrument during playing such as Jonathon Impett's *Meta Trumpet*, which utilizes ultrasound for directional tracking of the relative position of the instrument during performance. This is a conventional trumpet modified with a range of sensors to integrate the original instrument and performance mannerisms with a sophisticated Max/MSP based interactive music environment. We have already introduced an excellent example of a hybrid as opposed to a retrofitted instrument, building on the design of a traditional instrument while combining both haptic and non-haptic sensing to integrate a sophisticated interface controller. Dan Overholt's *Overtone Violin* (Overholt, 2005) In addition to a range of tactile sensors for haptic control the instrument features a 2D accelerometer, 2 channels of Sonar and a video camera integrated with an on-board microprocessor. The *Overtone Violin* also incorporates an instrument-specific data-glove, where the sonar transducer continuously measures the distance between the player's bow hand and the head of the violin, a similar method as used by the jazz singer mediating a hand-held microphone discussed earlier but applied to capture bowing gestures in conjunction with the accelerometer. Each of these examples uses techniques for *direct gestural acquisition* as actual physical gesture or movement is recorded in addition to the sound generated by the instrument. *Indirect gestural acquisition* describes an approach where a gesture is detected through a detailed understanding of the instrument's material properties and the subsequent audio analysis of the sound produced i.e. the gestures a performer used to make specific sounds with an instrument can be identified by recognizing a repetition of that sound. This method is more processor-intensive, due to the speed at which the analysis needs to take place, but algorithms can be used to optimize this process and the potential for more sophisticated learning or adaptation within the software environment is possible. Impett's *Meta Trumpet* combines *direct* and *indirect gestural acquisition* (Impett, 1994). The differences between the Overtone violin and the Meta Trumpet are worth further consideration. The Overtone violin is built on an instrument that already has a high degree of expressivity and a complex control system embedded in the way the conventional instrument is played. To add further complexity and layer additional direct gesture sensing allows the interface to produce more sounds and layer acoustic and synthesized material. An engaging performance has to be experienced first hand to appreciate this interplay. There is also the potential that arbitrary performer actions can lead to a cacophony of uncontrolled sound where the performers skill is perhaps to retrieve a former balance. This process cannot really be described in terms of interactive music, nor is it a significant compositional approach. It is a significant technological development but would benefit from a mediating software framework. *Meta Trumpet* is a more distributed model, in the sense that the instrument is extended by the addition of sensors both on the instrument and in the performance environment. The simplicity and elegance of valve movement, is combined with performer breath-control, which is extended by adding interactive music elements. These additional elements are triggered by performer behaviours in response to system-generated material, a dialogue is tangible and the independence of the mediating software provides a framework for the performer to explore. Instrument orientation and performer proximity is combined to form an evolving compositional framework. The system is esoteric and not presented as a new instrument for other trumpet players, it's strengths lie in the iterative design of a conversational framework for musical dialogue between a player and a generative composition. This example demonstrates machine musician ship whilst extending conventional notions of live performance with software environments.

Marcello Wanderley (Wanderley, 2003) has written extensively on the topic of gesture acquisition for digital musical instruments, and has undertaken a thorough survey of new musical interfaces that use both direct and indirect gesture acquisition. This work also considers the inherent HCI and design issues that can define interaction within a musical context. The broad range of approaches considered includes real-time control for synthesis, devices for score control, graphical interfaces that use established graphical user-interface metaphors, physical interfaces for post-production and, of particular relevance to adaptive social composition systems, interactive installations where sensing of an individual's or group of participants' interaction is used to input values for an audio/visual/haptic system. Dance-based systems and gaming interfaces are also discussed, in which direct sound manipulation may not be a priority but the approach is still significant in terms of interaction through gesture. The examples discussed are representative of the field, and many are also well documented in peer reviewed conference proceedings detailing system specifications, technologies, and documentation of core features or compositional approach. Less well documented are the interaction or design strategies for developing these new approaches to such systems, as many go beyond the more established human/computer interaction modalities more widely understood by the HCI community. There is a clear opportunity for those developing such systems to articulate their methodologies, documenting the development process and publishing findings based on performers' experiences of these new digital instruments. This suggests that those developing new hybrid systems should consider integrating data collection for analysis and refinement of interaction design.

"On the other hand strategies to design and perform these new

instruments need to be devised in order to provide the same level of control subtlety available in acoustic instruments."
(Wanderley, 2003 pp.1)

Non-Instrumental Gesture describes a range of technologies primarily based on video tracking or environment sensing. These software-based interactive multimedia systems range from simple event triggers based on 2D grid references compared to a captured live video feed overlaid on a performance area to sophisticated vision systems that incorporate multiple cameras, pattern recognition and, more recently, Artificial Intelligence (AI) to interpret live video data. These un-encumbered systems tend to mediate performance data through continuous sensing. Live video analysis and sensor data are combined to communicate with software through a range of context-specific mappings or mediated interactions between people and software. Just as there are numerous methods for data acquisition and parameter mapping there are many alternate ways of decoding human movement described as semantic gesture, Selligman discusses the expressive mapping of participant actions in the multimodal artwork *Medea.*

"Gesture generally refers to dance movements and sometimes to specific body expressions. However, gesture can also be considered a structure with definite semantics defined into an abstract space, as it is here. For example, a musical phrase is a gesture that expresses an emotion using only musical parameters, where music is the abstract space."
(Selligman, 2004 pp.5)

Key Terms: gesture acquisition.
Instrumental gesture, Non-instrumental gesture, Haptic sensing, Non-haptic sensing, Direct gesture acquisition, In-direct gesture acquisition.

1.2.2 Motion - Presence and Influence

A sound environment intended for social interaction needs to be able to distinguish between motion, presence and gesture to predict intended influence over a system. Just as open-source software, such as *Eyesweb* (Website reference 7), has trajectory prediction tools to enable meaningful interaction to be adapted, both direct and indirect events need to be mapped, interpreted and reprocessed in the context of the current gesture(s).

" The system can make decisions based on the incoming information from analysis and the acquired knowledge. Such decisions may concern the kind of expressive content to produce and how to convey it, and can be related for example to the narrative structure of a performance."
(Cammuri et al. 2004 pp.38)

Within a mediated public space or a Smart room it is equally challenging to design tracking strategies to distinguish between active participants and passive observers. Within an adaptive compositional system, video-based gesture capture can be very effective for identifying anticipated movement. It is not necessary to implement a highly specialized system for 3D tracking of individuals in a multimodal environment (Focken and Stiefelhagen, 2002) when the context of the motion or interaction is collaboratively mediated or when the system has adaptive potential.

"Today, it is possible to go further on by programming the relationship between an instrumental gesture and the sound result. So, instrument designing may become a creative act."
(Merlier 2004 p.5)

There are considerable creative opportunities for mediated social works where the designers are able to extend these gestural recognition methods with an understanding of human actions in relation to known behaviour as opposed to purely aesthetic, performance or control parameters. Some considerable development (Cammuri et al., 2004) has taken place in the video analysis of posture within mediated spaces, ranging from literal mapping of hand gestures directly to 3d simulation (Zhu et al., 1997), to extended libraries for developing new full-body gestures. Research at ATR Media Integration and Communications Research Laboratories (Takahashi et al., 1999) presented a hybrid system that combines silhouettes processed from a live camera feed using the Kalman Filter with a virtual simulation or 3D model of body orientation. This hybrid approach resulted in a robust method for recognizing posture and body orientation allowing a repertoire of human characteristics to be integrated. A structured experiment to show system reliability was presented. This work has since been developed but even at it's early development stage the potential for capturing intended meaning through recognized postures was intimated. If we consider the wide use and stylization of poses in performance to communicate meaning or emotion to an audience the potential to communicate this directly to computer-mediated environments is very apparent. Logically we could also implement strategies to interpret or map this action as a behaviour, an intended response to system variables presented to a participant through audiovisual feedback. It is not such a great step to then consider emergent behaviour or subconscious interaction as a viable communication channel for parameter mapping. These issues are more openly discussed in mediated artworks where the primary audience is non-musicians and the social context of participant actions is included in the development of the work.

"Those inside the sensing area (if more than one person is allowed) share in a collaborative "performance," their social interactions contributing significantly to their experience of the work. What is the psychology of this participation? How can installation artists engage, prompt, and empower amateur "performers" who have no prior knowledge or particular expertise? How does the computer program and content facilitate action and encourage response?"
(Winkler, 2000 pp.1)

Winkler provides a considered analysis of these elements within two motion-sensing installations and establishes factors for consideration in the design and evaluation of such experiential works where the participant becomes a significant part of the content or process.

These questions relate specifically to strategies for inclusion where an open work or interactive mediated environment is presented. New methods for gesture-based interaction allow a flexible approach to design of such systems. Winkler discusses a number of strategies relevant within his own works, listing factors for consideration when devising these inclusive forms. In the case of motion-based interaction within collaborative works, it is useful to consider the two primary states of interaction, *presence* and *influence* before assuming specific musical gestures will be enacted.

The *presence* of a person is easily detected, though this alone does not necessarily mean they are engaged with the work or relating to the presence of others. Perceptual cues in the form of audio or visual feedback can communicate or acknowledge this individual presence and elicit a further response. A logical strategy is to assign a property or value to this person, an identity within the context of the mediated environment, a tangible response. Strategies need to be in place to motivate participation, *explore-ability* of the interface method should not distract from the layers of *influence* open to the participant, i.e. the exploration of the technical control should not dominate the learning process as the participant adapts to or learns from the system. With an adaptive approach, the system design should include strategies for communicating *presence* and *influence* to participants; ideally, the system demonstrates its adaptive nature through response and collaboration in an accessible form.

Key terms: motion detection modes and participant behaviour.
User-presence, User-influence, Explore-ability.

1.2.3 Environmental Parameters and Context

One of the challenges of *real-time composition* for a specific venue, event or location is establishing a tangible musical relationship between the system, participants and the environment (Camurri “et al.” 2000). This is significant if the composition process is intended to be adaptive and open to influence. We are familiar with optimizing environments for sound reproduction, adapting a live performance to the acoustic limitations of a given venue through instrument, performer or monitor placement. Live mixing allows for the acoustic limitations of an environment to be overcome, compensated for. This assumes that a consistent reproduction is desirable, that the core material of the performance is predetermined. Numerous examples exist where the qualities of an environment, be they natural or manmade, are factors in both the conception of a composition and the delivery of it. From the humble wind-chime suggestive of water in a Zen garden, through sound sculpture and public sound art to current diverse approaches to *sonification*, composers have considered ways of harnessing *environmental parameters* to mediate these works. The wind-chime example illustrates how one natural parameter, air movement, can be harnessed to symbolize a perceived missing element, water. It is the careful selection of material and placement of the chime that reinforces this mapping of a single parameter. Of course, this air movement is a variable allowing the sonification to suggest a range of qualities and association. This variable is also subject to influence, a gate or doorway entered, a group of visitors and so on. Many interactive works follow the same model. A variable generated by the environment, is used either directly with a tangible mapping, or indirectly to mediate another element of the process. This provides variation that is contextually relevant and perceivable by an audience. This is not simply a reactive approach as a number of parameters can be identified, harnessed and mapped to create new compositional material. The typical parameters for mapping might include presence, proximity, direction and number of participants. Incidental sound qualities or participant interaction can also be harnessed using amplitude, pitch or beat detection. Of course, the physical properties of the environment may encourage different ways to experience it; just as the interface method itself, can create and mediate new experience (Rokeby, 1998). Many approaches use video tracking, pattern recognition and other *data analysis* to apply real-time data from an environment to a compositional process. Other data such as air temperature, air movement, humidity, light levels, magnetic fields, vibration, pressure etc have also been explored. More recently, building management systems data intended for monitoring services and maintaining environmental conditions have been harnessed to generate variables for composition. An evident problem with this approach is that many of the values generated are compositionally meaningless and removed from their context illustrating minor abstract data variation over extended periods of time. Compositional strategies for mediating data within these systems, is an aspect of this research (Livingstone, 2004). Dedicated sonic installations and intelligent interactive rooms are more established as a dedicated physical location or custom system design provides a higher degree of compositional focus. An alternate approach that harnesses remote data to form new compositional material from online weather forecasts is Garth Paine's Weather Sonifier (Website reference 20)

Key terms: environmental context.
Real-time composition, sonification, *data analysis.*

1.2.4 Ubiquitous and Pervasive Interfaces

Other systems such as the web-based works of artist Stanza (Website reference 21) take the output of CCTV or surveillance cameras to generate compositional material transforming data to create new visualizations or sonification. Such approaches to transforming data often explore our relationship to technology or the way it pervades our lives. Some practices actively reveal hitherto hidden technologies or processes; others subvert existing technologies (Magnanensi and Rolfe, 2005) using circuit-bent educational toys, transforming the audio output into new forms through electronic intervention, live sampling and deconstruction of the original devices as live performance. With these works, the aesthetic or compositional intent is often simply to reveal technological processes or initiate a process of creation from uncontrolled sources. Other approaches extend existing personal technologies such as mobile phones, personal digital assistants (PDA's) and Palmtop computers by harnessing Global Positioning Systems (GPS) for location or trajectory data. A key feature of these works is that they transform an existing signal or data set into new sound material using *generative* or *transformative* processes. These processes range from application of algorithms to juxtaposition of material from different sources to create something new. Some of these works seek to augment reality through location-based narrative as in the binaural Sound Walks of Janet Cardiff (Website Reference 22) others harness participant movement and collaboration to create a *malleable music* (Tanaka, 2004). Just as public data is captured through a network of sensors and can be transformed for musical output (Gaye et al., 2003), personal data devices can be used both for location-based tracking, collaborative scoring of stored music fragments and as new digital interfaces or instruments in their own right. A number of these devices support third-party software development kits in addition to crude sampling, mixing and sequencing software preinstalled for consumer use. A number of tools have also been created for the mobile music community including *pd2j2me* (Scheimer, et al. 2004) which cross compiles MIDI based applications created in the Open Source environment Pure Data to Java 2 Micro Edition (Website reference 23), which is supported by many mobile handsets.

Key terms: pervasive processes, public works.
Generative, Transformative, Malleable.

1.2.5 Instrument Design for Collaboration

The term instrument is generally accepted, in musical terms, to describe a physical object that can be manipulated to make sound, ranging from the simplest percussive object to the most complex stringed instrument. Years of refinement in terms of materials, craftsmanship, playing styles and the subtleties of the physical interface have led to literally hundreds of variations. Many musicians have further refined their instrument through minor modifications, or simply through years of playing causing a gradual but significant wear in response to their playing style. Others have gone on to design or commission their own custom variations. New technologies have allowed further developments, the evolution of the electric guitar for example, from the earliest single coil pickup to the most sophisticated hybrid nylon strung MIDI guitars with the playing feel of a carefully handcrafted bespoke instrument and the performance possibilities and customization of sound of a high-end synthesizer. The key point here is the effectiveness of the interface in translating an intended action into a sound. The humble woodblock can be struck in a variety of ways with different materials to achieve a surprising breadth of sounds, yet it is possibly the simplest and probably the oldest handheld percussion instrument. Adding a microphone and processing the analogue sounds captured can lead to a complex array of sound textures, percussive structures with a high degree of personalization by the performer. Similarly adding beat detection or measuring time between hits can produce more variables to affect the processing of these sounds adding infinite variety from a very basic action, but does this create a more engaging instrument, or simply add layers of complexity that cloud the effectiveness of the original interface. The humble woodblock has been evolved using wide and varied shapes, beaters and materials developing an extensive potential repertoire and rich sonic variation from a simple act of selection, timing and striking. Does adding multiple layers of gestural interpretation and additional synthesis really add any significant musical extension for the novice or group who use them? The subtlety of the original instrument is that slight variations in timing and striking are rewarded by tangible sonification that is refined by listening to a wider musical context. Current examples of new interfaces for exploiting physical surfaces or textures using an *exploratory model* include the *Sonic Scanner* and *Graphonic Interface* (Overholt, 2004). Of course, the term 'instrument' in a Computer Music context also refers to software instruments or physical controllers for hardware or software based sound generators. Leon Theremin's innovative capacitive field interface of the twenties, possibly the earliest form of custom electronic interface for sound generation and gestural performance, has inspired numerous designs for electronic instruments and controllers designed for highly complex individual performance (*Wanderley and Orio,* 2002). There are many recent examples of custom electronic controllers with integrated sound generators that can be described as instruments in their own right. Don Buchla's *Thunder* (1990), *Lightning* (1996), and *Mirimba Lumina* (2000) are carefully crafted and sonically refined innovations that continue to inspire the field. However, fewer examples exist of groups of custom instruments designed for collaborative and social creation of sound. Certainly at the forefront of the NIME (New Interfaces for Musical Expression) field, specialized examples can be found for manipulating live sound material. Weinberg's *Beatbugs* or more recently *Itur* (Weinberg et al., 2005) interfaces are an established example for collaborative improvisation. Increasingly children's toys include examples that provide novel interaction that goes beyond established instrument models of wind, string or brass. Perhaps this is because the average home is more likely to have a television and a computer than an instrument, but toy developers recognize that effective tangible interaction with sound engages younger audiences. The *Bopit* developed by Hasbro is one of these (Website reference 24). Effectively it combines a series of hand actions; press, twist, spin, strike, bend, to activate pre-recorded sounds, a simple onboard chip sets sequential tasks, patterns for the child to repeat etc. The 'interface' can also be used to 'freestyle' encouraging the child to create new patterns and compete with others. The *Bopit* is superficially engaging as an intuitive interface for novices but the included sounds (mainly percussive spot effects) are of such poor quality that the toys have limited appeal beyond their initial novelty factor. Obviously groups of musicians working with their own custom instruments will improvise, perform and create new material together and composers continually seek new approaches for creating and arranging sound material but my specific interest is in accessible interfaces for non-musicians to explore the medium of sound together with some form of adaptive mediation. An established strategy is to adopt algorithmic approach for extending simple actions in handheld devices. Weinberg describes a collaborative performance using *Beatbugs*, novices and live performers.

" A variety of transformation algorithms were developed in an effort to provide expressive and intuitive musical control for novices. Some of these algorithms utilized direct mapping between continuous

bending gestures and fundamental musical aspects such as pitch and rhythm. Other algorithms used more sophisticated stochastic operations in an effort to allow players to control aspects such as melodic similarity or rhythmic density."
(Weinberg and Driscoll, 2005 pp.21)

Historically it is within the area of percussion that one finds small and large-scale groups effectively interacting without formal instruction; structures and patterns evolve and transform over time, new textures and timbres wax and wane by a seemingly subconscious interaction, perhaps initiated by a single more experienced player. Perhaps it is the simple fact that as each instrument is relatively limited, each player has to explore the subtleties, the range of possibilities to maintain their interest and sustain a valid contribution to the whole. Above all each player must listen attentively, whether a novice or an experienced player. It is perhaps no surprise that the many successful new instruments or novel controllers for group interaction of non-musicians (Blaine and Fels, 2003) include those based on a percussion model. For effective social interaction, many interface designers believe that a simple learning curve balanced with an engaging yet effective control over a musical range is a reasonable starting point. For sustained engagement, it is also the interplay between players and how a multi-user system or network of interfaces relate and transform this interplay that provides a significant design challenge within an adaptive system. This is less obvious in terms of system design and clearly requires a design strategy where the software element of the system can recognize and adapt known or learnt gestures in order to be perceived to be responsive to players. The more established *design principles* (Cook, 2001) the core concepts of *efficiency*, *learn-ability* and *user-role flexibility* (Jorda, 2004) combined with the significance of expressive range in terms of *instrument resolution* and *expressive depth* (Settel, Lippe, 2003) provide a clear framework for musical interface design for collaboration. *Explore-ability* (Wanderly, 2001) is also a key feature, both as an evaluation factor in terms of an interface's musical potential but also as a method or strategy for training both the instrument (or mediating software) and the performer. In order to establish a viable design process, in terms of collaborative musical interface design, we can prioritise these identified elements based on the intended context for interaction. Ultimately it is the nature of engagement, the relation between *gesture and mind* (Paine, 2002) that the designer must address if these interfaces are to develop beyond existing models of music and performance. It follows that the context of interaction can only be effectively mediated if multiple channels of communication and exchange are incorporated into system or interface designs. Chapter 3 articulates a series of design strategies to integrate these principles.

These design strategies are applied to the author's digital instrument design Orb3, as discussed in Chapter 4. The overall concept was to develop a system design including physical interface objects, mediating software and ambisonic diffusion, by applying these principles and key concepts while adapting approaches from related fields. For example, techniques for identifying and triggering conventional musical events in interactive music software can be extended to provide short and long-term memory of gestures within an adaptive framework (Livingstone and Miranda, 2004) extending *learn-ability*. The location of participants in relation to interface objects can be tracked and monitored to recalibrate *instrument resolution.* Performer behaviours can be captured and compared in order to extend and modify *expressive depth. Explore-ability* is used to develop the soundscape players can interact with, mapping environmental parameters of the performance or installation environment directly into the compositional space to extend *performative engagement* (Paine, 2002).

The definition of instrument in this field is much broader than previously recognized. It includes physical objects with embedded sensors that capture and transform data. These interfaces are dependant on a more complex integration of software, video tracking and sound diffusion. As a definition of instrument this is not so radical, Lucier's composition "I am sitting in a Room" (Lucier 1995) effectively turns the room the piece is performed in into an instrument. This is the logical development of musical interfaces that extend beyond the *excitation-sonification* model of acoustic instruments and apply a multi modal approach to instrument/interface design for co-creation, developing an adaptive social composition.

"If dynamic morphology is applied to the design of responsive and interactive instruments and installations, it becomes clear that the system design itself must be dynamic, and that during an interaction, an instrument must be able to change in fundamental

ways to produce timbres that were impossible at its inception. In other words, it must be possible for it (in accordance with the nature of interaction) to evolve into a new instrument altogether"
(Paine, 2004 pp.20)

Paine addresses a core point here, in that many developers of novel controllers and new digital instruments are not learning from the previous work in machine musicianship or interactive music, with mediating software or composition added as an afterthought. Many designers appear to focus on technical development, improved processing or sophisticated direct gesture acquisition within the established instrument model. This produces novel devices, but that in itself does not contribute to new musical structures, as these devices are connected to conventional music software processes using familiar fixed interactions. As instruments or interfaces become integrated with more complex systems or computer-mediated installations the tangible handheld element is often only a collector of data, a sensing device exchanging a variety of variables for software transformation, synthesis and reprocessing. Since the first computer games, physical interfaces for collecting and mapping human responses to mediated content or virtual experiences, a wide range of hand held devices have been designed. And as the established model of control over a character or vehicle becomes dated, new interactivity is created using previously unexplored genres and simulation. Of particular relevance are game formats that go beyond a learnt control model, extending the game world through novel controllers (Blaine, 2005) and new motivational strategies or game-play to reinforce immersion through interaction, manipulation and creation of sonic material. This convergence motivates innovation and offers complimentary frameworks for managing collaborative music systems. Researchers are increasingly investigating strategies for understanding performer or user *intent* and *response* in order to develop effective *haptic impressions* or forms of feedback that are not dependant on conventional acoustic instrument properties (Gillespie, cited in Cook, 2001). For example, modifying parameter mapping or implementing learning strategies based on observation of interaction, principles of music cognition and developing new system behaviours to motivate exploration.

"A few artists have gone so far as to devise installations requiring cooperation amongst players to realize their work (Ritter, 1997). However, the potential for a work to foster social interaction is an essential artistic decision that will not be appropriate for particular works, or may be limited by a given space, or the capabilities of software and sensors."
(Winkler, 2000 pp.27)

On developing this research concerning public collaborative interaction with mediated sound, a symbol-based approach for tracking participant location and group gestural control was developed. A large-scale open atria connecting lecture theatres, seminar rooms and social spaces was used as an example to articulate this compositional approach. The Portland Square building at the University of Plymouth features three interconnected atria with a grid based video tracking and multiple speaker system. This is intended to support a range of experimental projects. A single overhead camera and three tiers of four speakers are installed at floor level on the first, third and fifth floors of two of these atria. A novel element of this location was the inclusion of building management system data as a resource for developing novel works in response to environmental data. Parameters such as temperature, water usage and air circulation are made available via a website. A web-based interface to this collaborative resource entitled 'Arch OS' (Website reference 54) provides data from the overhead cameras and access to embedded diffusion software to distribute sound material to the speakers located in each atrium. The social interactions explored in relation to this system are more fully described in *Composition for Ubiquitous Responsive Environments* (Livingstone and Miranda, 2004) This peer reviewed paper was presented at the International Computer Music Conference in Miami 2004 and is included in appendix 2. The following summary highlights the core elements and initial findings of this work.

Through observation of individual and group behaviour walking through public atria, strategies to evolve new behaviours by combining short and long-term memory within software were proposed. Although more often these terms are discussed in relation to human cognition (Levitin et al., 2002) this design proposed a method for interpreting the motion and relative locations of people as gestures that could be described in lists and compared to video analysis of a public space. This lists could be stored in short term memory so gestures were contextually relevant, repeated lists could be sorted and organized into classes and activated when conditions were repeated on another occasion. The lack of a tangible interface in this system design motivated an alternate approach to collaboration through individual or group positions or trajectories of participants in a public space captured via a video feed, with sound objects created and diffused in response to conscious and subconscious interactions. This approach begins to explore *auditory perspective* (Chowning cited in Cook 2001) through synthesis and diffusion of sound objects intended to motivate behaviour of participants.

" In fact, the understanding and exploration of these issues suggests somewhat magical musical and acoustic boundaries that cannot be part of our normal acoustic experience, yet can find expression through machines in ways that are consonant with our perceptual/cognitive systems."
(Chowning cited in Cook 2001 pp.275)

Within a public space of this scale, basic transformations of live data, participant locations and individual trajectories can be implemented relatively easily as the technical infrastructure is provided. Additional software can be written to mediate a wide range of data to create alternate sonification or feedback systems. The initial testing for such a system was carried out using Portland Square atria b, Plymouth in 2004. The mediated composition was intended to be responsive to individuals and small groups, while using building management data to provide a sound environment, which participants could interact with. However, once the basic control, response and synthesis elements were resolved it soon becomes apparent that the context for layered interactions and meaningful musical exchange was unresolved. The fixed position and orientation of speakers in this location, did not allow the use of ambisonics. The stereo field of each pair of speakers is located far above people moving through the environment so only a general trajectory of certain frequencies could be perceived. The extensive glass surfaces and high ceilings added further complexity so sound diffusion could not be used discretely to motivate or mediate interactions effectively. The building management data provide is representative of changes over significant time periods, so although these parameters can be mapped to generate sound material, making this material contextually relevant is a significant challenge. The volume and movement of people at various stages of the working day means that the resulting data quickly becomes generic and differentiating between participating individuals and groups is not feasible unless a number of controlled conditions are introduced. The multiplicity of parameters in large-scale social environments, have to be dramatically simplified or they become too complex too quickly.

An effective design strategy was needed to manage this complexity. Emphasis was placed on tacit referencing of symbols for interacting with the system by creating a simple *gestural vocabulary* using *line*, *triangle*, *square*, *and circle*. For example, when video-tracking groups of people using an overhead camera, simple classes of gesture can be created. Individual trajectories can be mapped to direct panning of a sound relative to an individual's direction while walking in a *line* through a given region, applying a simple *control model*. A higher level intervention can be triggered by identifying individuals standing still, a simple rule can be implemented to link relative static points into groups, a group of three form a *triangle*, a group of four form a *square*, a group of five form a *circle*. The significance of these simple symbols is that they can be easily mapped to tangible musical events, if appropriate feedback is provided. A *triangle* can be interpreted as three variables by calculating the distance between points; these variables can be mapped to create a new note *sequential model*. Basic MIDI note parameters, pitch, velocity and duration can be used to trigger a new sound. This alerts participants to the fact that *proximal interaction* generates new sound material *relational model*. Classes of gesture or parameter mapping of symbols can be contextualised in a tangible way for participants. Forming a *triangle* and generating a note can happen unintentionally, but is rewarded, attracting attention of those who have chosen to stand still. This shift in attention is useful to move a passive individual into an active participant. When participants recognise the application of the pattern, exploration is motivated. The pattern is repeated in different locations, new notes are created new participants are invited to join in *exploratory model*. If we visualize the principles of cellular automata, the process of individuals moving relative to identified positions and making simple decisions that reveal a larger evolving pattern is surprisingly similar. Further rules can be added to allow a higher level of interaction that is contextually relevant. For example, an individual joining a standing group of three can modify a *triangle* to a *square – organizational model*. This group action to form a more structured set of *proximal relationships* can be classes as a different gesture with previous parameters being remapped and a new parameter being introduced. For example, instead of simply triggering a note (*triangle*) the distance between points can be used to control synthesis parameters of the current instrument the note is playing, *transformative model.* Once a group of individuals is identified as a symbol, the associated musical event continues until the group separate or is identified in a new formation. A second significant realization was that one needed to motivate effective participation beyond simple control models. A number of approaches to *direct gesture*, *indirect gesture* and *proximal interaction* are required. These would need to be combined both at a technical level and a conceptual level to respond to interaction and behaviour. This phase of the work introduced the concept of conscious and subconscious interaction, providing examples of how different participant behaviours could be interpreted in context by combining interaction models with classes of gesture.

Key Terms: Collaborative interfaces
Efficiency, *Learn-ability, Explore-ability, User-role flexibility, Instrument-resolution*, *Expressive-depth.*

1.2.6 Virtual Composition Spaces

An alternate approach to collaborative interactive music can be seen within the design of virtual worlds, where the characteristics of a virtual space are reinforced by effective sound design (Website reference 25) or where character emotional states are conveyed by interactive music (Tayler et al., 2005). These types of spaces show how effectively interactive music can reinvent the roles of performer, instrument and composer. The simple exploration of a previously unexplored virtual location by a group of individuals, such as in a typical multi-user game space, can efficiently illustrate the complexity of interactive music. In this case, the composition is mediated entirely by software in response to the virtual location, movement, orientation and actions of both individuals and groups. Typically each participant in these spaces experiences a personal *soundscape* diffused using surround sound, so spatial orientation and the position of game elements and the relative location of other players are clearly referenced within the individual's mental model of the game environment. This is evidently a highly immersive form of interaction as navigation of these worlds alone motivates *explore-ability.* Clearly in this context any spontaneity in player reaction to these mediated *soundscapes* is secondary in the players mind to actually playing the game, however as with many media forms it is only a matter of time before these environments are used as compositional spaces. Indeed the visual potential of these virtual environments has already been reinvented, the form known as *machinima* (Website reference 26) where filmic sequences are created through skilfully choreographed player movement and original character 'skins' and level design, by recording the modified output of these environments. This output from multiple virtual cameras or player viewpoints is then edited to create 3D animated movies. Just as advanced users are able to modify the graphical elements, textures, models, lighting and physics within these worlds the sound material and the methods used to create it are being reinvented, enabling the development of interactive music composition spaces, where former players become both listeners and performers. Of course this general example illustrates the possibility for extensive network-based multi-user compositional experiences. Many systems of this type are currently being developed (Yeo et al., 2004), where a number of parameters are influenced in real-time through networked participants who access the environment via a web-browser, and have the ability to contribute new material for others to influence and respond to in real-time. Many examples of web-browser-based and multimedia systems implementing such multi-user interactions exist, the established models and paradigms for these multi-user instruments (Weinberg, 2002) are well documented.

1.2.7 Summary (section 1.2)

This section extended the initial discussion introducing technological approaches used in related works. Fundamental principles such as *presence, motion* and *influence*, for classes of gesture acquisition and motion capture were presented. Levels of expression were considered in the context of a cognitive linkage between the performer/participant and the presented instrument or musical framework. New interaction principles and published concepts (*efficiency, learn-ability, instrument-resolution and expressive-depth, explore-ability, intent and response, haptic impressions and auditory perspective*) from the wider field of Computer Music were introduced. Examples of developmental works demonstrating these identified principles were discussed. Interaction models (*exploratory, sequential, organizational, relational and transformative*) were identified in related Tangible interfaces that can be used to express actions and behaviours of participants using collaborative or shared sound environments. The phenomenon of *emergent behaviour* was considered in a broader context. The concept of mediated interactions or *process driven collaboration* was identified as a strategy to allow participants to *play, intercept and make contributions*. The principle of *co-creation* as a defining feature of systems intended for social interaction within a mediating framework was also established.

The following section identifies a range of *System Models* and presents a series of design issues referencing specific examples of tangible interfaces intended for collaborative interaction.

1.3 System Models

We will not consider the wealth of compositional approaches that make up the broad spectrum of Computer Music, but instead we will consider approaches with direct relevance to collaborative forms of interactive music or sound installation and significant historical works. Within the field of new interfaces for musical expression, there has been a continued research interest in establishing compositional frameworks for non-musicians that motivate collaboration through musical experience (Blaine and Fels, 2003).

Within an adaptive composition for collaborative interaction, key questions arise:

a) What features of the process of composition can be distributed between novice participants?
b) Which strategies can be used to invite and motivate interaction?
c) How can non-musicians be engaged in meaningful musical exchange?
d) What is the nature of collaboration within such a process?
e) Where do these interactions take place?
f) How can one go beyond a *call and response* model?

The answers lie in the novel approaches that explore new, technologically mediated situations, which endeavour to extend the creative process. New forms of composition and performance explore the periphery of musical practice while challenging convention. At the same time, we must consider our audience. If they are to participate then the forms of interaction, the roles we expect them to assume and the experience provided needs to be accessible and engaging.

"What separates interactive installations from other types of art installations or interactive performances is that the work is only realized through a participant's actions, interpreted through computer software or electronics, and those actions do not require special training or talent to perform."
(Winkler 2000 p24)

The following section explores the questions presented above in relation to systems with an emphasis on collaborative and social interaction. The general technologies these systems use are introduced, with specific examples that facilitate multiple approaches to gesture acquisition.

1.3.1 Performance Systems

Approaches dedicated to enhancing live performance by extending an individual performer's expressive grammar and influence over audio visual material are increasingly common and have led to established genres of augmented performance (Sparacino et al., 1999, 2000). With a proliferation of open-source and commercial solutions for tracking a performer's actions, the implementation of such systems is relatively easy to accomplish. Within Max/MSP software, the additional Jitter library can be used to implement a range of video analysis methods, some third party developers also provide additional externals or add-on software for video analysis, pattern recognition, and object tracking. Of these, 'Taptools' by Timothy Place (website reference 11) and 'CV.jit' by Jean Marc Pelletier (website reference 12) at IAMAS (Institute of Advanced Media Arts and Sciences, Japan) provide low latency solutions and robust examples for a wide range of methods that can be applied to Computer Music and performance situations. What is not so easy to achieve is a meaningful application of these methods to an individual performer's interpretation or response to a musical context. Parameter mapping between tracked regions of live video feeds and generative processes may be an engaging extension of a performer's influence on the digital realm (Dobrian and Bevilacqua 2003), but it is all too easy for these works to emulate other works if the underpinning processes are not evolved beyond the established models. Tangible cause and effect may reward the performer for exploring their influence over their mediated environment but it is debatable whether these extensions into the digital domain are rewarding for a passive audience, primarily because these systems are designed to be experienced, not simply received. The kinaesthetic linkage experienced between an accomplished performer and a well-crafted audio-visual composition can be compelling, particularly when this interaction evolves within the performance (Machover and Chung 1989). Some established performance collectives in this area have further defined their existing practice by developing their own custom software, which is designed to allow choreographers a direct interface for mediating the performance through dynamic mapping of performance elements to interactive scores (Rovan and Weschler, 2000; Weschler and Weis, 2004). This interaction through intervention in real-time processes and mapping of actions to events between performers or participants and a system provides an engaging model for placing a less physically active participant in an influential role. This is useful because presence alone can be used to shape a compositional context. Different responses or roles can be identified and constructively integrated to reinforce more challenging musical relationships between participants and a system.

1.3.2 Performer/Machine Systems

Rather than extending an existing aesthetic or performance model, *performer machine* systems often challenge the performer to adapt to a technology or environment (Bongers, 1998), reconfiguring their creative responses. It is often the adaptive qualities of the performer that engage the viewer; the compositional exchange is evident to audience members.

"At these concerts, the audience becomes involved in the compelling energy of the performance, the relationship between physical gesture and sound, and the musical communication between the three performers."
(Bongers, 1998 pp.13)

A new generation of video, computer, and console games player/performers have also been exposed to a performer/machine experience that goes beyond their previous character, weapons or avatar control featured as in many commercial titles. Small developers have tapped into the potential of novel controllers to extend the interaction with already highly immersive game worlds. Innovations to the gaming community such as Sony's *Eye Toy* gaming environment (Website reference 27) use simple video tracking to integrate players more dynamically through physical actions. As the Computer Music community creates new hybrid systems, small games developers design custom controllers such as *Gametrak* (Website reference 28) for direct gesture acquisition. One game developer in particular, Harmonix (Website reference 29) represented by Josh Randler at Cybersonica 2006, Dana Centre London, has produced a series of games products that use a custom interface, packaged with the game product itself having previously produced mould-breaking titles that challenged players to sequence, sample and playback audio content through established game-play interactions. Industry veterans Nintendo have also identified the potential to innovate their products through more sophisticated interfaces (Website reference 30), which incorporate the same sort of sensors used by the NIME community, such as force sensing resistors, accelerometers and infra red. Typically, these sensors are combined to capture direct gestures and player proximity and map player actions into the game world. In fact the Computer Music community has been modifying standard games controllers for many years, with some researchers actively encouraging collaboration between novel interface creators and an industry which has a long history of mass marketing new controllers to extend interaction (Blaine, 2005).

1.3.3 Collaborative Systems

Just as audiences have been absorbed by innovative performances and technologically mediated live musical practice, a new form of interactive music has been established. The development of custom controllers and hyper-instruments for individual use has an extended community. Initially collaborative works for non-musicians followed a group percussion model with numerous handheld devices being developed within the NIME community (Blaine and Fels 2003). The range of currently accessible collaborative composition systems falls into several forms; handheld interfaces for groups, the distributed network (Barbosa, 2003) and the Tangible interface.

During the process of this research, the author has tasked undergraduate interactive media students to develop 'Sound Practice' assignments as part of their degree studies. These assignments were focussed on developing a novel controller or Tangible interface with a supporting sound application developed in Max/MSP to process collected data, with the intention of exploring interaction with mediated sound. The techniques and methods for gesture acquisition and the terms used to classify different interaction modes were presented to final stage students. They were then able to prototype novel interfaces and mediating software to create generative compositions with a degree of user influence. A series of circuit designs for sensors such as digital compass (relative orientation) accelerometers (speed and direction of movement), ultrasound (relative distance) and bio-feedback (heart-rate and galvanic skin response) were provided. In addition, collections of simple to use components such as light dependent resistors (ambient or local brightness), bead thermistors (ambient or local temperature) and tilt sensors (activated at set angles). Workshops in basic electronics were provided, demonstrating efficient ways of modifying commercial game controllers to provide custom interfaces for use with these sensors. Other established sensor interfaces were also provided for testing and prototyping. A notable outcome of this process, running from 2003 to 2007, was that the prior musical background of a student had little effect on the quality of the outcome. Students were able to engage with higher-level design principles, and were actively encouraged to move beyond direct mapping of a control to a literal parameter or predictable interaction. Students were required to submit a short paper format critique of their work relating it to examples and research presented in related conference. In some cases, the standard was particularly high and the insight these students were able to develop into the nature of interaction design was greatly improved. Evidence of this can be found in a collaborative live open performance of a selection of these interfaces and software that was featured at the Port Eliot Literary Festival in 2005. The performance featured several handheld Tangible interfaces and a virtual composition environment manifested as a 3D stereoscopic immersive environment in which one performer could co-ordinate the material produced by his peers. A stream of sound could be sampled and associated with a virtual representation that the mediating performer could control using video tracked hand gestures. This allowed the audience to observe both the generation of material and its orchestration concurrently. Audience members were invited to take turns using the Tangible interfaces, generating new material for the immersed performer to mediate. It was a credit to the students concerned (Eggins, Feldman, Hayward and Hackney) that their initial explorations in this field could be so effectively integrated into a larger collaborative system which guest musicians were also seen to interact with effectively. This first collaborative outcome, built by novices for novices, was developed considering the principles and terms used to describe such systems.

This process effectively demonstrated the value of empathic design and the potential of shared design strategies using defined terms to focus creative and technical development. The significance of providing an accessible vocabulary allowed students to engage with complex processes and identify appropriate techniques for creating new digital instruments. A significant element of the software design was consideration of multiple modes of interaction and a variety of sensor data. Using shared design strategies for data collection, transformation and formatting enabled different hand held or worn interfaces and proximal interaction using motion tracking to be integrated effectively. Tangible interfaces designed for novices are usually limited to a collection of common hand-held interfaces following the same interaction model. Another established format is to provide a shared surface, generic objects and video tracking. It is less common for the characteristics of performer machine systems to be incorporated in a collaborative sound environment for novices.

> *"Interactivity and real time may be in vogue now, but time and communication are the essence of music. These are concepts that are essential in our art form, independent of technology. Musical performance has always had multiple modes of interaction: between performer and instrument, between conductor and orchestra, and between musician and audience. The stage has always been a real-time environment."*
> (Tanaka, quoted in Bongers, 1998 p3)

An extensive collection of tabletop 'stages' or composition frameworks has been developed for small group interaction between novices and performers alike. These systems, introduced previously, follow established models for motivating interaction; *object placement, removal and orientation* and *visual feedback* (Jordà, 2003) are typical of the form.

(fig. 3 *Reactable* Jorda et. al. ICMC 2005)

These multiple modes of interaction offer easily defined interaction models within these tangible interfaces: *Exploratory, Organisational, Sequential, Relational* and *Transformative* (Livingstone and O'Shea, 2005).

These shared stages allow participants to perform all the established elements of *sampling, sequencing and mixing;* often using author provided material or contributed samples from other participants. Some systems enable users to *control and modify* individual elements through *signal processing*. Fewer systems provide *real-time synthesis* or complex compositional interaction such as evolving new musical behaviours. A representative collection of tangible interfaces was identified and categorised by interaction model

(Livingstone and O'Shea 2005), these tangible interfaces were identified as *Exploratory; Sound Garten* by Wolf, 2003 (Website Reference 31), *Organisational; ISS Cube* by Quarta, 2003 (Website reference 32), *Sequential; Blockjam*, (Newton-Dunn et al., 2003) and *Relational*; *Audiopad* (Patten et al., 2002). There are numerous new tangible interface systems in development but the majority fall into the categories presented. The paper (Livingstone and O'Shea 2005), proposed a new category; *Transformative* or '*Freesound'* to identify the interaction models evident within more complex tangible collaboration systems such as *Reactable* (Jorda et al., 2006) which include networking and higher level compositional interaction for both novice improvisation and scored performance. The fact that many of these examples are table-based motivates small-group interaction that develops from turn-based or competitive play seen more commonly in traditional board games, and just as board games have superficially simple rule sets for framing interaction, a well designed compositional process can have higher level adaptive elements. Initial work in this area evolves composition as gameplay unfolds between a human player and the computer. Software associates moves with musical forms evolving a neural network (Miranda and Zhang, 2005). The high-level strategies of *attack, counter attack, subterfuge, distraction* and *anticipation* evident in many board games can all be considered as potential compositional strategies to motivate collaborative interaction between system and players. Of course, diverse models of collaborative game-play are well established within the computer games community and there is healthy crossover between independent games developers and the design of novel music controllers (Blaine, T. 2005).

1.3.4 Summary (section 1.3)

The final section of this introductory chapter proposed design questions specific to System models within mediated sound environments. Specific examples were cited presenting identifiable interaction models, while introducing potential methods for extending collaborative interaction with tangible objects through gameplay and identifiable behaviours.

1.4 Chapter Summary (Collaborative Sound Environments)

There are wide ranges of approaches that attempt to offer novel interaction with compositional structures to engage new audiences or establish a new interface or software design for musical collaboration. Many of these examples, despite being technically innovative, are compositionally unresolved. Hand-held novel interfaces tend to follow long established percussion principles, and while technically effective, do little more than add new sound processing to an established social process. Tangible interfaces appear to offer a more versatile framework for groups of novices to engage with new musical contexts using shared 'stages'. Accessibility is reinforced by using generic objects and shared 'stages' that superficially follow a collaborative model of turn-based board games. A limitation of many of these systems is that they don't apply the engagement of social interaction or complexity of layered events that traditional board games exhibit. The generic Perspex puck and tracking symbol simply becomes another re-configurable controller with stylized visual feedback on a shared surface with limited expressive potential. Highly developed hybrid instruments combine the expressive qualities of a traditional instrument with custom electronics and extensive synthesis to create new material. And yet when performed to an audience these new instruments are often too challenging both to play and to listen too, a multiplicity of parameters can be mapped but in what context, for what purpose. A wide range of gestures and interactions can be captured, processed and transformed but why is this desirable?

Other subsets of Computer Music offer a wider range of technologies for gesture acquisition to create new musical material. An initial example of a Design Strategy to use a symbol based *gestural vocabulary* with social groups in a shared compositional context was discussed. The terms presented contribute to the emerging field specific vocabulary and are used in the following chapters. These approaches offer alternate system models and use different classes of gesture to categorise interactions and contextualise behaviour.

The initial research question was explored.

'*Can a musical interface be designed that engages the novice but has the expressive qualities and personalisation of a traditional instrument?'*

Example interfaces, instruments and systems were discussed that demonstrate some each of these characteristics individually but no single example provides an integrated solution.

It is evident from the literature that a variety of solutions exist but in order to resolve this question, design strategies are needed to integrate these characteristics effectively and resolve the following questions.

1) Can methods of score following and interpretation of MIDI events (*interactive music*) be used as a model to identify different users or user actions with a tangible interface or novel music controller?

2) Can the characteristics of a musical dialogue between a specific musician, instrument and software (*machine musicianship*) be used to respond and transform these different novice actions in a shared musical context mediated by software?

3) Can the processes developed to interpret a multiplicity of parameters (*performer machine systems)* be applied to a reduced set of musical interactions that are intuitive to learn through a *gestural vocabulary* using a Tangible interface?

4) Can principles be drawn from software models of behaviour such as *sociability*, *self organization*, or *pattern propagation* (*algorithmic and generative approaches*) inform synthesis of sound material or mediate behaviour to establish a meaningful dialogue with novices.

5) Can sonification of novice gestures or relative positioning of sound objects (*sound diffusion*) be used to establish a dynamic compositional framework. Can *proximal relationships* be designed

to motivate a range of musical behaviours by combining hand-held tangible interfaces within a diffusion environment?

To resolve these questions and develop effective design strategies a vocabulary of relevant field specific terms is required. A core problem is that each subsection of the field applies generic terms or descriptions to specialised contexts.

The following explanations of field-specific terms are grouped into *Features, Qualities, System Models, Interaction Models and Behaviours.* These terms are used throughout the main text of this thesis and form the basis for design strategies presented in chapter 3. These terms are contextualised in chapter 2.

1.4.1 Glossary of Terms

Features

- *User number*

  Individual, group or distributed network.
- *User role flexibility*

  Collaborative interface supporting different interaction models for participants supporting learning and efficiency (Jorda 2004))
- *User mapping*

  Parameter mapping by user
- *Adaptive mapping*

  System/interface re assigns parameters during live interaction (Livingstone 2005)
- *Direct motion tracking*

  Performer/participant proximal interaction/body gesture (Tanaka 2004)
- *Indirect motion tracking*

  Relative or proximal tracking of performers/participants.
- *Direct gesture acquisition*

  Data collection through integrated sensors. Typically, these are mounted on an instrument or embedded in a novel controller.
- *Indirect gesture acquisition*

  Analysis of sound output to determine gesture based on knowledge of instrument properties or playing context.
- *Live synthesis and/or diffusion (Transformative interactions)*

Encourages participants to *play intercept, change or make contributions* (Bandt, 2004)

- *Intended audience/end User*

  General Public

  Performer/Musician

  Novice (individual)

  Novices (group)

<u>Qualities</u>

- *Gestural Vocabulary* (Mathews 1984)

  Interface supports wide range of established and new musical gestures.

- *Nourishing* (Machover 2002)

  System/interface motivates continued discovery and creativity through audiovisual feedback

- *Dynamic Context* (Winkler 1998)

  System/interface motivates player/participant spontaneity within an evolving musical framework.

- *Player Paradigm* (Rowe 1993)

  System exhibits player like behaviour or machine musicianship.

- *Self-Organising* (Blackwell 2004) (Whalley 2005)

  An open or conversational generative framework; providing sophisticated interactions between people and improvisational systems.

System Models

- *Performance* model (Rowe 1993)
  Assigned interpretative role players/participants with reactive music.
- *Instrument* model (Wanderly 2000)
  Extended conventional instrument with additional sensing for solo performer – *explore-ability*
- *Unencumbered* model (Cammuri 1995)
  System uses video analysis for data acquisition (*direct* and *indirect* gestures within controlled environment performance/installation).
- *Compositional* model (Winkler 1998)
  System designer/composer provides compositional framework, typically mapping performance data to interpretative musical parameters for performer/participant interaction.
- *Distributed* model (Weinberg 2002)
  Novel interface forms part of a collaborative network manipulating shared content.
- *Collaborative* model (Jorda 2000)
  Typically, a tracking system monitors generic *Tangible* objects manipulated on a table surface with symbol encoding for parameter mapping of group interaction.

Interaction models

- *Control* Supports direct manipulation of musical output parameters.
- *Sequential* Supports linear ordering of defined musical elements, typically samples or events.
- *Organisational* Supports non-linear restructuring of defined elements or events.
- *Relational* Motivates musical relationships between objects or symbols through manipulation and structuring of elements.
- *Conversational* Motivates and sustains a musical dialogue between system and performer/participant.
- *Transformative* Extends musical dialogue by adapting content and processes in response to identified behaviour.

Behaviours

- *Exploratory* (Wanderley 2002) Performer/participant discovers new gestures/interactions during use.
- *Interpretative* Performer/participant assigns system actions/events to personal context/goal.

- *Transformative* (Modler et al., 2003, Thiebaut 2004)
  Performer/participant or system mediated abstraction of musical material.
- *Sociable* (Miranda 2001)
  Learning and memory modelled in software in relation to evolving repertoire. Either a system or performer/participant attribute.

This structured collection of field specific vocabulary consolidates these terms into identifiable groups.

The next Chapter, Adaptive Social Composition, establishes a compositional framework; referencing related works that provide a sophisticated range of socially mediated musical interactions. Conceptual approaches to scoring interaction processes are also presented. The field specific terms identified in chapter one are used to articulate the core elements of Adaptive Social Composition. The questions identified earlier in this chapter are further unpacked by considering specific examples that exhibit one or more related features.

Chapter 2

Adaptive Social Composition

With human-to-human interaction, there are already sophisticated structures in place that mediate interaction. Between improvising musicians the type of instruments played, prior musical experience and training, performance style and genre, even the venue and acoustics all contribute to the compositional process. These factors are not easy to isolate and identify, however equally well-respected compositional approaches have challenged these contextual structures and our responses to them.

John Cage's Imaginary Landscapes No. 4 (1951) designed for 12 radios and 24 performers where, positioned in pairs, each performer responds to a score making finite adjustments to either the tuning or volume knob. Each of these actions is meticulously scored, however the original tuning and volume parameters of each radio are randomly set removing choice or *interpretive response* from both the composer and the performers. The compositional strategy can be considered as an example of the *performance* model with specific scored events, however this controlled behaviour is contradicted by the unpredictability of the broadcast sources, structured compositional actions with integral *indeterminacy* is a feature. Alvin Lucier's *I am Sitting in a Room* composed in 1969-70, is one of Lucier's earliest mature works. It is equally distinctive as a compositional strategy; the room itself and its acoustic properties are the instrument.

The score comprises the following text:

I am Sitting in a Room different to the one you are in now.
I am recording the sound of my speaking voice and I am going to play it back into the room again until the resonant frequencies of the room reinforce themselves so that any semblance of my speech, with perhaps the exception of rhythm, is destroyed.
What you will hear, then, are the natural resonant frequencies of the room articulated by speech.
I regard this activity not so much as a demonstration of a physical fact, but more as a way to smooth out any irregularities, my speech might have.

(Lucier, 1995 pp.98)

The text provides the compositional framework through a set of simple instructions explaining itself to the performer, the performance can take place in any venue using the same method of recording and playback of the original spoken score and the consequent degradation of spoken content and revelation of the 'instrument's' resonant frequencies, a *transformative* process. The clarity of the score is a counterpoint to the subjectivity of its performance. Many recordings of various performances exist. Of Lucier's own published recordings the 1970 version features 15 generations while the 1980 version presents 32 generations the final line refers to Lucier's known stutter. This compositional strategy deconstructs the *instrument* model, introducing the concept of the local environment parameters (the room) being the instrument. This changes the way the performance is mediated; the performer cannot aspire to a virtuoso performance (*instrument* model) because after the initial spoken input the instrument effectively plays itself.

Both these examples emphasise a compositional strategy that manifests a mediated process outside the control of the composer, performer or listeners. The former implements a structure over random material and organises the behaviour of performers to structure its delivery and removes musical interpretation. The latter captures contextually sensitive unique source material and transforms it through iteration by the influence of environmental parameters in which it is manifested. Neither piece require conventional musicians, instruments or technologies to be performed, both provide compositional frameworks where sound material is mediated in real time. The process can be easily recreated but the results will always be different generating new material. An archive of the original piece can also be accessed online (Web reference 33)

The complexity of systems that mediate interaction between people and collaborative systems should not be underestimated. Parameter mapping between an interface and a virtual representation merely connects an action to an event. The context of this action informs our response and this is not easy to identify or model. Composition that is collaborative between a technology and participants requires a shared context with events that can be identified with behaviours that mirror our current understanding of these events.

"A technology is interactive to the degree that it reflects the consequences of our actions or decisions back to us. It follows that an interactive technology is a medium through which we communicate with ourselves – a mirror."
(Rokeby, 1995 pp.133)

Many composers have investigated the potential of evolutionary algorithms for generating compositional material (Moronia et al., 1999) exploring new musical events. Architectures that implement AI within a musical context also cast new light on the previously hierarchical roles of composition and performance creating new interdependencies (Cammurri, 2000). Often these approaches are described as interactive. Normally this interaction takes place through evolutionary algorithms and includes some editorial or selection process by the composer at different stages in the evolution. This interaction is enabled through compositional tools that allow different properties to be set and parameters mapped before, during or after the evolutionary process, both to generate score material and to synthesize new sounds (Miranda, 2001). In essence, the adaptivity is between virtual elements within a software driven process to generate new musical material with human mediation. Ultimately these innovative software approaches apply real world models of evolution to new musical processes through the creative application of evolutionary techniques (Dahlstedt, 2004).

The nature of agents and agency plays an increasingly sophisticated role in contemporary composition. The term *sociability* has been discussed in terms of the behaviour of musical agents (Miranda

2001). In this research, sociability was seen as a desirable feature with the level of an agent's *sociability* evidenced by its *repertoire*. To be sociable an agent's repertoire would be similar to that of its peers thereby providing a motivation for agents to *learn* and exchange *tunes*. Motivational strategies are equally important for human participants interacting with Adaptive Social Composition Systems to initiate *learning*, motivate *exploration* and reward *memory*. Without a specific score evident to structure performance, motivational cues are required to invite and sustain interaction. Identifying the responses to these motivational cues is a significant challenge. The environment itself can be considered to be an autonomous instrument, where a performative action or human input can be evolved to form new material and compositional relationships not present in the source material. The performance or sound environment can also play a significant part in mediating the way that interactions evolve by demonstrating the behaviour and attributes that participants can *learn*, *explore* and commit to *memory*. It follows that *sociability* is a desirable feature for the system to identify and reward. Therefore, behaviours that can be identified and interpreted either through *proximal* tracking of participants or *direct* and *indirec*t gesture acquisition are significant elements of the compositional process. The ability of the system to differentiate between *local* or individual actions and *global* or group actions is essential to develop effective social behaviours and motivational cues.

The convergence of compositional systems that integrate live processes with sensing environments for collaborative output has led to many new creative forms. A feature that these forms have in common is that they are human centred and often cooperative, this presents significant hurdles to overcome:

"The design of MIEEs (Multi-sensory Integrated Expressive Environments) is challenging and many research issues have still to be faced. For example, systems must be endowed the capability of interpreting performers' gestures, and in particular expressiveness in the context of where and when the gesture is performed. A MIEE should keep into account of spatial, temporal and content memory." (Camurri, 2004 pp.1)

In developing such systems, one has to establish a method for identifying gestures and establishing appropriate descriptors or *cues* for them. Algorithms need to be designed to encode or decode known gestures into system events. There has been considerable work in this area and established gesture libraries and externals exist for the eyesweb and MaxMSP software environments (Web References 7, 10, 12, 19.). Analysis of these system events in relation to user behaviour or response should be implemented to allow new gestures or behaviours to be integrated. Again numerous software externals have been developed for MaxMSP which provide specific functionality. An efficient way of training new gestures is to encode actions into lists by capturing data from a novel interface and sorting these lists to represent different gesture classes. These lists can then be learned, stored and recalled using a simple training process. A specific external is provided by Robinson, R. *MXJ nnLists* (Website reference 56.), this provides an artificial neural network with simple input, training and comparison of incoming data. Multiple instances of this process can be used simultaneously, a significant feature allowing a gestures to be trained by integrating data from a group of sensors. In principle, any data sequence can trained and applied to a wide range of compositional events, however identifying the start or end of an action encoded into a list requires a broader conceptual approach to interaction design. It is equally important to establish levels of influence or layers of interaction by relating different gestures to different events with appropriate content transforms. These principles are dealt with in more depth in chapter 3: Design strategies. In the following section the broader conceptual approach is unpacked.

2.1 System Analysis

As discussed earlier, a number of approaches exist that engage a performer or player in a process of shared composition. Usually there is a clearly evident interaction model and a tangible exchange between the performer and some software mediated Computer Music process: score following, signal processing, algorithmic or generative responses to human input. However, the challenge with adaptive composition is to create a live environment that adapts to repeated or new compositional events through collaborative exchange. This implies that a learning process is essential, where initially the software element is trained through exposure to a number of gestures or relationships that can be identified and mapped to potential responses or musical actions. These processes are more often evident in musical systems that are modelled in virtual communities or developed from neural networks. Patterns of behaviour are revealed through the interaction of virtual agents who individually have limited musical potential but collectively can learn, perform and develop new musical behaviours. A key difference with the compositional process that forms the focus of my research is the emphasis on social interaction between people, adaptive interfaces and a mediating framework; participants are not required to have any formal musical training. This means that the system has to have the capacity to learn and respond to new interactions from people engaged with it. The compositional process does not take place within a self-contained simulation, it is manifested within a collaborative 'Cybrid' (Anders and Livingstone, 2001). A *control model* provides a limited framework and lacks the collaborative exchange between players and a system, however it is useful in terms of training as a novice participant intuitively adopts this mode of interaction as most music software is designed around a control model; logical responses or feedback can be given in response to recognized actions. The first problem is how to change this balance of control, to move from a *control model* to an *interpretive* one. For the system to be adaptive it needs to go beyond a call and response model, it needs to influence player actions and contribute to the compositional process. Logically the player is expecting feedback to an action to communicate the system response, for example, changing a property of an existing sound or restructuring an event. It is only a small step to enable the software to modify this response. This small step is very significant, as it must be perceived as an intentional action not a 'bug' or fault in software operation.

Social groups of human players interacting musically have the benefit of past experience, knowledge of each other's musical interests, styles or genres and through attentive listening during play they can easily change the direction or emphasis of the exchange by adapting their behaviour. So, with an adaptive social composition system it is equally important that the system has this potential. By observing a call and response model one can identify an element of gamesmanship, creating a move or action that challenges the other player to change their behaviour, to extend their response. This approach is useful as it informs the way the player/system interaction can be extended. The system needs to recognize and respond to gestures logically, but over time or with repeated actions on the part of the human player, the system should adapt its responses to challenge the player, to motivate a new action, evolving from a *control* model and establishing a more adaptive response, a *transformative model*. Clearly, for this to be possible the system needs to have past experience of such exchanges and a strategy needs to be in place to allow new responses and musically perceivable actions to evolve. Experience in terms of software mediation of events is learnt through either a training process or exposure to new events that can be attributed to new behaviours. This implies that the system needs different forms of memory and the ability to relate actions or events to each other, whether initiated by the system or a human player. So a compositional framework needs to be in place that links a virtual or software based framework for these exchanges to take place with a physical environment for human players to engage with where both parties can perceive the influence of each other on the collaborative output. The design of the environment itself becomes a significant factor in the design of the compositional process. In a practical sense, the properties of this environment will need to manifest a series of musical events that are significant to both parties, that relate to each other's model of the environment. Of course, there are well-established frameworks for exchanging musical information between people. The musical score has many forms but is effectively a tangible framework that encodes actions, gestures and relationships for storing, sharing and performing musical material. Within an adaptive composition intended for social interaction the environmental framework and underpinning software infrastructure becomes the score, for example; when we experience the environment, we are able to identify a number of possibilities and choose how to enact them. The environment also shapes our responses and modifies our behaviour. As we turn the page, new material unfolds. This encoding and decoding of musical actions and intent bound together by a process of exchange in a live and ever changing sonification of collaborative interactions poses a significant challenge in terms of overall design. A modular approach to development, that implements different interaction models that are perceived by players and tracked using novel interfaces supporting a range of interaction models in a shared context is a logical approach. In fact, it is probably the only approach that will allow the different process to be revealed. It is also far easier to simplify this complex process into individual but complimentary elements that can each be modified, trained and given varying degrees of autonomy as the compositional process evolves, and clearly, for the collaborative process to manifest adaptivity or motivate new behaviours in players it must have the potential to evolve. So, the system design must include properties of both the human players and tangible modular elements that combine to create a shared musical process. An improvisational learning environment that sonifies gestural interaction develops from previous musical exchanges. Such a system design will also need a series of strategies to motivate and engage new players in this shared process, making tangible relationships or explicit responses to actions that evolve over time, or identifying new behaviours; moving beyond the *interpretive model* of interactive music to a *transformative model* where conscious and subconscious interactions are part of the compositional process.

A number of approaches have been identified for this thesis, by observing the use of, and interaction with, novel controllers and custom or modified instruments. Representative selections of these devices have been reviewed during conference attendance at NIME (New interfaces for Musical Expression) Vancouver 05 and also at ICMC (International Computer Music Conference) ’04 Miami and ‘05 Barcelona. Many of these devices were presented publicly and made available for delegates to experience first hand through workshops, demos, concerts and scheduled improvisation sessions, complimenting the formal academic delivery. Many of these approaches are well documented through published papers but in the majority of cases these interfaces follow a *control model* and do not manifest any adaptive features through software mediation between people and autonomous interfaces. A proportion of these interfaces are designed for use with *interpretive* interactive music software, fewer examples exist where unpredictable data, such as a camera feed of the audience affects synthesis parameters during performance, although there are notable exceptions. Installation based sound works tend to take more account of audience interaction and cater for more varied forms of interaction either through positional tracking, gestural mapping or object placement where the relationship between elements is influenced by audience presence, location or actions manifesting a *transformative process.*

The field of tangible interfaces has some excellent examples where social interactions with an evolving soundscape are mediated through a 'Freesound' model (Livingstone and O'shea, 2005). Motivational strategies become very significant with these systems which are designed for 'hosting' a shared composition, as with the performance featuring two *Reactables*, one in Barcelona and one in Linz, shown at ICMC 2005 with two performers manipulating objects on both tables remotely. The *Reactable* team, led by Sergi Jorda (Jorda, 2006), had begun to adopt motivational strategies to include exploration of their system by conference attendees. The conference theme was 'Freesound' and as part of this event, delegates were invited to contribute to the 'Freesound Project' initiated by Bram De Jong (Website reference 34). Jong (Website reference 34). This required delegates to produce a sound and upload it to a shared database of sounds for use by anyone at the conference. This generated a wealth of diverse material in an efficient online-shared resource. The *Reactable* (Jorda, 2006) team then generated an individual tag or symbol for the reverse of each delegate's conference badge which could be tracked as a control object by placing it on a *Reactable.*

(fig. 4 Reactable with delegate tag using symbol identification)

Each badge was assigned to a sound from the Freesound project as the base material from which a new composition could be formed by moving other tracked objects or symbols in relation to it to modify effects parameters.

What is significant in this example from the perspective of Adaptive Social Composition Systems, is that it provides a method of integrating an audience member, motivating them to contribute and interact. It gives them focused feedback in the form of their own ‘unique’ sound and encourages them to explore its properties and relationship to other sounds through simple intuitive interactions, for instance, moving it in relation to other sound mediating objects which are also heard simultaneously. This apparent simplicity is enhanced as others join the process, adding additional sounds and exploring the various features of the system. There is a degree of adaptivity evident in the human agents, interacting with the system and modifying their sound in relation to other participants and also an element of social mediation as turn taking, call and response and other social exchanges or collaborative models are explored. The system itself is only designed to facilitate these interactions, it does not autonomously influence or mediate the shared material or respond to interaction beyond a *control model*. It applies logical mapping of performer or player actions through tangible interface objects to musical parameters.

Other systems exist that follow an equally social interactive model where the sequencing of a simple sound or sample is shared collectively. These systems are quick to learn, responsive and enable participants to move quickly through exploration to control of structured sound material. Probably the best-known example of these systems is the *BeatBugs* (Weinberg, et al. 2002). These handheld tactile wireless interfaces have been used in a series of performances with groups of children who quickly learn and explore new ways to create percussive patterns collectively. Again, from an Adaptive Social Composition Systems approach the significant factor is the emergent behaviour manifested by the children while using the *Beatbugs*. One interaction described by research students involved in the project was the observation of passing a sound from one to another. In terms of the *Beatbugs,* this event was simply a timing issue to trigger the event. The performers interaction only required the correct sensor to be activated in time with other participants. However, the children exaggerated this with an intentional gesture of the *Beatbug* in the direction of the next participant and as the performance was refined, this action became exaggerated, described as a swooping motion. It is no surprise that a shared rhythmical pattern generates movement in participants, and of course, many hand held traditional percussive instruments exploit movement directly as part of the performance. But the *Beatbugs* were not designed to respond to this emergent behaviour. What we learn from this is that technology has the potential to motivate new interactions. The human element will adapt to a technology and thereby change its context and demand more from it than was originally anticipated. Therefore, an element of adaptivity within such systems is highly desirable. The ability to recognize new patterns of interaction or new performer behaviours is one way to achieve this. Tasks that explore mental models and make them 'visible' or which sonify relationships dynamically, provide effective reinforcement.

Other interaction models can also be used effectively to build a more engaging dialogue between performers and systems. Different training strategies exist within related learning systems. For example, if we ask a performer to whistle a middle C, software can identify the pitch and a screen prompt could indicate higher or lower until the performer hit the note. This simple process engages the performer in a task that the software can measure and respond to. Designing abstract tasks, that can be performed and monitored in real-time, enables a higher level of integration and response between performers and a mediating system. A more challenging task is to relate sound properties of discrete sound objects to one another. In this scenario, we are able to create new ways of relating sound elements that both performer and system can repeat, model and apply at a later stage.

2.1.1 Technologies, Strategies and Methods

Drawing on documented examples of novel interfaces or new digital instruments identified through review of published works on related systems and through attending technical demonstrations during field specific conferences, the following technology, strategies and methods were identified.

For multi user or social interaction with a degree of control or influence over sound creation and manipulation of their own and other collaborator's sounds a networked or shared interface can be highly effective, supporting collective control through *user number* and *user-role flexibility* (Jorda, 2005). The design of the interface itself is very significant and needs to apply ergonomic design principles. For non-performer interaction, video tracking of gestures or participant location does not always lead to meaningful interaction. A degree of influence may be experienced but this 'black box' approach tends to alienate participants and does not motivate musical collaboration easily. Handheld interfaces are far more effective in translating an action or gesture into a sound event. Our intrinsic motivation to select, control and place objects in relation to one another is an intuitive way of exploring or mapping relationships. A compositional approach that provides a balance of tangible control with an evolutionary approach to learning or relating new relationships also has the capacity to motivate emergent behaviour. The physical size and shape of an interface is a major factor in the way its intended use is perceived and also in how it is held. Unencumbered interfaces are easier to distribute and exchange during group interactions and lead participants to explore new processes while taking responsibility for their element or contribution. Individual contributions and different gestural expression should be recognized by the interface or control software, further motivating exploration and refinement of musical process, learning and exchange. The closest example that supports these interactions is a hand-held percussion or simple wind instrument where hand held independent objects with intuitive finger placement and player orientation vary the sounds produced. So, the interface design for a social composition system can be extrapolated from this model. Independent wireless handheld objects that are sensitive to different users motivate individual exploration while enabling social exchange. Physical orientation and the relative location of objects should have a perceived relationship that can be modified or influenced by the user through direct control or object placement. To motivate new interactions' objects should have tangible feedback or ways to indicate status. Emergent behaviour can play a significant part in extending object use and parameter mapping if there is a means to identify new gestures created by participants. Similarly, behaviour of objects can be used to motivate or reconfigure participant interactions. Adaptivity is highly desirable within such systems if shared composition rather than score following or call and response exchanges are envisaged. Tempo plays a significant part in building relationships between elements and allows participants to identify discrete elements more effectively. The ability to control sound placement as well as tempo in response to gestural control motivates further interaction and enables individuals to distinguish their contribution from that of others and to adapt their approach in relation to other elements. A software method for capturing known gestures and learning new gestures is required. A method for delivering a mental model of compositional elements to participants is highly desirable, where simple tasks can be used to focus user interaction and control through cognitive tasks related to sound control that demonstrate interface adaptivity. Sound synthesis is a significant factor in facilitating shared composition. A method for creating new sound material or sound objects that can be individually placed through live diffusion will enable users to identify, control and transform each element. Certain compositional genres lend themselves to this type of process, e.g. music concrete, electro-acoustic composition or live experimental forms of electronic music performance where musical material is generated in real time and mediated by tangible interactions.

A number of significant studies in relation to real time interaction have been published within the field (Lippe, 2002) documenting fundamental *principles and paradigms* (Wessel, 1991) for multi user music interfaces. The number of participants and the differences between online and local networked instruments are detailed in Weinberg's Taxonomy (Weinberg, 2002), which also documents methods for organizing group structures or possible topologies for shared performance. A classification space for *remote and co-located* and *synchronous or asynchronous collaborations* for computer-mediated interaction was proposed by (Barbosa, 2003). The collaborative experience of novices in relation to *play-ability* and *learn-ability* has also been studied (Blaine and Fels, 2003). Of these studies, each offer insights into specific forms of group interaction where either software or a novel interface is provided to facilitate musical collaboration. Solutions are offered for structuring interactions across distributed groups using networked software or for organising groups of novices using novel interfaces. Key terms identified in chapter 1 feature in these studies but the terms are not explained or grouped into accessible language outside each specialist field of enquiry. The form of co-creation proposed as Adaptive Social Composition (combining qualities, models and behaviours to mediate novices, adaptive interfaces and a shared compositional framework) is not presented in these studies.

In chapter 1, the following questions were identified to unpack the core elements required to establish an adaptive social composition framework. From the literature review, explanation of terms and discussion so far, we can begin to address these questions:

a) What features of the process of composition can be distributed between novice participants?
   Pitch: identification/modification (<u>control model</u>)
   Timing: Event triggering, turn taking (<u>sequential model</u>)
   Diffusion: Positioning sound objects (<u>relational model</u>)
   Influence Sound properties: simple transformations in context (<u>transformative model</u>)

b) Which strategies can be used to invite and motivate interaction?
Enable different interactions through same interface (user-role flexibility)
Provide musically relevant feedback to user actions (reciprocal)
Provide variety of gestures in defined classes (gesture library)
Promote learning and exchange of gestures or sound material (sociability)
Develop visible interface behaviours (absorb and adapt modes)

c) How can non-musicians be engaged in meaningful musical exchange?
Combine user-role flexibility with proximal interaction
Combine conscious and subconscious interactions so new musical behaviours can be learnt

d) What is the nature of collaboration within such a process?
Playful, exploratory, turn based, (self organising).

e) Where do these interactions take place?
Within a shared framework, a balance of gesture capture, interaction design and shared composition

f) How can one go beyond a *call and response* model?
Motivate co-creation (evolving rule set) through conceptual framework of identifiable qualities, models, and behaviours.

2.1.2 System Requirements

A tangible adaptive social composition system will need a minimum of three wireless interfaces to motivate group interaction. Each interface will need to incorporate sensors for capturing both *local/direct* user gestures and mapping *global/indirect* location of each interface. Each interface will need to encode analogue or digital data from a range of sensors and transmit this data wirelessly to the composition software for processing to mediate real-time performance (Bongers, 2000). The composition software will need a method for tracking the relative positions of interfaces and mapping the gesture acquisition data to sound diffusion in response to user *local* control or *global* influence. Ideally, composition software will also need to modify the position or orientation of interface elements to motivate participants, interaction or to demonstrate functionality through autonomous interface behaviour. The system adaptivity must be tangible to participants combining auditory and visual feedback to reinforce user or system created relationships. Each interface must be able to operate autonomously in response to other interfaces if there is no participant. This autonomous behaviour should not be disruptive but could draw on previous interactions from an evolving gesture library. Environmental parameters have an influence on participant interaction, and interfaces should have a similar adaptivity implemented to respond to environmental parameters to influence sound material or compositional process in a coherent context experienced by participants. The sound material created should motivate shared composition and not contain extended pre-structured sequences or fixed samples. Ideally, sound material should be perceived to be in context of current interactions. A distinct range of frequencies and textures can be generated as base material for shared composition that is related to current system activities and influenced by participant gestures and motion behaviours.

"There is no fixed ordering to the human- computer dialogue. There is no single permitted set of options (e.g. choices from a menu) but rather a series of continuous controls. There is an instant response to the user's movements. The control mechanism is a physical and multi-parametric device, which must be learned by the user until the actions become automatic. Further practice develops increased control."
(Hunt et al. 2000 pp.210)

Hunt and colleagues efficiently summarise the ideal requirements for a system where feedback is perceived by the participant in real time as an integral part of the composition or performance process of new gesture based systems (Hunt et al., 2000). Human computer dialogue is presented as a shared process with an element of learning or extensibility. These requirements are equally relevant to adaptive social composition systems, although as we have discussed previously strategies for facilitating and motivating this dialogue are also essential. Equally important are the inclusion of algorithms for controlling musical properties of mediated sound objects such as *pitch relationships* and *time scaling* (Arfib and Verfaille, 2003) that can motivate participant behaviour. Karl Heinz Essel has provided a range of compositional tools referred to as the *Real-Time composition library* or *RTClib* (Website reference 58) that includes such algorithms. This collection of software externals adds compositional functionality to the MaxMSP environment and is ideally suited to mediated frameworks were classes of identified gestures are assigned to compositional processes designed to provide musician like responses to the actions of novices.

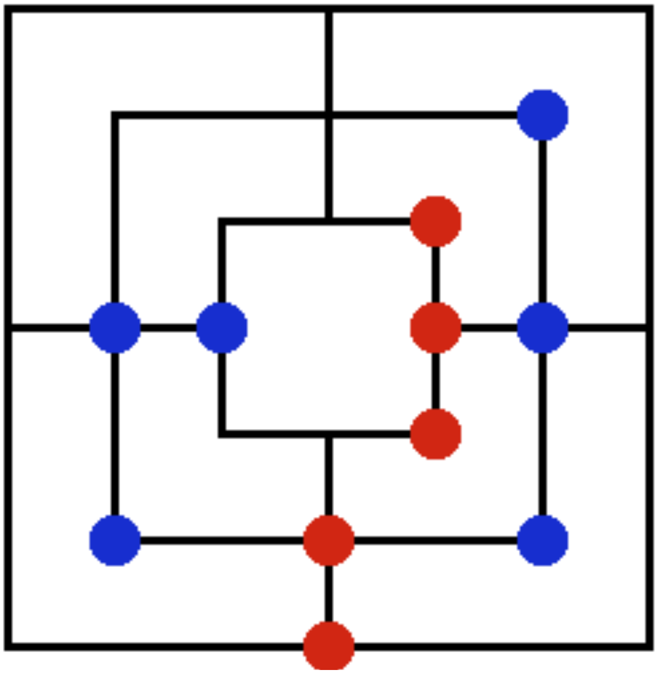
2.1.3 Summary (section 2.1)

This section introduced distinctive scoring approaches by composers Cage, and Lucier whose compositional practices feature embedded role or task assignment as compositional process. *Instrument* and *Performer* models were discussed in relation to these examples developing a theoretical framework for adaptive social composition systems. Further features representative of adaptivity within related research were considered: *sociability* between agents, and the concept of agent memory or *repertoire* as a motivational factor within collaborative composition systems. The design of handheld controllers was considered in relation to these factors with discussion of related interfaces.

The following section considers behaviours that can be identified and learnt through simple gameplay interactions using a traditional board game to illustrate learning proces and identify terms used to extend design strategies for adaptive social composition.

2.2 Conscious Interaction

Conscious interaction can be motivated through a set of rules with a known parameter space. Traditional board games also provide some answers to the problem of communicating the rules or structures that facilitate interaction within a shared space. Only the simplest board games can be mastered after reading the rules. It is the experience of playing that allows the complexity and flow of interactions between players to unfold, or new strategies to be identified. Observation of this process and the level of information required to reveal each new relationship is a productive strategy as the turn taking and unfolding nature of certain board games support multi user interaction. Even within competitive board games for two players, a range of interaction models and behaviours can be identified. The process of teaching or learning a board game through example and experience is not unlike teaching or learning an instrument, and includes a basic introduction to the parts of the instrument and their function, how to hold or manipulate elements and how to group actions together to make new events. It is no surprise that many tangible interfaces for Computer Music or sonic art have been created using object placement, symbols, movement, and proximal relationships for mediating musical material by non performers, enabling intuitive interaction and collaboration to a greater or lesser degree. Whether the choice of these formats has been deliberate on the part of designers, or subconscious based on a previous mental model of structured play is unclear, but the popularity of table-based systems with novice users is increasing. These tangible interfaces enable musical interaction without having to learn conventional music software while providing a face-to-face or *social* context for musical communication. Let us consider one example board game and consider the frameworks evident. Our choice should not be too complex with multiple moves and complex playing strategies (chess), and it should not be so simple that it provides no opportunities for players to adapt new behaviours (snakes and ladders).



(fig. 5 Nine Men's Morris)

Nine Men's Morris (figure 5) is unique in format as play is structured around lines and intersections as opposed to an 8 x 8 space or chequered board. The play area usually consists of three squares drawn one inside the other and evenly spaced in two dimensions. The corners of each square or connecting points are used for placing counters. Each side of a square is divided in the middle by an additional connecting point. This additional point is connected to the adjacent square's equivalent point. The game follows the conventional turn based two-player format of many competitive games. The premise is simple: capture/remove your opponents pieces by making lines of three of you own pieces. You cannot take a counter that is already in a line of three; you can only take a piece each time you make a line of three. Each player begins with nine counters. Play unfolds in three stages. *Placing*, *Sliding* and *Jumping*. Pieces can be taken throughout the game; play ends when one player has only two pieces and can no longer make a line of three.

• ***Placing***

Players take turns to place counters on connecting points attempting to place three in a row and/or block opponents from placing three in a row. A turn consists of placing a counter on an empty point. At this stage, each player is exploring the game space and observing the placement strategy of their opponent (the interaction model is *exploratory*). Players soon learn that there are optimum ways to place counters if they want to secure a line of three. The simplest strategy is to place two counters on opposing corners of the same square. Placing a third counter on a further corner of the same square creates two opportunities and cannot be blocked. Players quickly learn to *anticipate* this pattern and block it before the third counter is placed. Once the game space has been understood the process becomes more strategic (the interaction model is *organisational*). A second common strategy is a 'T' shape created across the intersecting line between squares. The player places a counter on either an end point (corner) or intersection (centre). If the opponent does see this pattern, the third counter is placed on the intersection adjacent to the corner creating two placement possibilities to make a winning line.

• ***Sliding***

Once all counters have been placed, play shifts from *organisational* to *relational* and players continue to *attack* (create line of three), *counter attack* (block opponents) or *distract* (create patterns or opportunities for multiple lines) by alternately sliding counters along lines to the next available intersection or corner. A turn consists of sliding a counter to an adjacent empty point.

• ***Jumping***

When a player has only three counters on the board they are no longer restricted to sliding moves, they can jump to available points. This effectively turns the tables as the ability to jump makes it far easier to complete a line of three, and the player with the most counters can no longer easily block due the increased freedom of movement. To win they will need to *distract* their opponent by creating potential line patterns that their opponent will have to block to stay in the game forcing their opponent to *anticipate* potential moves. If the player with three pieces plays well it is still possible to reduce their opponent to an equal three pieces where the game comes full circle.

What this shows is that a balanced rule set within a structured framework can motivate a range of behaviours. Some are predictable actions that players learn; the environment determines others as the game stages evolve causing players to adapt. This particular game also has short and longer-term strategies that can be brought to bear once the opponent's skill has been assessed. Further rules can be introduced to refine gameplay.

This game has been taught to many children who quickly learn how to play. A typical approach is to introduce each sequence of play, and after 5 to 10 games the child will have learnt the basic rule set and have begun to notice repeating patterns that ensure a winning line can be made. Of course just as there are stages of play, there are levels of complexity. But within this *sequential* framework a number of *transformative* processes in terms of player understanding are revealed. The teaching role requires two states, one where examples are given in the context of the junior player's understanding, engaging the player while allowing them to *absorb* new interactions, the other where challenges are provided by playing just beyond the ability of the new player and observing their response and actions, *adapting* the playing style to provide a learning experience. The objective is to sustain engagement and provide a learning experience that motivates interaction, mediating the process. 'Nine Men's Morris' was identified as a basic structure for implementing proximity and movement behaviours within the author's design for a group of three adaptive interfaces called Orbs (technical details of the final prototype for this novel interface is provided in chapter 4, Proof of Concept. Initially, these Orbs were developed using modified line tracker robots implementing three motion behaviours related to game stages to engage participants. The stages *place, slide, jump* were translated to *lead, chase, avoid* using computer controlled remote control cars to explore motion behaviours as part of an evolving rule set. Two modes were identified for providing a learning framework for novices interacting with these Orbs, *absorb* and *adapt*. This is explored in chapter 3, Design Strategies). As a design strategy the behaviours *place, slide* *and jump* are useful in terms of dynamic tangible interfaces where the relationship between placed and moving objects is an integral part of the interaction. Innovative touch-screen interfaces such as the Lemur (Website reference 35) are marketed as customisable control surfaces, and these systems allow the familiarity of analogue control surfaces with the flexibility of a sophisticated tactile visual interface. Although such interfaces are ideally suited to versatile personalized parameter mapping for individual performance there is the potential for complex adaptive collaborative behaviour to be supported through a coherent design strategy developed from an understanding of gameplay models.

The board game chess provides far more challenging problem solving and presents a number of psychologically mediated variables (Hartson and Wason, 1983). The level of complexity introduced by multiple objects or playing pieces with varied patterns of movement provides significant proximal and distance related problems. This in turn extends the range of possibilities on every move with sequences of moves planned and anticipated well before they are actually made, depending on the experience of the players. Part of the art of chess is to *lead* your opponent to make moves they believe to be their own, to threaten your opponent by *chasing* valuable pieces to anticipate your opponent *avoiding* traps, forking moves and other subterfuge.

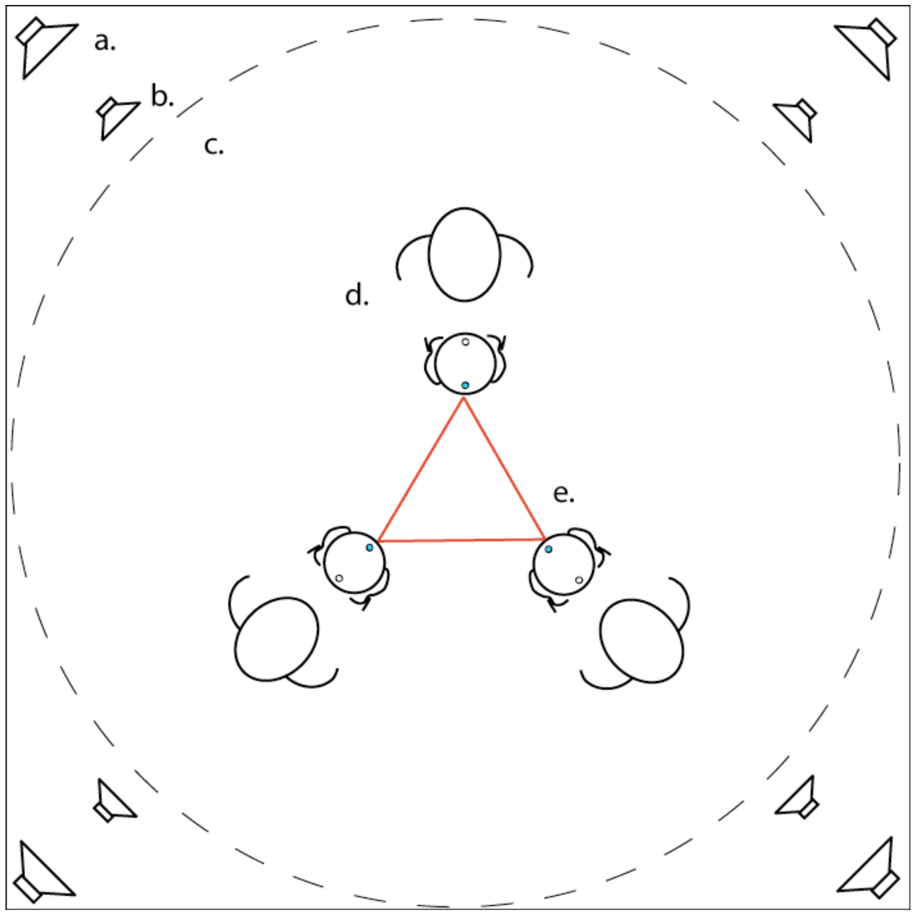
The spatial orientation of pieces and relative positions over time, become far more sophisticated, and proximal interactions are harder to anticipate due to the combination of *placement, sliding and jumping* all possible depending on the piece chosen. These types of moves or events are remarkably similar to the embodied actions in instrument playing, encoded through a *score* and manifested through performance.

It is important to note how a novice player can be introduced to a potentially complex set of interactions within a shared framework. The individual moves of the pieces can be easily learned, the way they interact can be unpacked as a game unfolds offering choices and the potential to plan more complex sequences. New *proximal relationships* are discovered. The actions of players are identified and categorised in context by the opponent (*classes of gesture)*. These actions can be intercepted and transformed by changing anticipated responses and adopting new playing styles (gamesmanship). This presents an intriguing potential for designing collaborative musical frameworks. This thesis proposes a design approach that combines a mediated environment with novel interfaces and a limited amount of participants. This follows the board game or shared stage model adopted by many so called 'Tangible interfaces' but extends it by introducing multiple layers of *instrument resolution* to extend engagement. This integrated approach provides an evolving rule set to enable participants to *play, intercept change or make contributions*, within a shared composition. Just as in a board game, stages of play, types of actions and context of actions can be designed as a collaborative framework. Participants can learn new interactions, as a process of *co-creation* unfolds. This overall design approach has been referenced earlier and developmental stages are documented in appendix 3: published papers. This Adaptive Social Composition framework was conceived to explore the interaction modes discussed in this chapter extending the principles and stages of gameplay and learning as exemplified by the board game Nine Men's Morris. The Orb3 overall design follows the model of environment as instrument established by Lucier and is summarized here for clarity. In chapter 3, Design Strategies, the Orb3 overall design is unpacked to establish new design strategies for Adaptive Social Composition.

2.2.1 Orb3 Overview

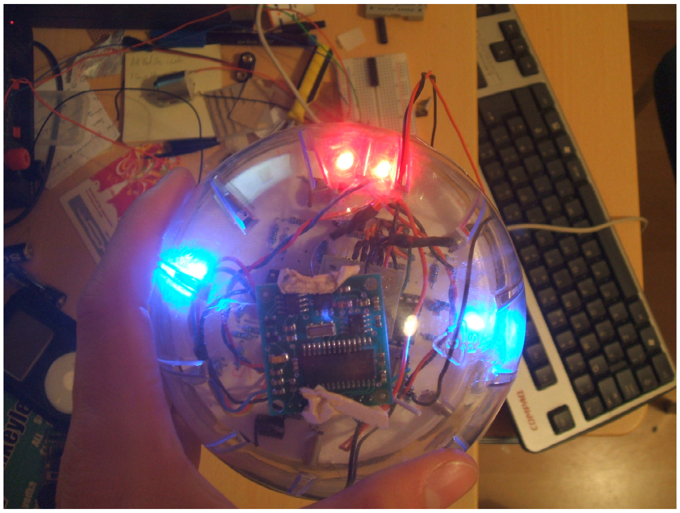
The Orb3 environment design (Livingstone 2005) was conceived as a model for developing a new compositional form, Adaptive Social Composition. It comprises an eight channel sound environment, which occupies a 4m square cube. Speakers are positioned (equally spaced) in each corner of the cube, facing the centre. This installation is a dynamic sound environment for three to five standing participants to interact collaboratively within a diffused real-time composition using ambisonic diffusion.

Interaction is layered, supporting both direct and indirect modalities. Three 200mm diameter PVC balls form a wireless distributed interface, where each ball or 'Orb' interface has a cluster of sensors for gesture acquisition. Each Orb interface can operate either independently of participants, using basic mobility and motion behaviours developed from the *sliding, placing and jumping* behaviours discussed earlier. These are converted into autonomous behaviours by integrating simple robotics into each interface;



(fig. 6 Orb3 environment plan view)

The diagram above (figure 6) shows three participants holding Orb interfaces (d) Their positions form a triangle (e). An overhead video camera tracks these group behaviours and monitors the florr area (c). Two tiers of four speakers are provided at floor (b) and ceiling (a) levels. Speakers are positioned facing the centre of the arena allowing sounds to be diffused in relation to participant positions. This integrated system design is fully explained in chapter 3.



(fig. 7 Orb interface prototype)

A sound object is created and assigned to each interface, as each Orb explores the environment, it's sound object is diffused relative to it's current position This autonomous behaviour is described as *Absorb* mode, onboard sensing is used to influence the diffusion and characteristics of the Orb's assigned sound object. In *Absorb* mode, the motor behaviours echo the interactions of a board game, revealing potential interactions to participants. This follows the model of learning apparent when a child learns to play a game by observing an evolving rule set and interacting with it in stages. When an Orb interface is picked up the mobility behaviours are disabled and the sensing mode is switched for local gesture acquisition (figure 7) . The position of each of the three Orb interfaces and up to 5 participants are tracked by an overhead camera, to enable different behavioural actions and responses in relation to diffused sound objects to be identified, forming an adaptive compositional system.

This integrated environment (fig.5) is referred to as Orb3 since the compositional framework is extended from three Orb interfaces. Orb3 refers to the overall integrated environment design and an individual interface is described as an Orb interface (fig 6). Technical details of the Orb interface design are presented in chapter 4 *Proof of Concept*. Conscious interaction, within a mediated sound environment, relates to participant intent either to engage with a process or motivate an event. The deliberate decision to participate and to affect the compositional process can be shown by physically intervening within the presented sound environment. The first stage of this process can be identified as proximal listening, that is the participant moves to different positions in the space attentive to sounds produced and the location of sensors or interfaces. This first stage of interaction can be detected via an overhead tracking system using 2-dimensional motion detection. It is easy for a human observer to identify this conscious behaviour, however software identification and categorisation requires a resolved strategy. For this *exploratory* action to be registered as significant, within a system such as Orb3 it is not enough for the system to register an additional moving object. It has to map the proximity of this new object relative to current system events. These system events can include several elements such as the properties or number of sounds being played, the location of these sounds relative to the identified listener or participant or the position of Orb interface units relative to this new collaborator. These strategies are further developed in chapter 3, where a controlled experiment is presented that identified different patterns of response in participants, by altering the sequence, location and characteristics of sounds played through an eight channel speaker system. A *call and response* model can be introduced between the new collaborator and the mediating system can be used to deliver a tangible reaction to this new presence. For example, if this new *active listener* has moved towards or away from a particular sound object the system response could be to reposition this sound object, assigning a new diffusion trajectory. This reactive approach can certainly motivate additional directional movement of the participant and the relative position interpreted by software as behaviour. The software library CV.jit (Website reference 12) for Max/MSP is used for analysing video data to identify proximity and motion of participants. These movements can be mapped to known system behaviours as used to mediate Orb interaction, for example: Is the participant *chasing* or *avoiding* the current sound object? Just as a board game may have a set of primary rules for placement and movement of pieces in a *sequential* turn-based model, participants can be motivated beyond a call and response model by identifying potential 'moves' as they become familiar with primary rules. A form of short-term memory can be implemented within the mediating software, comparing a system sound-object trajectory with participant relative proximity to initiate the next move or response. This is achieved by archiving data streams form different sensors embedded in each Orb interface using lists; sequences of numbers captured in real time and written to a text file for comparison with stored lists. Max/MSP provides numerous ways of handling such data, the simplest object to use is MTR (multi track recorder). This object simply assigns each data stream to a channel and records or plays back input data in response to messages. Pattern recognition can be used to compare and categorise data streams as known or new gestures. Examples of how these processes can be implemented are provided in chapter 3. Equally, participant trajectories in relation to sound object proximity can be used to map and compare current and previous behaviours enabling the system to archive repeated participant proximal responses. These repeated responses can be effectively written to long term memory, e.g. if conditions match, the software can store participant response data in an envelope or list for later comparison or mapping to alternate compositional elements of the soundscape as the compositional process unfolds. The delivery and timing of such exchanges is significant if participants are to relate events and adapt to new motivational cues or behaviours created by the system. During the Orb3 environment design process a number of different software techniques were explored to identify user actions and transform musical context. Of these techniques, many are processor or resource expensive, so solutions were sought that optimised or simplified gesture identification or mediated musical contexts when known gestures were identified. Small scale artificial neural networks were implemented using *MJR nnLists* (Website reference 57) externals for MaxMSP. This provides an easy to use framework for comparing incoming data streams with known or previously taught sequences, tolerance can be set so gestures captured by an interface can be identified within scaleable parameters. Previously in this research a simple model of short term and long term memory was used using standard MaxMSP data objects to compile lists from gesture data and replay them when certain conditions were met. These lists can be easily written to text files so gestures or sequences can be stored, recalled and re-applied to other compositional contexts. The Real-time Composition library

RTClib (Website reference 58) externals for MaxMSP were used to provide simple algorithmic transformations to individual sound objects. An alternative solution to recognition, analysis and generation of musical structures and events *artificialtango* (Website reference 56) was also tested. The *artificialtango* library for MaxMSP by Olaf Matthes provides a sophisticated tool for complex audio and MIDI analysis. It is based on the FTM library from IRCAM which requires additional installation. The higher level features add additional layers of complexity that were not necessary to resolving musical exchange between novices. A significant difference in the Adaptive Social framework described is that an alternate process of interactive evolution is established. Effectively the human agent is constrained by the rules of a physical sound environment, but can evolve new behaviours within it. The mediating software has the higher function of selection, identifying which behaviours are relevant and can be used, or archiving newly evolved behaviours in response to system intervention. Another core difference is that the evolutionary actions take place in the physical domain, and are not modelled within a software simulation. The mediating process combines software-identified classes of gesture with modification of interface behaviours to motivate new collaborative interactions.

Generally, scientific research has four goals (Levitin, 2001) description of behaviour, prediction of behaviour, determination of the causes of behaviour and explanations of behaviour. In designing adaptive composition systems, it may not be a primary goal to prove a particular action or sequence of events is related to a specific behaviour, particularly when interaction between people and adaptive software is concerned. Nevertheless, where considerable time and resource is invested in such systems it is only logical to implement experiments that inform interaction design. Of the three established frameworks for establishing scientific studies; *controlled studies, correlational studies* and *descriptive studies,* one has to decide which will be the most productive in a given situation.

a) A controlled study usually measures the difference between two groups of participants, and random assignment of individuals to groups and consistent test conditions are required. One would expect the two groups to be drawn from a larger group of similar background and experience, however the method of recruitment can significantly influence the experiment and care should be taken.

b) A correlational study is useful for identifying patterns of occurrence or correlations between events with a random assignment of participants. Care needs to be taken that other factors outside the experiment do not bias findings.

c) Descriptive studies differ from controlled and correlational studies in that they seek not difference but patterns of information that show how something is. The primary goal is often to describe something that has not been described or shown before, or to show pattern. Typically, a questionnaire is used. User responses are then grouped and analysed for similar response patterns. The interaction design problems within an adaptive social composition framework are related to the likely actions, behaviours and responses of individuals to both system events and in relation to collaborators. So describing the reactions of individuals exposed to the same set of challenges is a valuable way of identifying patterns of behaviour and the prevailing conditions that motivate it.

For the purpose of this research a) and c) were chosen. An experiment was designed to map a participant's ability to locate individual sounds within a cube format 8-channel diffusion environment. This experiment included a controlled test a) using analysis of variance to identify significant patterns in relation to two controlled variables, sound type and sound location, the method and results of this experiment are presented in chapter 3: Design Strategies. This experiment was developed from an adaptive game prototype (PoundtheSound 2006) after observing beta testing of the basic game. The original concept was developed from the cube score exercise presented earlier. Rather than tasking participants to visualize points in space as a conceptual task, or to reposition sound objects to spatially represent coordinates of a *perceptual construct* as a cognitive exercise. A hand held controller was used to track participant behaviours in response to different motivational cues; sound placement, pitch, timing and motion to examine participant response to different sound material and sound placement. The diffusion format of the test environment places the subject in a three-dimensional soundscape, where sounds can be positioned in fixed or moving locations to motivate different responses and identify common behaviours. The original game featured a standard handheld game-pad controller. Using a symmetrical generic controller for initial perception and sound location identification tests constrained subjects interaction and provided task focus for accurate data collection. The participant was asked to listen to isolated sounds and map the perceived location by pressing a button indicative of the emitting speaker location. The participants stood in the centre of a speaker configuration with 8 equally spaced and balanced speakers pointing directly at this central position, four at floor level four at ceiling height within a four-meter cube. A successful identification was rewarded with a non-location specific sound and a point accumulated. If a sound was not identified, a point was deducted and a negative sound played. After a successful number of identifications, a new level or stage was reached with increasing speed of playback and sample selection demonstrating basic adaptivity. After initial observational tests, it became apparent that different players evolved a different strategy or behaviour to increase their likelihood of locating each sound correctly. These strategies included moving from the central position, covering one ear, tilting the head to one side or moving to an alternate position after each sound was broadcast. This showed how a 'player' would evolve different strategies to improve their interaction with the system, in response to the given task.

To implement a more resolved experiment, the basic engine was rewritten to include more sophisticated data collection than simple scoring of correct matches. This allowed the previously randomized collection of samples to be categorized by difficulty of identification. It was also noticed that the sequence and location change of samples played altered the difficulty for different players, depending on their amount of exposure to the test and the strategy they adopted. One of the experiment variables was the potential for the participant to move within the environment. This is an interesting behaviour revealed by observing the initial tests, so a *correlational* study approach was added allowing one group to move around and another to stand centrally during the test to identify the effectiveness of these different behaviours in identifying sound locations. The experiment was designed using automated data collection, a specific environment (controlled test) and a short questionnaire (descriptive test) was used to identify interaction and responses. A detailed account of experiment design with sample questions and supporting data is provided in Chapter 3 *Design Strategies*.

The following data was collected to build a detailed account of each sequence of interactions using a sequential model the data was compiled into lists using software written in Max/MSP to mediate the experiment. The headings below summarise the data collected.

- *Event* (timestamp for organising lists to track user interactions)
- *Sound* (documents which sound sample was played)
- *Location* (location each sound was played at (1 of 8 available))
- *UserResponse* (sound location identified by user)
- *Response Time* (time delay measured from start of sound played to user identification of sound location).

This quantitative data was archived with observation of subject behaviour (overhead video capture of subject during test). A qualitative self-report was completed after each test. This approach can be used to inform sound design parameters for Adaptive Social Composition systems by providing examples of how non-performers respond to a range of sound objects and sound placements. It also reveals a range of methods to make interaction more rewarding through effective response, or to extend engagement by increasing the challenge through the introduction of more complex material. Strategies such as:

attack (increase speed between sound playback)
counter attack (select and move relative sound location)
subterfuge (filter sound, e.g. increase reverb),
distraction (mask sound, i.e. introduce background noise or motion)
anticipation (change context or sequence of sounds played)

The design of such an experiment is an essential part of the development process when dealing with less predictable social situations or interaction between people and adaptive software, and significantly informs system implementation; refining strategies and methods for extending engagement.

The test software, implemented in MaxMSP was designed so that it would archive the variables of each test session in a list that can be recalled by the software and sorted in different ways, allowing different variables to be compared across a single test or all tests. It allows single features such as sound location (speaker location of sound played) to be compared with all player response times to show whether one spatial location is easier to determine than another. The system can be set to play either a single sound for all tests in all locations, to play groups of sounds with similar characteristics sequentially or to randomize sounds. Each of these modes enables experiment variables to be controlled and isolated. This test software can be used for a range of descriptive studies where the test population is exposed to different experiments relating to *exploratory, organisational, relational* and *transformative* tasks. The data collection and retrieval are very flexible. Any variable can be isolated and compared, while each experiment can be supported by a correlational study to cross reference the validity of the data against the designed task. Each task is designed around the core interaction models identified in related tangible interfaces for collaboration *exploratory, organisational, sequential, relational* and *transformative.*

2.2.2 Spatial Orientation

Given the design strategy of *proximal interaction* as a compositional parameter, spatial orientation is necessary to enable a participant to contextualise their interaction with, or influence over, an adaptive social system. During the testing process discussed it was apparent that even the mapping of buttons on a control pad posed different problems to different users. What appeared to be a logical mapping of buttons to speaker locations for one subject had to be explained to another. If an interface is not either intuitive to use or logical to explain it is inefficient and distracting to use. Within the commercial games industry standard mappings for conventional actions to control a character or object and its behaviour with the game environment have evolved over the past thirty years. Despite this nearly all commercial games allow the player to edit the control pad mappings to their own preferred set up to customise control models and user preferred interaction behaviours. Alternatively, a specialist interface is supplied for the product or the game is designed for a specific dedicated controller. This is not the case with most novel interfaces for musical expression, where the conventional model of learning an instrument is more common. Using spatial relationships between individual participants and the environment in which they are interacting is a method more familiar to performing arts, installation or dance based performance technologies. This potential has also been identified within multi-user instrument systems. Novel interface design for specific performance or installation situations has the potential to motivate new interaction behaviours. As a consequence, immersion or engagement can be refined through observation of these interactions or responses. We previously discussed proximal interaction in relation to sound objects as a strategy for motivating *conscious interaction*. A dancer, whose body can be considered as her instrument with an extensive repertoire of expressive possibilities, can adapt to different sonified spatial relationships to interact with a system. Often this is established through video tracking or gesture analysis, establishing a tangible dialogue between movement and sound (Weischler et al. 2002). However, the adaptive system design described here is not intended for dancer system interaction, but for social interaction based on simple movement and musical exploration. *Learn-ability* is a required feature of systems for novices so a method for enabling this process is critical. Structured tasks that follow established interaction models have proven to be effective and many of these models build on early play experiences. The three Orb interfaces within the prototype Orb3 system are designed to reveal their potential by demonstrating how they influence the generated sound environment. Their linear motion has a direct relationship to sound object diffusion when they are operating autonomously and their motion behaviours of *lead, chase, avoid* are represented through sound placement. This global behaviour is tangible as the scale of motion is relative to the scale of diffusion. These behaviours are limited to two dimensions until an Orb is picked up. If a participant holds an Orb interface, they are given three-dimensional gestural control of sound object placement. They would have already learned the significance of proximity so the intuitive movement or orientation of the Orb interface is a natural extension of this previous behaviour. One strategy to introduce this new set of parameters is to locate the Orb's sound object relative to its physical location. As the user picks up an Orb the sound object generated by it's current location and environment parameters is focused on the user. The sound produced is more clearly defined with more literal mapping of user location to sound diffusion. The users local hand movements collected by dual axis accelerometers influence sound object diffusion. Personalisation of the current sound is achieved by monitoring basic biofeedback (skin temperature and pulse) collected by the held Orb's onboard sensors. These collected variables directly influence sound placement and apply velocity and momentum behaviours. Temporal elements are added to the current sound object in response to biofeedback variables, reinforcing the relationship between the current sound and the user holding that specific Orb interface. This strategy is used to motivate conscious interaction (movement and proximity) while harnessing subconscious interaction (skin temperature and pulse) to further influence and personalise sound synthesis. This layered interaction model is derived from the evolving rule set of Nine Men's Morris discussed earlier: stages *placing, sliding and jumping.* Spatial orientation is an established method for mapping temporal gestures to real-time synthesis and diffusion-based performance, for by example using a graphics tablet encoding small gestures from exaggerated body actions through both stage performance and collaborative installation. A number of these works and performances were presented as part of the 2006 International Computer Music Conference, 23+7 Studies in Experimental Electronic Music, Sala Metronom, in Barcelona September 2005.

2.2.3 Perceptual Constructs

An effective strategy for non-musicians collaborating within a social adaptive system is to initiate the visualisation of simple geometric structures in relation to actions within the compositional process. A shared mental model of a potential sequence of events can reinforce collaborative actions, in effect a set of relationships encoded in symbols, a score. Alvin Lucier's 'score' for the piece *I am Sitting in a room* (1969-70) (Lucier, 1995) is a list of instructions that initiates a process that is transformed through mediation of an environment. In response to first learning of this approach in 1996, a new conceptual 'score' was designed for individuals to visualize *points in space* in relation to the physical environment.

"A point in space seems perfectly objective. But how are we to define the points of our everyday world? Points can be taken either as primitive elements, as intersecting lines, as certain triples of intersecting planes, or as certain classes of nesting volumes. These definitions are equally adequate, and yet they are incompatible: what a point is will vary with each form of description."
(Varela et al. 1991)

At this stage in the research, the author was considering the problem of designing visualization tasks to generate shared mental models that have tangible reference points in the real world that can be identified and mapped to refine human computer interaction within cybrid collaborative sound environments.

Cube Score (Conceptual visualisation task)

> *Visualise a point in space, within the room you are currently in*.
> *Visualise points until you have eight points in total.*
> *Move these points to form the corners of a cube.*
> *Orientate your cube in relation to the physical environment.*
> *Describe where your cube is.*
> (Livingstone, 2001)

The score works effectively when delivered line by line as a set of tasks to a group of willing participants. It engages participants in a

range of cognitive tasks with a degree of *indeterminacy* that is framed by their physical location and concludes with verbal self-report by individuals.
Cube score deconstruction:

- *Visualise a point in space, within the room you are currently in*.

This *exploratory* interaction engages a different form of attention as the visualisation takes references physical space and the process is inherently proximal, i.e. when we imagine a flower we create a symbol or instance of flower in our imagination, however when we are asked to visualise a co-ordinate we need to relate it another element or it has no meaning. As we engage in the first step of the task, we need a framework or common reference to structure the process.

- *Visualise points until you have eight points in total.*

The second instruction provides motivation to explore, to reposition points relative to each other and reinforces the process of visualisation.

- *Move these points to form the corners of a cube.*

The third instruction manages our attention by introducing a shared mental model, a geometry we are familiar with. It also redefines our mode of interaction from *exploratory* to *organisational,* providing a more familiar *control model.* It also introduces a *sequential* process as individuals visualize each point moving into position. Of course, some individuals will find this mental exercise easier than others and will simply visualize the points moving simultaneously into position to represent a cube.

- *Orientate your cube in relation to the physical environment.*

This fourth instruction acknowledges that we have succeeded in the task and identifies the cube as belonging to the participant, it also reminds them that although the cube is their creation that they can also move it and *relate* it to the environment. This *relational* process follows an *interpretive* model. Individuals have freedom to orchestrate the relationship of each point of their perceived cube to the environment.

- *Describe where your cube is.*

The final instruction invites individuals to explain where their cube is to other participants; this process identifies the individually perceived or virtual cube sequentially. As each participant reveals the position, scale and orientation of their cube, this further reinforces the significance of a *relational* process in motivating interpretive behaviour in participants. The original *points in space* have been transformed into a contextualised symbol that can be readily identified. The relative position, scale and orientation of these symbols as described by their creators represent groups of *conscious* interactions that have contributed to a shared *transformative* process, however, the mediating environment or room architecture has also contributed to this process.

At this stage of research, the author was also concerned with implementing a system that would allow a computer to map the co-ordinates of each individual's visualized cube. This was achieved by writing a sound placement application. After initial calibration of listener position to speaker placement, the listener was invited to select 8 different sounds to represent the 8 points visualized. The user could position these sounds by controlling panning, pitch, and reverb. These Parameters were representative of position, height and depth. The software played each sound as a loop with mouse control influencing the three parameters. Thus, the user could recreate the visualised 'points in space' as individual distinct sound objects. Each process could be recorded and replayed creating a simple composition from the visualization task, an example of *embodied cognition* (Brooks, 1991; Clark, 1997). In this example, the mediating system facilitated the cognitive process and motivated emergent behaviour through physical interaction with perceived elements. This software prototype *soundspace* (Livingstone and Swain, 1997) was designed as an installation so that users could use a mouse to drag graphical points off a computer screen into physical space. Each 'point' was represented sonically as a sound object, a distinct user-selected sample. Each sample as perceived by the participant could be repositioned and mapped by the software to create a simple cube representation, pitch (x) was varied to represent the relative height of the 'point', panning (y) was used to represent horizontal position. Combined volume and reverb indicated depth (z) allowing each 'point' or sample to be placed. This early prototype allowed a perceptual construct created from the original cube score to be transformed into an integrated sound object, constructed from eight samples, the position of which was objectively mapped as 3 dimensional coordinates. The placement of stereo speakers and room dimensions were also mapped in software. Points making a cube sound object were initiated by mapping sample pitch to relative height of point perceived using the scroll wheel altering pitch (x). Horizontal placement was mapped to panning parameters via cursor left or right movements (y). To move a point or sound closer or further away the user dragged the mouse towards or away (z) from them changing relative volume and reverb mix. Although the sound diffusion was simplistic, the human computer interaction was effective with users able to identify points and associated sounds. This system provided an interpretive method of recording a *relational* process. Previously recorded single point sessions could be played back and layered allowing up to eight points or sounds to be positioned relative to one another as heard by the listener. Multiple users were not supported, although the software was useful in prototyping a mediating technology for this type of perceptual interaction. This approach has informed the design of perceptual tasks used for subsequent experiments to identify relationships between participant actions and musical behaviours that form the design strategy for adaptive social sound environments.

The original cube 'score' was presented at Consciousness Reframed 2000 Conference, Newport, Wales (Livingstone, 2001). During delivery of the paper, conference delegates were asked to engage with this process. After a verbal introduction, individuals were invited to report their cube position. The reason for requesting a voluntary self-report in this way was to reveal the range of *interpretive* or *transformative* strategies that individuals apply during interaction within a perceptual task. The varied responses are significant because they effectively illustrate how extrinsic knowledge informs the way each individual responds to the task. Each individual created a construct related to their past experience or specialism. By describing this abstract process, each individual provides a different perspective on the same process. This can be true of collaborative music where extrinsic knowledge of musical relationships shapes the behaviour of individual collaborators. The cube score was designed to motivate these different responses by encoding the extrinsic information through an accessible task. The extrinsic influences on the visualization process are revealed when participants describe the location of their cube in relation to the physical environment. By associating perceived *points* with a physical relationship or relative position that others can plainly observe a collaborative exchange of abstract responses or behaviours can be revealed. During the Consciousness Reframed 2000 presentation, the first volunteer reported the largest possible cube, visualised to be floating above the audience and constrained by the height of the auditorium. The second contributor, (an architect) described his cube as penetrating one wall of the auditorium and also occupying the adjacent room. The third contributor (a researcher with a background in Psychology and Human Computer Interaction) described his cube as a diamond rotating on one corner inside his head. These three responses to the same score demonstrated three significant factors; that individuals were responsible for ‘their’ cube, that individuals related past experience to cube conception, and that a different interaction model could be used to categorise each individual perceptual process. The first cube reported follows the instructions (the visualized cube is contained within the room, no features are added) - *control model*, the second modifies the instructions and adapts them to his own context (the visualized cube penetrates an external wall. As an architect by profession this individual's perception of the room includes a three dimensional understanding of the space. The architect subconsciously demonstrates this ‘bigger picture’ by playfully extending the score parameters - *interpretive* model. The third contributor extends the act of interpretation by adding additional parameters, a creative response that adapts the cube visualisation to a dynamic object - *transformative* model. Clearly, this original body of enquiry was speculative utilising an informal but public self, however it has been useful in establishing a design strategy for motivating behaviours that can be tracked, archived and transformed through proximal interaction. It has proved a useful way of illustrating the *transformative* process of mapping abstract data and interaction between people and systems within a common framework.

This form of mediated interaction could be described as an experiment in *distributed cognition* (Hollan et al. 2000) in the broader sense encompassing interactions between people, resources and an environment. The environment informs each participant’s reactive process. Common resources are provided so communication can take place between people and the environment through a mediating framework. This initial research proposed that a mediating framework can be implemented in software replacing the subjective self-report of individuals with embodied interaction, evidence of which can be recorded and replayed. This research exploring sonification of *perceptual constructs* combined with the strategy for an evolving rule-set to motivate new interactions and behaviours provide a framework for implementing an adaptive social composition framework. Hollan, et al. (2000 pp.190) present the following observations as a framework for distributed cognition in relation to designing new forms of human computer interaction:

- *People establish and coordinate different types of structure in their environment.*

- *It takes effort to maintain coordination.*

- *People offload cognitive effort to the environment whenever practical.*

- *There are improved dynamics of cognitive load-balancing available in social organisation.*

These core principles helped shape the interaction design for the prototype Orb3 compositional framework, reflecting on the nature of interaction with interfaces, collaborators and sound environments. This form of collaborative interaction within a shared compositional process for non-musicians can present a number of challenges to participants. Presenting appropriate challenges or problems to solve that fall into the musical interaction models *exploratory, sequential, organisational, relational* and *transformative,* can effectively motivate *distributed cognition.* The four observations can also be considered in terms of learning modes. The first two establish that people need time to *absorb* new contexts. The second two identify that people *adapt* to new contexts more effectively through collaboration. This realisation led to the design of interface modes for the Orb interfaces for context sensitive parameter mapping of sensor data. The first mode *absorb* is active when an Orb interface is operating autonomously. An Orb's sensors are calibrated to collect environment data and it's relative position. It also exhibits one of three motion behaviours; *lead*, *chase* or *avoid.* This allows participants to observe potential actions in context and allows more complex interactions to be revealed overtime. *.* The second mode, *adapt,* is active when a participant is holding an Orb interface. The sensor parameters of the held Orb are remapped to direct actions or gestures, these gestures are interpreted and classed as either *sequential*, *organisational*, *relational* or *transformative* events. As the player learns new gestures or adapts to new musical contexts layers of interaction are revealed. Collaboration reveals new relationships. Understanding the context of each problem or challenge can be useful in identifying the likely responses and thus mapping them to appropriate software-mediated events. A musical sequence or structure that is interactive offers numerous possibilities or problems for a participant to solve. Identifying these problems or potential interactions is a rewarding part of the creative process. The quality of the end result of a process is often determined by the behaviours exhibited during it (Getzels and Csikszentmihalyi, 1976), so it is a significant design strategy to incorporate current participant behaviours within an adaptive system design. Of course one of the dangers of embedding problems to be solved within a collaborative process is that they are either too simple or too challenging for participants to resolve, so it follows that attributing behaviours to identified actions is one strategy for implementing adaptivity effectively. As problems are presented contextually, depending on current identified behaviours, possibilities unfold over time in relation to each participant's behaviour.

2.2.3 Summary (section 2.2)

This section considered conscious interaction using a traditional board game as a model to introduce a group of actions (*placing, sliding, jumping)* that can be applied within a range of interaction models *(exploratory, organisational, sequential, relational and transformative)* within an evolving rule-based framework of identifiable behaviours. *Distributed cognition* was introduced in relation to these behaviours referencing related research fields. A overview of the Orb3 compositional environment design was presented. Conscious interaction was discussed using an original score 'Visualise a Cube' to introduce the strategy of process-driven collaboration using *Perceptual Constructs* to map perceived objects to diffused sound samples presenting alternative perspectives to interaction design within shared frameworks. Experiment design principles were also discussed, introducing three established forms; a control study, a correlational study and a discursive study.

In the next section, consideration is given to indirect actions and strategies for differentiating between classes of behaviour; examples are considered from the design process of the Orb3 compositional environment.

2.3 Subconscious Interaction

One of the questions that arises when observing different behaviours or identified actions in participants is how to distinguish between the intentional and the unintentional. Within a conventional set of musical interactions roles are typically assigned either by direction or mutual agreement, an individual's musical abilities are known to a greater or lesser degree. A range of musical behaviours are already in place and can be easily identified and categorised with different interaction models. For example, reviewing a new score or musical passage through playing *(exploratory),* distributing roles or musical parts amongst players *(organisational),* decoding of a score to performance *(sequential)*, adjusting one's performance to that of one's peers or the acoustics of the environment *(relational)* and modifying material and responses in relation to actions or behaviour of others *(transformative).* These behaviours are learnt through experience. Different frameworks are experienced as part of both formal and informal musical training. Musicians actively seek new behaviours as they learn to *interpret* and *transform* musical material but ultimately they mediate their own actions consciously. There are also secondary behaviours that may not be an intentional part of the process, for example, some musicians exhibit secondary behaviour that is perhaps a by product of the act of making music with their instrument: head and body movement, breathing and facial expression. These *body gestures* are more usually exploited within dance or other performing arts (Sparacino, 2000). Within a non conventional collaborative sound environment for novices it is common for a system to be designed around one interaction model with groups of interactions devised around these anticipated behaviours which the interface or software mediates. What is less common is for a system to respond to unintentional behaviours on the part of participants that can be meaningfully integrated into a collaborative framework. In fact this approach reinforces learning and indicates system potential by revealing valid actions of which a participant was previously unaware. One strategy is to motivate actions that fall into identifiable interaction models. For example walking into the tracked floor area of the Orb3 environment is classed as an *exploratory* action; it is registered by a subtle change in sound material relative to an individual location. This is further motivated by the mobile interfaces which position themselves in proximity to the individual depending on each Orb's current mode *lead, chase, or avoid.* This in turn motivates further movement on the part of participants. These movement interactions are initially unintentional in relation to the sound produced, and the sound interaction can be said to be subconscious as their proximal influence on Orb behaviours is registered with resulting feedback but not deliberately controlled by the participant until the relationship is fully perceived. In fact the visible Orb motion behaviours are a *subterfuge*, a *distraction* using the visual dominance of moving elements (a form of attention management) to focus attention and motivate *exploration* while using these proximities to transform current sound diffusion, implying control behaviour of each interface in relation to tangible sound objects.

2.3.1 Behaviours

Dividing potential participant actions into interaction models such as *exploratory, organisational, sequential*, *relational* and *transformative* allows appropriate system behaviours to be mapped, grouped, and stored in memory and enacted in response to identified actions or to motivate new actions using a range of perceptual cues. Behaviours are identified as gestures representative of a known action encoded in a list relative to current system parameters and concurrent events. In designing a shared framework that combines human actions and system events, it is useful to consider these interacting elements as patterns of behaviour. For example, the grouping of an intended known gesture (matches data to a previous known action) to a diffusion event (algorithm for spatial manipulation) is considered as a *relational interaction,* since the participant communicates intent through manipulation of the interface within known parameters. A *sequential* behaviour is designed for tracking tempo relationships in reaction to system or collaborator events, for example leaving structured spaces with reduced background sound to motivate input. An *organisational* behaviour consists of recognizing and arranging sound object placement and can be mediated by the proximity of an interface or participant in relation to another participant or interface, a parallel to the way cellular automata can be used to evolve music (Miranda, 2001). An *exploratory* behaviour is identified on entering the system environment: influencing other objects through presence. Equally, when an Orb is picked up the initial movement and gestural possibilities are revealed through exploration and system feedback. Just as a typical cellular automaton is based on a grid with a set of rules resulting in various evolutionary structures, a floor area can be divided into a grid for simple video-based tracking of objects in relation to a time signature. We have already introduced the three autonomous motion behaviours for individual Orb interfaces *lead, chase* and *avoid.* These elements have the potential to generate different spatial relationships between Orb interfaces. If we see the positions of these objects as if they were cells on a grid, events can be triggered in relation to these proximal behaviours. If a fourth object, a person enters the grid their proximity could add another neighbour, evolving a new behaviour. People will not necessarily register such a process. This interaction could be described as subconscious, and the principle can be adapted. A symbol-based approach for organizing proximal behaviours was proposed at ICMC 2004, which used a standard set of symbols, *line, triangle, square* and *circle* mapped to different control parameters. The system responded to either the conscious hand gestures drawn with a finger or by recognizing the same symbols in groups of participants via an overhead camera (Livingstone and Miranda, 2004). This strategy provides simple patterns that can be identified both by participants and by the mediating system, for example if three Orbs form points that make a straight *line* a global event can be triggered, panning all sound to a position aligned with the Orb interfaces. If a triangle is recognized the length of each side of the *triangle* is measured and mapped to envelope parameters to transform the current sound. Again this action can be assigned as a global event showing observers that although each Orb has it's own relative sound object there is a background layer of sound that is influenced by collaboration.

Using overhead video tracking to monitor relative participant and Orb locations additional patterns can be detected. A square can only be formed if a participant enters the tracked area, acting as the fourth coordinate, an unlikely event if interfaces are continually leading, chasing, avoiding but possible when Orbs are picked up for use as controllers allowing three active participants and one passive observer to self organize and make the *square* symbol. The symbol would not be possible with three autonomous Orbs and two passive observers, and is more likely to occur with three active participants holding an Orb each and two passive collaborators, motivating inclusion of the fourth observer through proximal interaction. This demonstrates that participants learn through interaction and collaborative actions to establish a viable repertoire of actions to influence the shared composition. These symbol-based relationships may be revealed over time by observing Orbs or participants but initially these global behaviours can be described as *subconscious* while participants *explore* the system. As participants transition to a more *interpretive* behaviour some of these symbol-based relationships are absorbed by participants and manifest the results as conscious interactions when participants *organize* their positions. *Local* behaviours relate to individual actions while holding and gesturing with Orb interfaces. The same symbols are used to indicate known control behaviours as gestures tracked by the onboard sensors. Moving an Orb interface at various speeds in a single direction produces a *linear* action, a panning trajectory for the current sound object. Moving in angular lines that join produces a triangular action, modifying the current object's envelope *attack, decay* and *release*. Moving in a *square* is more difficult to achieve and constitutes a higher-level task, which modifies synthesis parameters of the current sound. Moving in a *circle*, as a continuous fluid motion plays the current sound object, like running a wet fingertip around the rim of a wine glass. Subconscious behaviours are revealed by the relative location of the Orb/participant as they lead, follow or avoid other Orbs/participants but do not enact direct gestures. Each Orb is equipped with simple biofeedback measuring skin temperature when held. This parameter is used to identify when an Orb is picked up by or passed to a participant. The temperature change is mapped to equalization with a lower threshold set by default calibrated to average environment temperature. As a result of this tangible action, direct feedback is given by modifying the EQ and thereby giving definition to the held Orb's relative sound object. As skin temperature is not easily controlled this change is classed as a subconscious interaction, and the Orb adapts sound parameters to the new participant. A more refined application of Biofeedback using Galvanic Skin Response (GSR) can also be used but was not implemented in the prototype stage. Producers of the Wild Divine game system (Website reference 36) include three finger tip GSR interface and a task-based game framework using sensors for measuring breathing, pulse and changes in temperature to be calibrated against a narrative game environment. Players learn simple breathing and biofeedback control as they progress through a range of tasks such as levitating a character, moving objects and aligning elements. Although stereotypical in representation, these tasks work effectively and can be learnt by novices during gameplay. This has some significance in the field of novel interfaces where collaborative controllers lack personalization or extended repertoire of controls and expressiveness, as identified by researchers in the field developing systems based on an *instrument* model. These subconscious body states can be feasibly tracked through Galvanic Skin Response (GSR) sensing embedded in the Orb interfaces, the second interface prototype featured heart rate or pulse detection using the blood flow method (light is sent through the finger tip and variations in brightness are detected using a light dependent resistor). GSR can be used to identify involuntary shifts in attention be measuring skin conductance, two electrodes or contacts are positioned so a discrete signal can be sent between two finger tips, the signal is monitored for peaks. These peaks can be mapped to context sensitive events. This, in principle, enables emergent behaviours to be identified by comparing control parameters of Orb movements (*direct gestures*) against biofeedback from the participant integrating further classes of subconscious interaction.

2.3.2 Attention Management

Attention management is used in some well-known experiments in visual perception and attention in relation to *inattentional blindness (*Mack and Rock, 1998*)* whose research argues that there is no conscious perception of the world visually unless we are attentive. By directing attention, it is possible to hide material that is plainly visible when our attention is redirected. A classic example directs our '*focus of attention'* through a '*preattentive distraction task'.* A number of studies in the related field of visual awareness have shown how an individual's visual attention can be managed by direction through focused tasks. An example is provided from 'Gorillas in our Midst', an extensive collection of experiments of this nature are archived online by the Visual Cognition Laboratory at the University of Illinois, (Website Reference 37). In one particular example, viewers are advised that they will watch a short clip showing two groups of players passing a basketball. They are told to count the number of times the ball bounces. The indoor scene shows a group of players playfully passing a basketball. This takes place in a huddled group in front of a wall of lockers. After a short period, viewers are asked to report how many bounces they counted, and after reporting their answers, usually between 5 and 8 bounces are reported, they are asked if anyone saw the gorilla! The standard response is no, and the clip is replayed. Viewers clearly see a person in a gorilla costume walk into the scene behind the players. It waves and then walks out of the scene. It appears that we subconsciously edit the gorilla from the first viewing of the scene as we are focused on a *sequential* task. The gorilla is totally unrelated to this task and has no relationship to the context as presented; therefore, we do not see or register it, as we do not attend to it. Mack and Rock present primarily visual experiments in '*inattentional blindness'.* Of particular interest are their '*Spatial focus and "Preattentive" distraction'* tasks. Their later work on *Auditory deafness* indicates that there is an analogue between both auditory and visual perception and that '*inattentional blindness'* can be shown through parallel experiments. An MIEE (*Multisensory Integrated Expressive Environment)* designed for a range of interactions does not usually present one directed task to distract from another perceivable event. The nature of such environments is to decode numerous variables into tangible relationships and encode them into system responses. However, approaches used in the experiments revealing '*innattentional blindness'* show that our conscious interactions can be focused, that subconscious processes can be revealed. As we recognize these previously hidden subconscious processes we can modify our interaction and adapt to our mediated environment. In terms of the Orb3 system design, subconscious *exploratory* behaviours, i.e. proximity and moving toward or away from interfaces affects system parameters and generates feedback in the form of motivational cues.

2.3.3 Summary (section 2.3)

Strategies for identifying subconscious interaction of individual participants were introduced, examples of classes of behaviour were considered using proximal interaction and *indirect* gesture acquisition. The potential of biofeedback embedded in collaborative interfaces was outlined, describing how an adaptive interface can respond directly to individual input with biofeedback providing layers of personalisation, interaction and expressivity not normally found in tangible hand-held interfaces.

2.4 Chapter Summary

The core elements for a new collaborative form, Adaptive Social Composition, were established. Field specific vocabulary was applied in context to develop a new design approach to collaborative sound environments. These specialist terms, identified in chapter 1, were further unpacked and applied to describe related approaches. Compositional approaches were discussed to establish the model of environment as instrument, and to unpack potential roles of participants. Conceptual approaches to composition were discussed, showing how rules or constraints can effectively generate new musical relationships. The second research question posed in this thesis was expanded:

"Can the principles found in turn based board games be used to develop social composition frameworks that are intuitive to use in a collaborative musical context?"

Principles for structuring turn based or collaborative interactions derived from board games were established. Notions of gamesmanship and musicianship were compared and simplified to develop an accessible framework for novices. The author's Orb compositional environment framework was discussed. The importance of experiment design to identify new patterns of interaction and identify participant behaviours was introduced. The core concerns for providing a mediated framework for shared composition by novices were identified from cognitive psychology:

coordination, cognitive effort, load-balancing and social organisation Hollan, et al. (2000 pp.190). The terms *absorb* and *adapt* were introduced as interface modes to support learning.

The next chapter discusses the three core elements that are integrated to develop Adaptive Social Composition. The technologies typically used for developing novel music controllers and extended instruments are presented. Design strategies are presented that encompass interface design, environment, and principles for mediating software, to establish the need for a new approach to novel interface design. A detailed account of a controlled experiment using analysis of variance is presented as a design strategy for exploring interaction with diffused sound. The integrated approach to the author’s Adaptive Social Composition framework is discussed as a model.

Chapter 3

Design Strategies

The objective of this chapter is to establish new design strategies for Adaptive Social Composition. These are used to articulate a compositional approach that promotes co-creation within a collaborative mediated environment, using the *features, qualities, models and behaviours* established previously in this thesis. The objective of this chapter is to show how such a system can be implemented and why a new approach to hand held interfaces that have adaptive features is desirable.

The third research question is addressed:

'How would one describe such a system, what qualities and characteristics would it manifest and how would one design it?'

The third step of this thesis: *'to extend the core features of Adaptive Social Composition using a controlled experiment to identify new participant behaviours'* is evidenced by articulating the results and design process of a pilot study (section 3.1).

Firstly, interface technologies typical of the field for prototyping are introduced. A *pilot study*, based on the research models identified in the previous chapter, is presented. Section 3.2 *Interface Design* considers data acquisition, ergonomics, interaction models and references the handheld Orb3 interface. Section 3.3 *Collaborative Environment* presents overall installation and interaction context, including requirements for spatialisation and indirect tracking. Section 3.4 *Mediating Software* outlines a modular approach to software design and identifies key features. Examples of techniques and software externals that resolve many of the problems associated with such an approach are provided.

One of the primary decisions for creating new interfaces for systems that are dependent on live data input is the choice of sensor interface. The most established protocol for Computer Music software and hardware integration is to use MIDI. This well-established format allows multiple messages to be exchanged between external interfaces such as keyboards or mixers with software instruments or sequencers. These messages encode musical information into numbers but do not contain actual sound material so many MIDI based systems also combine audio analysis such as pitch and tempo detection, while more advanced systems require more complex audio analysis. Other protocols such as 'Open Sound Control' (Wright et al., 2003) enable flexible and robust methods of exchanging musical information over networks and between different software environments. Whatever the protocol, the initial task is to convert real world data from a sensor or actuator to a variable that can be mapped to a control parameter or behaviour within the software environment. The most obvious example of this form of data transformation is from a turn sensor or rotary pot. This simple resistor appears on most commercial mixer or control surfaces for modifying a range of parameters, typically volume, panning or effect levels. The resistor simply varies a voltage, typically between 0 and 5 volts and this voltage is converted to a variable. Within the MIDI protocol this would be between 0 and 127 with additional information such as controller number and MIDI channel being added to format the data so it can be mapped to the relevant software parameter, a direct physical action to a virtual performance variable. In addition to encoding note data (pitch, duration and velocity) a typical MIDI control message combines three elements: value, controller number, MIDI channel and can be mapped to any parameter the designer or composer can conceive of. This offers significant flexibility as any sensor, resistor or actuator that operates at equivalent response times and with the same current requirements can be used (Bongers 2000). Almost any interface that converts a voltage to a variable can be used. Many innovators in the field have modified existing commercial controllers to develop their own customised interfaces by removing the standard rotary pots and either hard wiring alternate sensors or adding some form of connector such as a screw terminal or mini-jack socket so that different sensors can be tested and used as required. Originally most of these external interfaces would need an additional MIDI interface to connect the device to a computer. Devices such as the *Phatboy* from Gmedia (Website reference 38) or the *Spin Doctor* from Kenton electronics (Website reference 39) feature 16 rotary pots and a switch for channel selection, with a single MIDI input and output. The MIDI input allows settings or profiles to be stored and modified in software and sent to the external hardware via a serial to MIDI interface. Doepfer (Website reference 40) also produced a wide range of commercially available voltage to MIDI controllers ranging from the Doepfer *pocketfader* to extensive multi-channel rack units. For one-off or custom controllers using alternate sensors many of these devices have been used. More recently USB (Universal Serial Bus) equipped units have been widely available. These units tend to be lighter than their metal-cased predecessors, sharing the USB port for both power and data thereby minimising external wiring and additional power supplies. These units such as the *UC16*, *UC32*, *UC33* (Website reference 41) from Evolution can be easily converted to a sensor interface and offer a low-cost robust portable solution for implementing new interfaces for musical expression. Other recent innovations in the commercial sector include the M-Audio *Trigger Finger* (Website reference 42), which features touch and pressure sensitive pads for finger tip control, ideally suited to percussion. These interfaces are all significant in their own right but many Computer Musicians and sound artists were limited by the standard control models that these interfaces perpetuate so independent developers designed more flexible units designed for those who were less interested in hacking or 'modding' existing controllers or who had specific requirements that the standard units simply did not support. The most common feature is control voltage or CV allowing both voltage to MIDI and MIDI to voltage conversion so that further external devices can be controlled, providing tangible feedback or the ability to trigger relays for mechanical control. These units include the *I-CubeX* from Infusion Systems (Website reference 43). The original unit featured 32 inputs, which could also be configured as 24 analogue inputs with 8 digital outputs, Infusion Systems also, provide a wide range of sensors for gesture tracking and performance control. This unit is mounted in a compact plastic case and requires a MIDI interface. Eric Singer developed a cheaper and more flexible alternative, the *Miditron* (Website reference 44) The Miditron features 20 connections that can be configured for either analogue or digital input and output, no sensors are provided but the manual includes useful circuit diagrams and lists of suppliers. The unit is mounted on a board and requires a MIDI interface. Both the I-CubeX and the Miditron provide a standard 5-volt circuit and can be powered from a 9-volt battery making them portable and versatile. Both have programmable ROM for storing performance/interface settings assignment and input scaling. The sensor connections, using either pins or screw terminals, are ideal for prototyping but less robust for permanent use. The *Arduino* (Website reference 45) system offers a more compact very cost-effective unit suited to building custom instruments or interfaces. The unit is small and connects directly to USB, so no additional MIDI interface is required. Of course, it is possible to integrate any of these units into a custom case with robust connections for bespoke instruments and/or live performance applications, and many effective novel interfaces or digital instruments have been developed in this way. Other high-end rack mounted systems such as Kronde (Website reference 46) with sophisticated sensing provision for full body tracking are also available, although is intended primarily for dance applications. The main limitation with all these systems with the exception of the Kronde, and the Bluetooth enabled Arduino, is that they are wired, with the obvious limitations and vulnerability that is inherent to wired systems. Again small developer Kenton electronics provides a solution via a wireless MIDI interface. Unfortunately, this unit also requires a MIDI to serial or USB interface adapter for connection to computer, but it is ideally suited to live applications for connecting conventional MIDI based instruments and hardware. Numerous performers have made use of computer track pads and pressure sensitive drawing pads for gestural control, writing custom patches to map pen coordinates and movement to parameter controls. Several examples of this approach were presented at the 2005 International Computer Music Conference, Barcelona (Experimental Electronic Music series, organized by Serji Jorda). Freedom of movement and robust data transfer are essential for handheld digital instruments and performance interfaces; developers within the NIME (New Interfaces for Musical Expression) community have begun to provide wireless solutions for voltage to MIDI or data conversion such as *WISEAR* (Swensen and Topper, 2005). An equally valid approach has been developed from the notion of circuit bending (Website reference 47), reengineering existing hardware to create a new interface. Computer Musicians and sound artists have made use of interfaces from the games industry to facilitate intuitive interaction by new non-performer audiences, others have established musical collectives using these interfaces instead of conventional instruments (Website reference 48). Standard joysticks and game controllers that combine analogue joysticks with digital actuators are particularly suited to modification and parameter mapping can easily be achieved to convert the control data to standard MIDI messages. Common approaches are to use *JunXion*, a software MIDI mapping utility developed by Steim (Website reference 49), for this purpose, or to use the human interface or 'hi' external provided with MAX/MSP computer music software to extract the control data from any interface connected to a computers USB port. After earlier works and experiments using basic stamp, I-CubeX, Miditron, Arduino and extensive testing of various modified music controllers the advantages and disadvantages of each unit were established. A low cost robust wireless interface was identified for developing the Orb3 interface prototypes. Wireless console and computer game hand held controllers (joysticks or gamepads) can be used effectively, stripping the units down to the primary printed circuit board (PCB) and microprocessor. These units are easily 'modded' adding specialist sensors and using software to interpret and scale the captured variables. Initially, wired versions of these controllers were tested, mapping the data to Computer Music software and building numerous patches in Max/MSP. These commercial game controllers are designed for sustained use in a wide range of situations where they take considerable abuse, tapping, twisting, dropping, etc. They have a two to three year life expectancy, are fairly robust and most importantly have fast response times. A standard Playstation 2 gamepad can be modified to provide 8 analogue inputs and 14 digital inputs with two digital outputs for force feedback for a fraction of the cost of the other voltage to MIDI interfaces.

<u>Summary of voltage to MIDI interfaces and alternative solutions</u>

Phatboy from Gmedia (Website reference 38) or the *Spin Doctor* from Kenton electronics (Website reference 39) *Pocket fader* from Doepfer (Website reference 40)

These units require additional MIDI interface, two MIDI cables and power supply, Intended for live events use and feature robust cases, can be modified to use a range of sensors and are useful for prototyping. They are not suitable for implementing novel hand-held controllers or extended instruments due to additional cabling and power requirements. These units range from £50 to £100+.

UC16, UC32, UC33 (Website reference 41) from Evolution M-Audio *Trigger Finger* (Website reference 42)

These units are versatile, intended for desktop and amateur DJ use. Units feature lightweight plastic cases and can be easily modified to use a range of sensors. Advantage is that communication and power is over USB so a single cable to a laptop or desktop computer is required. Can be used for novel controllers or extended instruments where size is not a limitation and they will be used near a computer. These units range from £50 to £100+.

I-CubeX from Infusion Systems (Website reference 43)

A long established unit; intended for experimental computer music and arts installation applications. Advantage of up to 32 analogue inputs, limited to 8 digital outputs. Sensor inputs are based on pin strip connectors, and are highly vulnerable; cannot be moved during use. MIDI data configuration can be customized and stored in memory so can be used independently with other MIDI equipment. Unit requires additional MIDI interface, two MIDI cables and power supply. Sensors provided are very expensive. New wireless version

overpriced when compared to newer competing systems. Standard system available for £300+, reduced input/output wireless system also available for £300+. Includes editor software for MaxMSP.
Miditron (Website reference 44)

Solves some of the flaws identified in the *I-CubeX* system. Provides 20 Configurable inputs or outputs (analogue and digital). Simple screw connectors offer stronger sensor connections. Shared earth screw cramped when all inputs are in use. No case. MIDI data configuration can be customized and stored in memory so can be used independently with other MIDI equipment. Includes circuit diagrams for DIY sensor making. Unit requires additional MIDI interface, two MIDI cables and power supply. Wired version available at £75+ New version wireless version available at £200+. Includes editor software for MaxMSP.

Arduino (Website reference 45)

An open-source platform based on Processing/Wiring language. A range of printed circuit boards for use with sensors or simple control applications are available. Excellent online community, technical details and example projects. No case provided, small screw terminals for connecting sensors. Offers a wide range of interface options including; USB, serial and Bluetooth for wireless applications. Small compact, robust design. Licensed under creative-commons. Ideal for developing individual custom interfaces, or extended instruments. Includes editor software and examples for numerous software packages including MaxMSP, PureData and Flash. Standard serial ‘Diecimila’ boards are available from £20+ Bluetooth boards start at £66.00+

Commercial game controller; *Madcatz Micron* (Web reference 50)

Standard game controllers offer an easy to customise, a robust solution for wired and wireless applications with smallest footprint of all systems listed. Up to 8 analogue inputs and 16 digital inputs can be adapted. Outputs are restricted to 2 digital. Connections are made directly to the commercial PCB so a simple method for sensor connections needs to be implemented. Data can be collected using the *hi* (human interface) object in MaxMSP using a standard USB port. Wireless units offer the best value and include a receiver that can be plugged directly into a USB port. Up to four wireless units can be used together for prototyping collaborative hand-held controllers. Ideal for low-cost prototyping and user testing. Some skill required to deconstruct packaged products and identify connections. Connections can be awkward to solder. Wired controllers are available from £5.00+. Wireless controllers are available from £15+

Each of these interface prototyping methods has been used extensively in both my masters and undergraduate teaching over the past ten years to teach an experimental approach to interaction with sound, developing novel controllers, installations and customising existing instruments. This activity was delivered through technical workshops and theoretical lectures exploring sound interaction and novel controller design. This period of exploration with groups of novices allowed me to identify which technical approaches were most relevant in a given situation. Through further testing of third-party wireless game controllers the 2.4Ghz, Madcatz *wireless micron* game controller (Web reference 50) was identified as a suitable interface for fast prototyping, providing the units are dismantled with care, and existing switches and actuators removed without damaging the primary PCB. These wireless units operate at 2.4Ghz and are self assigning, so as each unit is switched on it is automatically assigned to the next available radio frequency channel, with up to four channels or four independent controllers supported. In addition, if a unit fails the other units are unaffected and a replacement controller will automatically be identified and replace the failed unit when it is powered on, this is a significant benefit in live applications or installation environments where durability and easy low cost replacement are significant factors. The layout of the internal components varies between brands with certain models being far easier to dismantle and modify than others. The basic physical design is consistent; a wing shape, held in the palm of each hand with a joystick and four buttons for each thumb, with two buttons controlled by the first two fingers of each hand on the front of the unit. The standard unit does not respond to gesture in any way and the sequence of control is learnt by repetition and reinforced by the mapping of buttons or joysticks to actions repeated in hundreds of commercial titles. A stripped down game controller proved effective for developing radio frequency interfaces small enough to make handheld controllers with a wide range of additional sensors such as accelerometers, digital compass, light dependent resistors and tilt switches. After subsequent testing of interaction models and identifying appropriate sensors for optimum performance a custom

PCB was commissioned for the first full implementation of the Orb interface design (see chapter 4: Proof of Concept). The following design strategy summarises the findings of this section.

<u>Design Strategy 1: Interface platforms for prototyping novel interfaces:</u>

Identify which sensors will be used, how many inputs are required and whether outputs will be needed to relay feedback to the interface itself. Commercial interfaces such as *I-CubeX* and *Miditron* are effective for prototyping, and offer software compatibility. However, they are relatively expensive and connections are vulnerable if units are moved. If input only is required (to collect sensor data), a modified mainstream commercial USB music controller such as *UC16* may provide a more compact solution for sensor integration and interaction design. For hand-held interface development, *arduino* boards offer a compact solution, highly versatile and robust, suited to wired applications but less economical for wireless applications where multiple controllers are required. Conventional game controllers are a low cost low risk alternative and can support multiple wireless units at minimal cost, depending on sensor choice. Units require dismantling and customisation. Consider user actions and potential gestures *indirect, direct, haptic, non haptic*. Identify appropriate sensors for data collection *accelerometer, ultra sound, digital compass, tilt switches, force sensing resistors, light dependent resistors* etc. Choose appropriate platform for testing, then implement standalone interface/novel controller once technologies and interactions are resolved. *Concept, Design, Implementation.*

3.1 Pilot Study

To establish a methodology for implementing perception and interaction tests and to inform the novel interface design and recognition of musical behaviours a Pilot Study was designed. The Study was run under controlled conditions with 15 undergraduate students at the University of Plymouth. The eight channel cube-format speaker array outlined earlier (two tiers of four speakers, one at floor level and one at ceiling level equally spaced forming a cube, a standard for B format ambisonic diffusion) was used to test location perception of different sound material using a 'within subjects' test. Eight different sounds were created with different characteristics as illustrated in the images from 'Sonic Visualiser' below (Website reference 56). Different combinations of sound properties were combined; fast attack times, slow attack times, soft mid range pad sounds, high frequency clicks, low frequency textures, reverb and long decay times. Longer samples with a full rage of the previous properties were also used. Each sound was played four times in each of the eight locations in a randomized order. Each subject was presented with the same order of sound locations and the same order of samples comprising two hundred and fifty six events per subject. The focused task was for the subject to identify the sound source or location of each sound played by pressing a corresponding button on a game pad. The button mapping was logical and relative to the speaker positions. Subjects were given 1 minute to familiarize themselves with the call response format.

The format of the test was for the software to compile a list of events and responses for subsequent analysis, this list comprised;

- *event* (sequence in time)
- *sound* (sample played; between one and eight provided samples)
- *location* (location of sound playback; between one and eight speaker locations)
- *userLocation* (identified by user; between one and eight)
- *responsetime* (user response time from start of sample playback in milliseconds)

A standard symmetrical handheld gamepad was used as the control buttons are identical for both hands. This reduces the potential for errors due to a left or right hand bias in motor skills. Location response times can be measured to show if there is any dominance in user response time using either left or right hand thumb or index fingers etc.

If the subject correctly identifies a sound location, the sound is repeated in the same location (*learning reinforcement*) however if the sound location is not correctly identified no audible response is presented. After a second the next sound sample is played. Each user response is logged, both the identified location and the response time in milliseconds.

The experiment is designed around two *independent variables*: the <u>sound</u> sample played (8 levels/sounds) and the <u>location</u> it is played at (8 levels/locations). Two *dependent* variables were recorded: the location identification accuracy of sounds played (one of eight) and the user response time in milliseconds. The experiment had a full within-subject design, that is each participant responded to all 8x8

experimental conditions. The results for each user, a list of 256 events (4 repetitions of 8x8 experimental conditions) are processed producing a level of identification accuracy for each location and each sound sample. The response times for each location and each sound are averaged over the 4 repetitions. Fifteen participants did the experiment, and each session took approximately 25 minutes.

The controlled test was successful, showing that the experiment format and the method of testing were robust and that statistically significant data was evident from the subsequent data analysis using established methods. This means a suitable methodology, software design and experiment framework was successfully implemented to support perception and interaction tests. This is presented as the a significant contribution and design strategy for Adaptive Social Composition, when tasks are designed around the presented interaction models with explicit primary (conscious) and undisclosed secondary (subconscious) behaviours.

The results of this controlled test and findings based on the analysis of variance are explained in the remainder of this section. Radar style charts are used to show identification accuracy and user response times. Each sound is represented by a colour-coded trapezoid. Additional charts and data are provided in appendix 2.

RESULTS

An Analysis of Variance (ANOVA) for repeated measures was carried out on the data. The table of the ANOVA is shown here for the location accuracy (figure 8) and subject response times (figure 9).

Source	**Type III sum of squares**	**df**	**Mean Square**	**F**	**Sig.**
Sound	27.601	7	3.943	50.299	.0001
Error (sound)	6.585	84	0.78		
Location	1.705	7	0.244	3.55	.002
Error (location)	5.763	84	0.069		
Sound*Location	5.027	49	0.0101	2.862	.0001
Error (Sound*Location)	21.076	0.036			

(fig. 8 Location Accuracy)

Source	**Type III sum of squares**	**df**	**Mean Square**	**F**	**Sig.**
Sound	293429408.4	7	.41918456.92	9.073	.0001
Error (sound)	388092323.7	84	4620146.711		
Location	17232724.04	7	2461871.869	7.552	.0001
Error (location)	27382477.03	84	907249.56		
Sound*Location	44455226.44	49	907249.56	3.967	.0001
Error (Sound*Location)	134476785.7	588	228702.017		

(fig. 9 Response Times)

The independent variable sound was *statistically significant* for both the location accuracy and the response type dependent variables, respectively with $F(7,84)=50.229,p$ less than >0001) and with $F(7/84)=9.073\ p>.0001$ indicating that varying the type of sound played affected the accuracy of location identification and the speed of response. The independent variable location was also significant for both the location accuracy and the response type dependent variables, respectively with $F(7,84)=3.55p>.002$ and with $F(7,84)=7.552,p>.0001$. This indicates that varying the source location affected the accuracy of location identification and the speed of response. Finally, also the interaction between the independent variables sound and location was statistically significant for both the location accuracy and the response type dependent variables, respectively with $F(49,588)=2.862p>.0001$ and with $F(49/3.967,p>.0001$. This demonstrates that there is a complex pattern of interaction between sound type and its location source.

The speaker configuration used for the test provides an eight channel sound-field. The labelling and position of speakers (fig. 11) corresponds with a standard ambisonics (Website reference 15) installation for B Format. Sound diffusion is controlled by Max/MSP using a suite of diffusion externals developed at the Institute for Computer Music and Sound Technology (ICST) of the Zurich School of Music, Drama and Dance (Website reference 16). A portable M-audio Firewire multi-channel sound interface (Website reference 17) is used to connect the speaker system to a laptop computer. Simple linear single channel assignments were used in the controlled test. The test software is scaleable so additional tests can be developed for sound object motion trajectory perception interaction design using B format encoding/decoding in real-time for small groups of participants.

A Waveform and frequency spectogram for each sound is included to illustrate differing properties of each sample used (figure 10).

Test Sound Summaries:

Sound 1
2.9 seconds. Combination of hiss, noise and subtle percussive pattern with single high pitch beep.

Sound 2
1.9 seconds. Combination of hiss, sharp attacks and defined percussive pattern.

Sound 3
3.1 seconds. Mid range fast attack slow decay chord using 'Rhodes' like synthesiser.

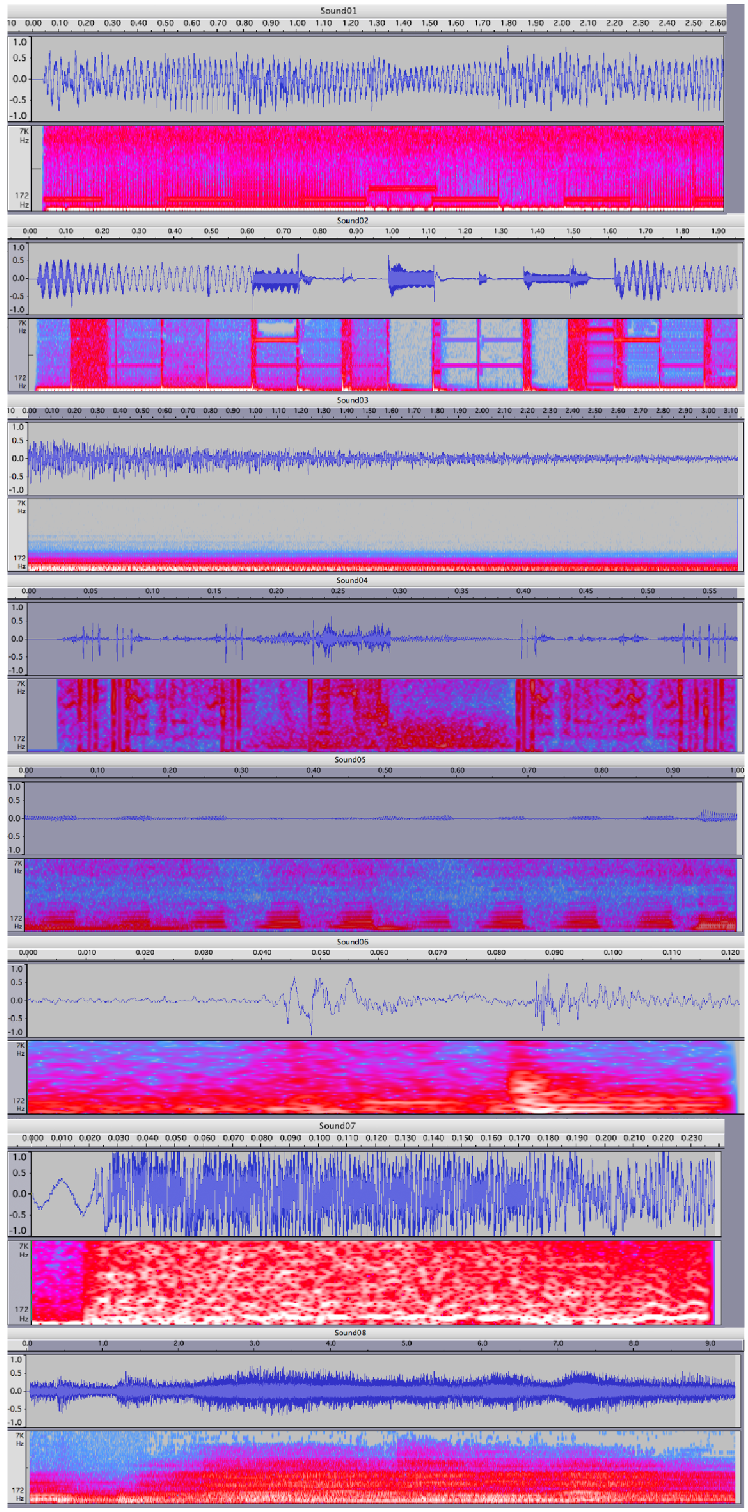
Sound 4
0.51 seconds. Clicking pattern using edited vocal sample with breath.

Sound 5
1 second. layer of abstract treated voice over repeating single note horn.

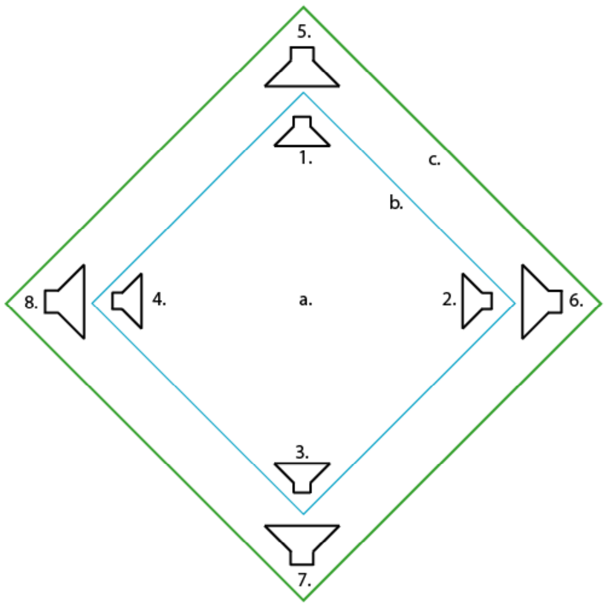
Sound 6
0.1 second. Hand percussion combined rolling tabla hits.

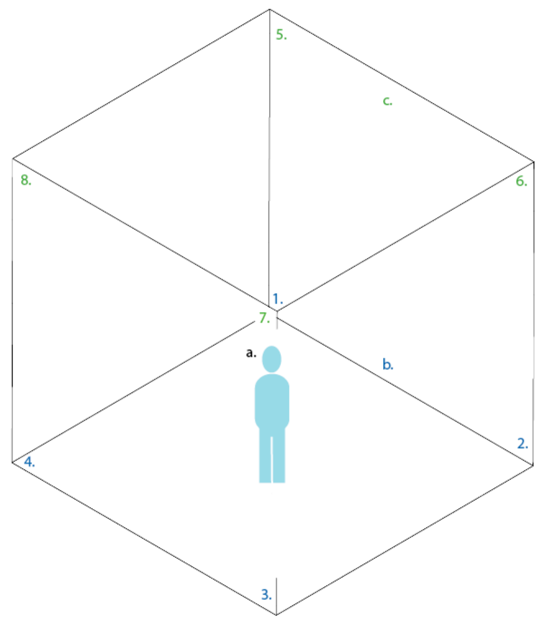
Sound 7
0.25 seconds. Harsh fast attack treated noise.

Sound 8
9.2 seconds. Piano motif, slow paced melody disguised in extended reverb with phasing. Slow release.



(fig. 10 Waveform of sound samples 1 to 8)





(fig. 11 speaker layout, plan & isometric views)

a. Central subject location (subject facing wall b/c)

b. Lower speaker tier (floor)

c. Upper speaker tier (ceiling)

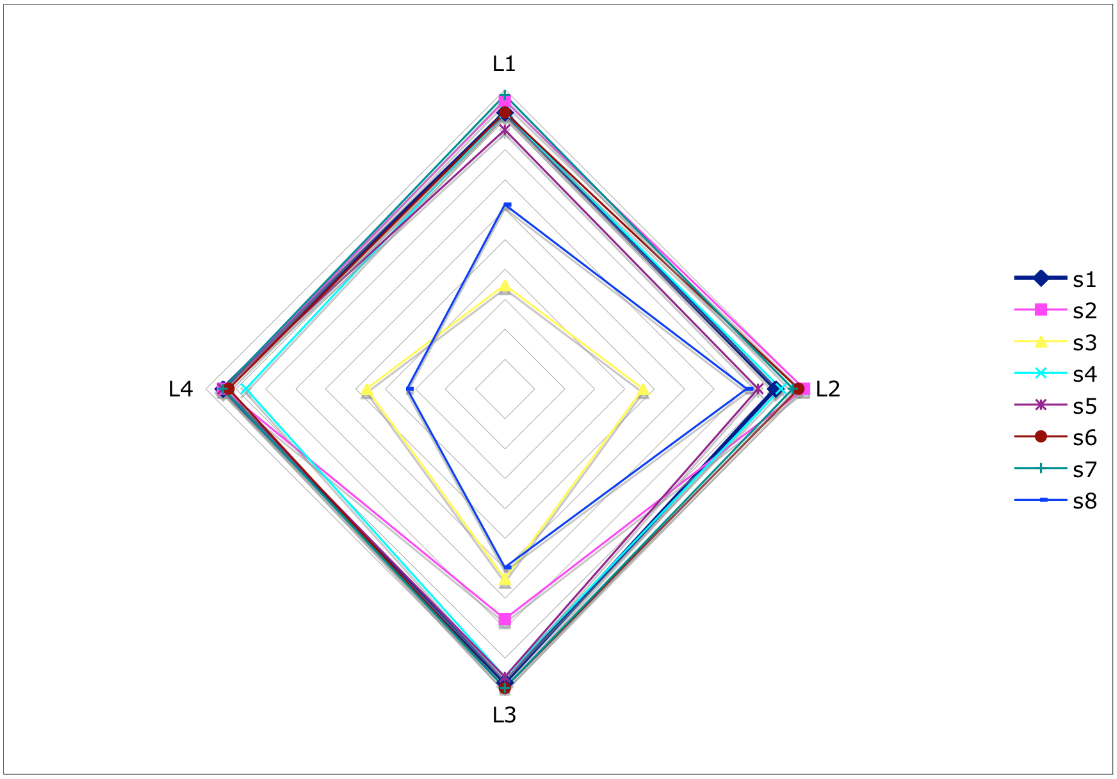
1. Front left speaker (floor)
2. Front right speaker (floor)
3. Rear right speaker (floor)
4. Rear left speaker (floor)
5. Front left speaker (ceiling)
6. Front right speaker (ceiling)
7. Rear right speaker (ceiling)
8. Rear left speaker (ceiling)

Speakers are positioned within a 3.5 square metre cube. Each speaker faces central listening position (a) at head height.

The following charts show the frequency of correct identification of sound placement in each location as identified in figure 11. Obviously environmental factors play a part in this type of test; each sound was tested in each location four times comprising a controlled test of 256 sequential events. The test was completed with 15 subjects. Two environmental factors were introduced to simulate subject relative locations, for example to understand when an individual is interacting with such an environment and has positioned themselves off centre or closer/further away from a specific speaker location, speakers 2: Front right speaker (floor) and 3. Rear right speaker (floor) were positioned 0.5m closer to the subjects central seated position.

This pilot study also indicates that each event is not isolated and that the preceding sound event can influence the perceived location of the subsequent sound. Certain sounds can be masked by relative placement, and certain characteristics can reinforce or distract from a particular relative location.

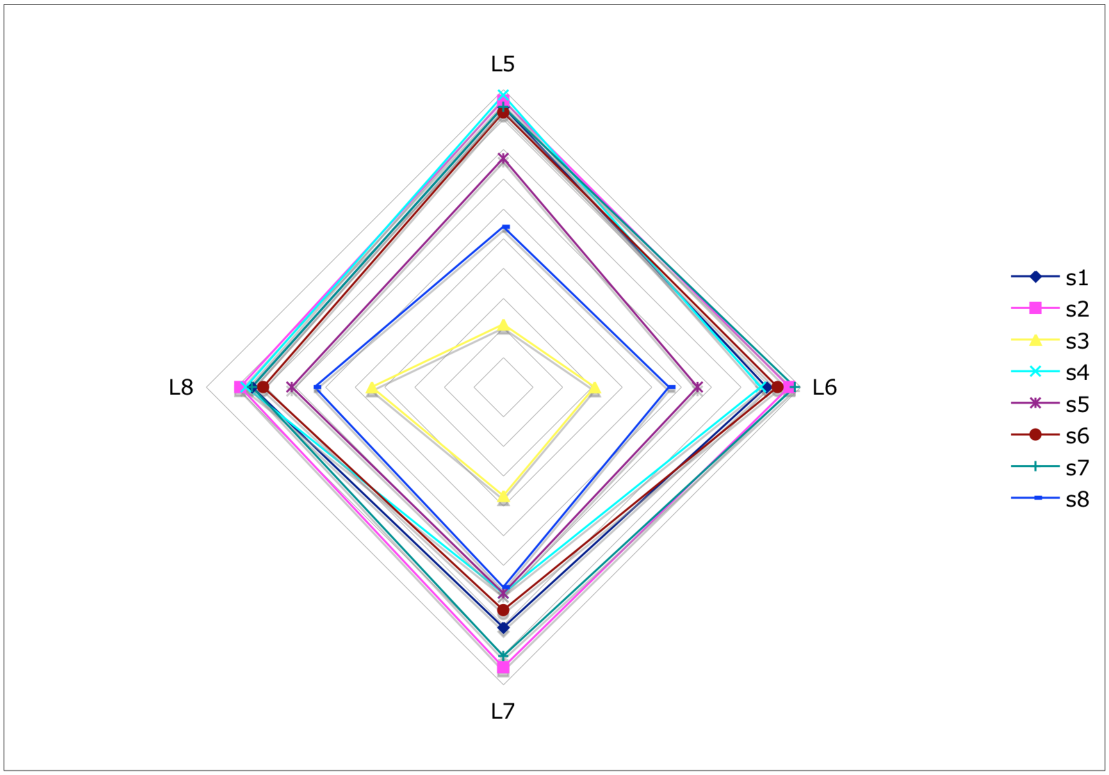
The first set of data shows the interaction between two *independent variables*, the sound location and the sound identity labelled lx (speaker position sound played at) and sx (which of the eight sound objects was played). The following charts (figures 12,13 and figures 14,15) describe in detail which sounds and locations produce accurate location identifications and faster response times. A top down view of the sound environment is shown (fig. 11) Speaker positions L1 to L8 form a cube as shown in fig. 11 (Isometric view). An environment diagram (figure 6) was presented earlier, the tests in this section were undertaken in controlled conditions in a dedicated lab providing an eight channel speaker system as illustrated.



(fig. 12 Sound Location Accuracy floor level speakers 1 to 4)

Sound three (yellow) was the hardest to locate on front left speaker (L1) Sound Eight was hard on speaker four, rear left behind subject. Subjects were placed centrally facing top right corner in representative Radar graphs.

Figures 12 and 13 show that sounds 3 (yellow) and 8 (light blue) were harder to locate than other sounds, and that due to the characteristics of these two sounds they were harder to identify in certain locations, for example subjects rarely identified sound 3 (yellow) in locations 1(L1) (front left floor) or 5 (L5) (front left s



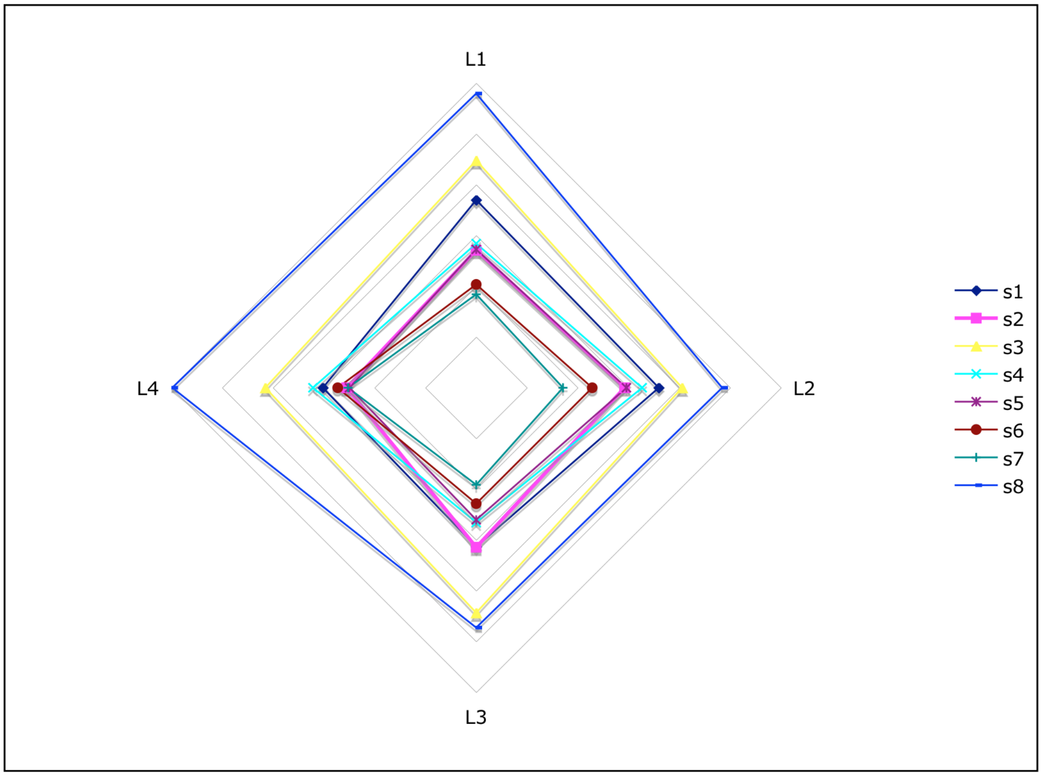
(fig. 13 Sound Location Accuracy ceiling level speakers 5 to 8)

Sound three was also hardest to locate on front left ceiling speaker (L5), followed by sounds eight and five, (L7) was hardest to identify overall, positioned rear right behind and above subject.

The data shows that sounds played on the floor level tier of speakers (speaker positions 1 to 4) were easier to locate overall with the rear left location proving harder to isolate than other floor level positions for sounds 3 (yellow) and 8 (light blue). It is clear that the top tier (speaker positions 5 to 7) were harder to isolate for most subjects, with location 7 (rear right ceiling speaker) proving the most difficult.

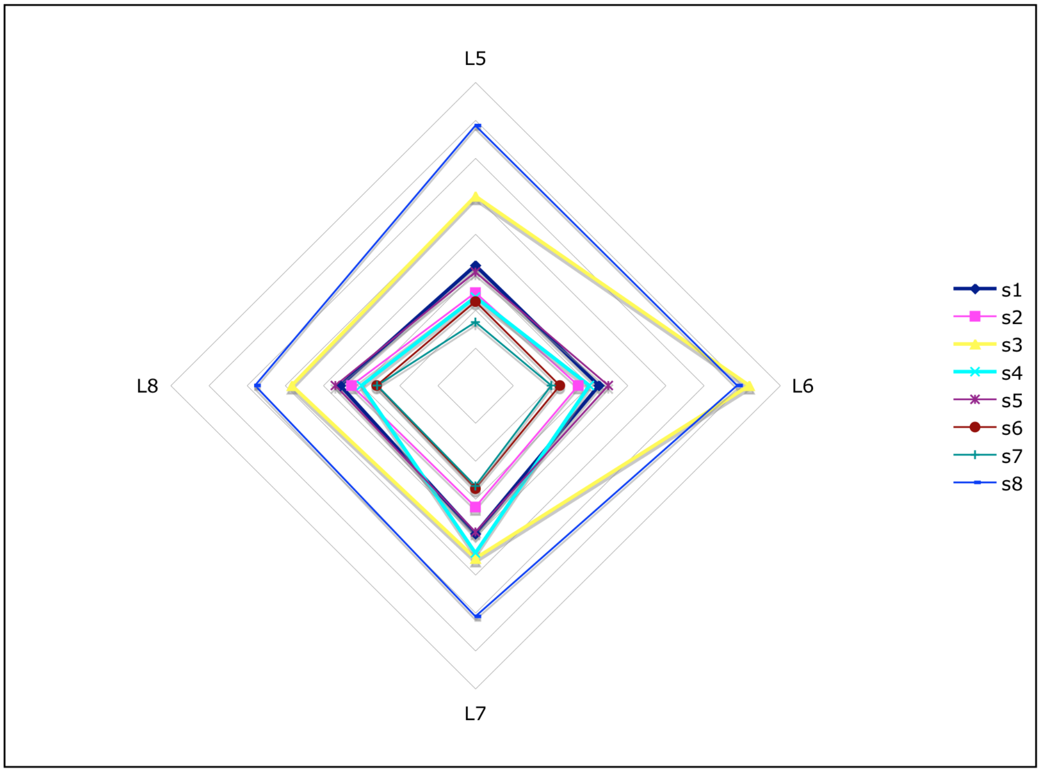
One factor observed during the test is that when a subject perceives a sound is behind them they look over their right shoulder, this subconscious behaviour appears to misdirect a subject's perception as the location they turn to face is location 7, yet this is then discounted as the likely source in the majority of subject responses. This is significant in terms of sound design for adaptive social composition as it shows sounds can be created that are easier or harder for individuals to perceive in different relative positions. By managing the balance of characteristics in a sound object dynamically and relative to user orientation it is possible to vary this perceived position-identification difficulty. As discussed in chapter 2, 'Adaptive Social Composition', strategies can be used to extend the *explore-ability* of a sound environment by allowing both participants and the mediating software to position sound material relative to participants to motivate interaction through attention management. Motivation strategies can be used to facilitate learning through call and response and can include an element of gamesmanship. By identifying the sound characteristics that allow a sound object to be deliberately 'placed' or 'hidden' within an evolving composition relative to other participant's orientation within a diffused sound environment, new compositional behaviours can be evolved. Equally, tempo characteristics in sound objects can influence participant behaviours, monitoring response times and adapting temporal structure of composition to subconsciously influence participant actions. If, for example, the mediating framework of such a system identifies a *relational* or *conversational* interaction model the interface sensor parameters can be adapted to control specific sound characteristics of real-time synthesis giving the individual participant an appropriate level of feedback that can be described as *nourishing* by rewarding exploratory behaviour that is significant within a *dynamic context.* If a lower level behaviour such as *sequencing* is identified, a logical linear response is anticipated to reward this form of timed control behaviour, interface sensor parameters are remapped to direct object control for sound placement.

The second set of data (figures 14 and 15) show the two *dependent variables.* These are the user identified location and the user response-time in milliseconds, the response-time is initiated when each sample is triggered, therefore longer sounds are expected to have longer response times as they may include the time the sample is played. This is because in an initial trial it was observed that subjects did not wait for sounds to complete before they had heard enough to confidently locate them. In light of this, some sounds were included with arbitrary initial material to change the listening mode, so response time is measured from the start time of the sample. Another environmental factor in such a test is the interface used for collecting the data. A standard handheld playstation controller with identical left and right hand controls was used for each test. For the demographic of subjects taking the test this type of interface is familiar so an extended training or familiarisation process is not required. Two minutes were allowed for users to familiarise themselves with the button-to-location mapping. This took a logical form that proved effective as is supported by the overall high accuracy rate. This shows that the interface was effective in the assigned role and did not unduly bias the test.



(fig. 14 Response Times Floor speakers 1 to 4)

Subject is positioned centrally facing side L1 to L2. Note that sounds three (yellow) and eight (blue) have the slowest response times. Sound eight has slowest response times on L1 and L4. It is interesting to note that the response times on the left of the subject (L1 and L4) are significantly slower than the right hand locations for the same sound. This is not the same with other sounds and therefore there is an implication that there are properties specific to this sound that hinder perception on the left of all subjects.

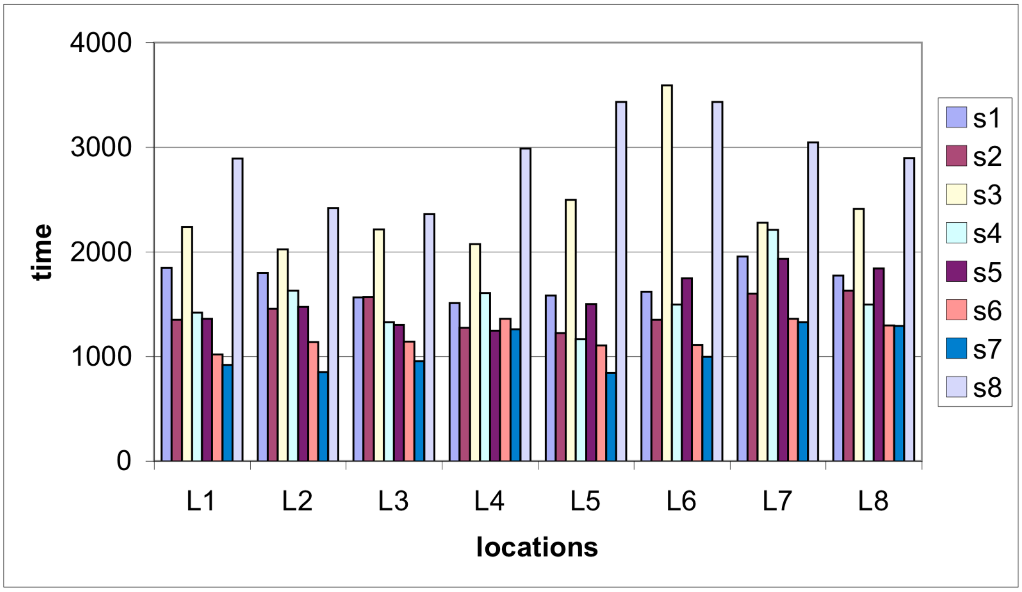


(fig. 15 Response Times ceiling speakers 5 to 8)

Note that sound three (the hardest sound to identify) has the slowest response time at L6 (front right ceiling speaker). Sound eight had similar slow response time on the front speakers L5 and L6 while other sounds have relatively balanced response times. Again there is an indication that sounds above subjects are processed differently as response times are not consistent across floor and ceiling speaker locations; Sound 8 had the overall slowest response times but the with the slowest response at rear left floor location but the fastest response directly above at rear left ceiling.

Figures 14 and 15 show response time for each location for correctly identified sounds. For the floor level speakers (1 to 4) the data shows that position 3 (rear right floor speaker) is dominant with the quickest average response time regardless based on the average of all sounds played. This speaker was placed closer to the subject in the test and acts as a control variable or reference. Sounds played on this speaker appear to be closer to the subject, and as a consequence of closer physical proximity, the perceived amplitude is higher and therefore easier to locate. A further factor is that a right thumb press identifies this location. Speaker position 2 was moved slightly forward and this is reflected in the overall average response times. However with location 4 (rear left floor speaker) correctly positioned the overall response time is slower than all floor locations showing that sounds played behind the left of the subject take longer to locate.

The full data for the pilot study including individual subject data and primary within subjects' data analysis is presented in appendix 3. The radar style charts presented in this section were generated directly from this data to aid the visualisation of sound source and timing events in relation to each location. The test software, data collection and analysis resolved for this pilot study form a methodology for further controlled experiments studying sound interaction, behaviour and perception in order to inform the design of Adaptive Social Composition Systems.



(fig. 16 Overall average response of each sound at each location)

Figure 16 shows the overall response times (milliseconds) for each sound object. Each interface button is mapped to a specific location (identical numbering for relative speaker position/sound locations as shown in fig.11 plan and isometric views) and each sound object is played in each location four times for each test. The average of all tests is presented; clearly showing that sound 7 was identified the quickest and was therefore the most distinctive sound. This sound has a fast attack and higher amplitude than all other samples. It is also the shortest sound motivating a quick response. Sound 8 has the slowest response time as there is no distinct difference from the other samples in the slow opening sequence, the sound has a mix of high and low frequencies with temporal elements making it hard to localise. Sound 3 was perceived as the second hardest sound to localise. This sound although simpler than sound 8 has a lower amplitude with slow attack and long release times, again this tended to be deceptive. The primary aim of the test was to identify the difficulty of perceiving different types of sound material in different locations relative to the subject. The secondary objective was to show whether the characteristics of sounds played influence user behaviour. This secondary feature is shown to be true as the response times correlate to the characteristics of each sound more closely than to the overall length of each sample and indicates that principles identified in turn-based gamesmanship are manifested as intrinsic behaviours in response to certain sound material. For example, in chess one player may set up a ‘forking’ move, by moving one of their pieces into a position where it can take one of two opponents pieces on the next move. The opponents response is constrained to moving one of the two threatened pieces, unless the have a counter move that motivates the attacking player to change tactics. Within a conversational model of sound interaction between participants and a mediating framework, a similar level of engagement can be achieved. Participants can learn to control the relative position or trajectory of a sound. They can also shape its’ characteristics, adapting the current sound object to a new context in response to current events. These events can be triggered by other participants shaping their own sound object through interaction, or through their subconscious influence on the overall context. This subconscious influence is created by layered interactions such as a player’s relative location or temporal patterns identified by the mediating software. Temporal information can be gathered from simple biofeedback sensing in a hand held interface or through motion tracking of participants moving within the environment. These local and global events form a compositional structure that has the potential to be easy to engage with but difficult to master. In the example of chess, the moves, consequences, and repercussions, of an action are anticipated by players to various degrees of complexity within a known rule-set. This complexity can encompass the number of moves or options a player has identified, or the likely behaviour or response of a known opponent in a given situation. This 'gamesmanship' depends on their previous exposure to both the game environment and previous games with the same opponent. By identifying a range of participant response to different sound events in a controlled environment, it is possible to provide an adaptive improvisational context. This set of intrinsic behaviours and possible responses can be used to motivate musical collaboration in novices, using a turn-based or conversational model of interaction. The format of this experiment can be applied to a number of discrete designed sound events using the same experiment software and data collection/analysis methodology.

It would be appropriate to run an identical test with a new group but swap the speaker locations to identify whether the response time difference is due to left hand right hand reaction times or left brain/right brain perceptual processes. This possibility is far more clearly indicated on the top tier of four speakers all speakers were equally positioned. There is a significant difference between position 5 (the front left ceiling speaker) and position 6 (the front right ceiling speaker) sounds were identified by either the left or right index finger respectively but the difference is far greater than between positions 1 and 2 which are also identified by the same fingers. The implication is that sounds played on the top left location are perceived faster than on the top right location. As with the accuracy test, location 7 (rear right ceiling speaker) was the hardest to identify with the slowest response time. Considering the lower tier data one would expect position 8 (rear left ceiling speaker) to be equally slow, however it is faster than both the equally distanced right hand speakers indicating that either a behaviour (such as head inclination to right shoulder on sounds perceived behind subjects) or a difference in left/right perception of localised sound objects.

Learn-ability has also been identified as a significant feature of new digital instruments or collaborative music systems. By designing the sound synthesis methods from a perspective informed by the affect of certain characteristics of the test sounds used, it is possible to map the control or synthesis parameters of these elements, to higher-level interaction models. Facilitating an evolving set of parameter mappings relative to each current participant's behaviour. If, for example, the mediating framework of such a system identifies a *relational* or *conversational* interaction model the interface sensor parameter mappings can be extended. This could take the form of enhanced sensitivity, extending synthesis parameters, or a higher degree of influence over the overall shape of the shared composition. Either of these actions can be achieved be the mediating software dynamically remapping individual sensor parameters in context using a structure of *event groups* (as introduced in section 3.3). The principle of this ‘extension’ is to give the individual participant an appropriate level of feedback that can be described as *nourishing* by rewarding exploratory behaviour that is significant within a *dynamic context.* If a lower level behaviour such as *sequencing* is identified, a logical linear response is anticipated to reward this form of timed control behaviour. Interface sensor parameters are remapped to direct object control for sound placement. By designing the mediating software framework to adapt parameter mapping of individual participant interfaces relative to identified behaviour, it is possible to design a framework that promotes a *self organising* shared compositional process. One way to achieve this is to identify *interaction models* to categorise participant behaviour and develop a compositional framework accordingly. Software can then mediate parameter mapping using a set of *event groups* (descriptions of previous gesture-event relationships used to re-route sensors to relevant software modules) in an evolving context, influenced by participants but shaped by previous exposure or encoding. Normally novel controllers or mediated performing environments are designed to function around an established repertoire of actions, identified in software through the use of a *gesture library* to map control actions to relevant musical events.

3.1.1 Quantitative Data

While developing this approach, software to support interaction and cognition experiments was written in Max/MSP. It was designed to collect and organize a wide range of sensor data into stored lists, directly monitoring any interface connected to a computer via the USB Port. Channel assignment supports analogue and digital signals so a wide range of sensors or alternate and novel controllers using can be utilized. The default data mapping was for a commercial wireless customised game controller that is designed for ease of use, combining analogue and digital sensing in an ergonomically resolved hand held interface. As initial prototypes were based on modified game controller PCB's, a number of further experiments were implemented efficiently using groups of undergraduate students as test subjects.

A number of objects are provided within the Max/MSP environment for data analysis. A specific object for MIDI based musical events analysis called *borax* (web reference 10) strips incoming MIDI data into the following streams: *Delta Time* between note on, *Event Number* associated with delta time report, *Duration Value* -sent with note off, *Event Number* associated with duration report, *Velocity* of incoming note, *Pitch* of incoming note, *Number* of Notes currently held down, *Voice* allocation Number, *Event Number* associated with pitch and velocity report. This object alone reports nine properties that can be archived and cross-referenced. When designing test software in Max for collecting quantitative data it is useful to consider the *borax* object (an external for Max/MSP) as a model. Clearly, the object itself is limited to MIDI events but it demonstrates that with a focused analytical framework multiple data can be extracted from a single process. By identifying user actions or gestures captured by an interface against system functions following this model, statistically significant data can be collated from simple task-based testing. The usefulness of the collected data depends on the level of clarity or focus of the task, and of course the *population* of test subjects. Observing a repeated task from a number of perspectives allows higher resolution problem solving for systems design. This can be cross referenced to extract secondary features but this data alone needs *qualifying* for it to be valid when secondary tasks are anticipated as part of a test but are explicitly required of test subjects. The tests are designed for novices with no formal musical skills, the typical target group for novel collaborative interfaces, but the same principles and models can be applied to more sophisticated musical interactions by extending secondary tasks. Different turn-based interactions following a call and response model are designed to motivate a range of behaviours in response to system events. Tasks can be categorized interaction model into groups such as *control*, *exploratory, sequential, organisational* and *relational.* The following design strategy consolidates this process:

Design Strategy 2: Interaction and Perception Testing

Design an experiment to identify interactions and participant responses with interactive content. Consider which study is appropriate; *Control, Correlation* or *Descriptive*. Implement data collection from a simplified version of a proposed interface or novel controller. Test anticipated *interaction models* and identify any new *emergent behaviours*. Use *Analysis of Variance* to analyse data, demonstrate results and to inform subsequent interaction design. Develop subsequent tests by observing user interaction within a controlled experiment. *Testing Observation, and Objective Analysis.*

Control Test

A *control* test is used to determine logical assignment of sensors and actuators to functions. Within an extended instrument model it follows that controls should develop from actions or gestures that are an extension of direct sensor mappings and therefore intuitive, extending *explore-ability*. For a novel interface, ergonomics and feedback should reveal functionality and respond to intended actions or gestures. Within an adaptive sound environment, a simple control test is to ask the user to match a pitch, volume or panning position of a played sample. The target sample is played continuously while the control sample is user activated so they can compare or differentiate between the two sources, using the described control action to complete the task.

Outcomes: *Accuracy, variability* and speed *of resolution* can be measured. A secondary control test is to task the user to match moving or variable elements over time or to map the same control action to different interface gestures and test for the most effective or most used method.

Exploratory test

An *exploratory* test can be used to identify users' ability to identify and locate sound materials and to measure the range of extension motivated by the interface. An example task is for the user to match a button to a sound played. The user is told that each button plays a different sound and that each button relates to a logical location. Each button plays the same sound for the duration of the test but button assignments are randomised for subsequent tests. The task is to match correctly each sound played by the computer by pressing the correct button. First the computer plays each sound individually, then the user is given a set period of time to press different buttons and trigger individual sounds. The task is to associate sounds to buttons pressed; *sound identification and location perception.* The test is set up with two layers of complexity – single sounds directly mapped to individual buttons, secondary sounds mapped to two button combinations. The user will learn of the two button combinations if they realize (*attentive listening*) that there are more sounds than single buttons and therefore try button combinations to find the missing sounds. These additional sounds are made by combining two previous sounds, so user success is dependant on focused interaction. The primary task is for the user to repeat sounds played by the system. Despite the simplicity of the process, the principle established can be transposed to any number of call and response events to develop new musical behaviours.

The quantitative data collected by the software shows the sequence of sounds played (*sound sample)* to number of sounds identified (*location accuracy)*, and the number of combinations tried to number of successful matches (*exploration)*. One can also identify most the most commonly identified sounds and the most common button combinations tried. It is also useful to identify and isolate different user strategies for memorizing sound triggers, i.e. cycling through buttons in a particular order (*sequential*). Results can be compared with a self-report question to confirm findings: number of matches (*accuracy*), spatial perception (*awareness level*) and sound combinations (*attentive listening*). This test shows that although button assignments can be learnt and used effectively, the interface requires practice and logical mapping of buttons to locations for effective interaction.

Outcomes: this test also allows the categorisation of sound types that can be located and identified by test subjects and aids sound synthesis design for this type of interaction. Once a proven framework for a common interface is tested, controls from an intuitive interface can be mapped to related actions. A visual representation of each sample used and the associated test data is was presented in the pilot study.

<u>Sequential Test</u>

A *sequential* test measures the user's ability to identify and sequence sound sample playback. The computer plays a sound and awaits a single button press or gestural response. An incorrect response resets the sequence, a correct response replays the previous sound and adds a new sound, while the user attempts to follow each iteration to recreate the sequence. The electronic game 'Simon' (Web reference 18) follows this model. Users could either memorise each sequence by recalling the order of illuminated buttons or by remembering the sounds played and their associated colour. There is a visual dominance within the traditional consumer 'Simon' memory game, following visual sequences over sound identification. This *sequential* model is useful for training the user, focusing attention and also measuring memory. Again, two layers of complexity can be used, correct identification of the *sequence* of played sounds and recall of *timing* events. The data collected shows how many attempts are required to build a sequence from listening (*short term memory*) and as a secondary or subconscious behaviour, whether the subject included the same time intervals between responses*.* The event time, duration and intervals can also be recorded in relation to user response, generating a pattern for each sequence that can be grouped by success rate. The test reveals different skill levels for recognizing different sounds, matching these sounds to actions and structuring of temporal events.

It also allows a range of strategies to be tried to increase attention to sound objects, and shows that certain sound properties are easier to localize within a group or within a sequence of sounds. This is achieved by recording the *event* (sequence in time), *sound* (which sound sample was played) and *location* (location of sound played on speaker array) comparing this with *userresponse* (which location the user identified) and *userresponsetime* (user response time in milliseconds from start of played sample). The data from multiple tests can be compared to show patterns in sound recognition; which sounds are easier to recall and associate. The higher-level secondary task of timing is not an explicit task but the test will record whether users attempt to match timing. This shows a level of self-improvement, going beyond the basic requirement and demonstrates a musical competence.

Outcomes: This test shows a user's ability to sequence and differentiate sound materials based on a *call and response* model. It can also provide an example of *sociability* in that sequential combinations of sound material can be presented and exchanged with the user. A higher-level test is to allow the user to create a sequence for the system to modify and repeat which the user then modifies and repeats. The system monitors the degree of change in each exchange and modifies the next transformation accordingly.

Organisational test

An *organisational* task is to sort sounds into groups, for example using four buttons to represent four distinct groups with named categories. The user is tasked to sort the played sounds. The sorting task can be used to identify user-perceived sound properties. The naming of groups needs careful consideration and should avoid specialist terms, but equally the categories should be different enough for the user to sort sound material; *melodic* (contains a sequence of notes), *noise* (no identifiable notes), *percussive* (identifiable timing pattern) *spatial* (has identifiable location panning/reverb effect). The data analysis shows comparison of multiple sorting tests revealing patterns or groups of sound with different commonly perceived features.

Outcomes: the test isolates different types of sound into difficulty classes. This is useful for developing a *conversational* approach to real-time sound synthesis where feedback is mediated to motivate sequential actions and is adapted to user gestures or identified behaviours. The pilot study also indicated that sounds within each class were perceived differently depending on relative sound placement. This was not expected, and shows how even simple controlled tests can be useful in revealing new potential strategies for adaptive sound environments, i.e. develop a tracking module that localises sound diffusion to user orientation to sustain or diminish the current interaction/exchange.

Relational Test

A *relational* test is intended to show patterns of association by users across sound object properties. *Pitch*, *Location* or *Motion* can be used to challenge the user to recognize and transform material in response to a series of mediation events combining *sequential* and *organisational* actions. A simple tuning test challenges the user to vary the pitch of one sound to another, using the appropriate interface control. A location test challenges users to identify the diffusion location of a sound by identifying it's perceived source (multi-speaker system) although rather than looking for a direct location match (as in a *control* test) the system monitors spatial relationships, for example: relative location choices. Asking a user to mimic sound object location and motion using a novel controller can test a user's perception of sound object trajectory. A simple task, to copy or mimic an event after listening to a previous example. A secondary relational test monitors complimentary rather than direct match (*control test*) responses: reversing, acceleration, and deceleration.

A pitch-matching task can be extended with sounds of varying complexity, testing the user's *pitch perception.* A secondary task is to build simple note relationships a 3^{rd}, 5^{th}, 7^{th} in response to user actions; the system changes the sample pitch each time a related pitch is achieved by the user, establishing an exchange dialogue.

Outcome: This group of tasks identifies a higher level of performance skill of ear-hand coordination. Simple transforms are used to vary the feedback and stimulate more *playful* or expressive responses. This can be qualified by user report attributing behaviours to the system or identifying limitations of the range of actions supported by the interface.

Each of the tests described performs two functions, testing and refining use-ability and identifying user behaviour or response through structured tasks. Each function can be evaluated by cross-referencing test results across groups of users. A control group approach as presented in the following pilot study, can also be used. Conventionally this is achieved by explaining a task in a different way to one group or selecting two different populations for the same test and noting differences. Using the experiment framework of control groups, it is also useful to present a standard and non-standard interface across two groups, or to present different types of sound material to each group. The first strategy offers function comparison across interface types and the second identifies patterns of interaction that can be influenced by contextually significant feedback, an essential element of adaptive composition.

These tests were designed as a result of implementing the original controlled experiment (pilot study section 3.1). Observation of participant responses was essential to develop further tests to resolve interface design and parameter mapping issues. The inclusion of *exploratory, sequential, organisational* and *relational* tests allows different interaction models to be integrated within a single interface.

Design Strategy 3: Interaction Models – Resolution & Depth

Establish which interaction models are relevant to motivate interaction between participants to extend *instrument resolution* and *expressive depth*.

Exploratory tasks can be used to identify users perception of events and ability to distinguish between sound sources and different types of sound material establishing response patterns and effective interactions.

Sequential tasks can be used to develop collaborative exchange using a *call and response* model. It can also provide an example of *sociability* by tasking participants to order and rearrange events in turns, these sequences can form *repertoires* of sound relationships or gestures that can be learned, shared and exchanged

Organisational tasks can be designed to identify user perception and level of attention to musical events establishing a turn based or *conversational* model.

Relational tasks can be used to establish basic musicianship, challenging participants to *play*, *intercept* and *transform* events using higher level musical skills. Identification of successful exchanges through pattern matching can be used to implement *machine musicianship* and enable different *gesture classes* to be developed.
If a Tangible interface is intended to have the expressive range of a traditional instrument, it must be designed considering a range of *interaction models* that support *gestures classes* or user actions in a shared compositional context: *Control, Sequential, Organisational,*

Relational. This enables participants to *play*, *intercept* and *contribute*.

3.1.2 Qualitative Data

Control, Exploratory, Sequential, Organisational and *Relational* tasks can be designed for a controlled test situation. For the results of such tests, the size and background of population can have a significant influence on results, a structured self-report for test subjects can cast new light on collected data. These *descriptive studies,* where test subjects interact within a controlled framework and complete a questionnaire, allow user intent and interpretation to be revealed. For example, data collection may show a particular pattern of control, implying a strategy on the part of the user. A questionnaire qualifies whether this was an intended action, a *conscious* interaction, a *subconscious* or merely an interface design feature. In the initial *exploratory* test two levels were described, single button sound trigger and dual button sound combinations. Repeated playback of dual sound material and response tracking during the test shows whether this was an accidental or learnt response.

A question-based survey is used to differentiate between an accidental combination and an intended action. More importantly, it is used to identify how the user found the combination during the exploration of the interface/sound dual association process, allowing a level rating for the tested feature to be identified. The following example question provides a self-report framework for the exploratory test detailed previously:

Q. 5 Exploratory Test: Secondary control.

> This test featured several sounds that could not be played by single button presses, how did you respond:
>
> a) I did not realize that some sounds could not be played (move to Q7)
>
> b) I heard sounds that I could not play. (Move to Q7)
>
> c) I tried combining buttons to play these sounds (exploratory unsuccessful)
>
> d) I found some sounds could be played with two button combinations (exploratory successful)

Q. 6 Exploratory Test: Method.

> How did you identify the two button combinations?
>
> a) Trial and error (random)
> b) Logical deduction (sequential testing)
> c) Combined the buttons that triggered these sounds (attentive listening, learning new behaviour)
> d) Another method: (please describe)

This sample question shows how an element of self-report within a structured framework can qualify quantitative data tests. Equally, this quantitative data can be used to verify the process claimed by the user as each approach generates a different pattern that is easily retrieved. The quantitative data alone could be used to imply a particular approach to this form of problem solving process. By inviting the user to qualify this through self-report and comparing responses of multiple users data patterns, different responses can be archived and categorised in relation to user intent. It also enables different levels of engagement to be recorded.

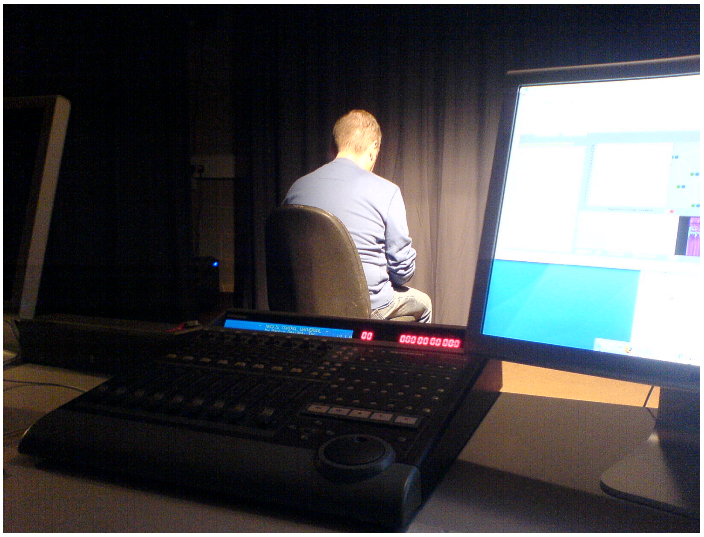
This combined quantitative/qualitative approach to contextualizing levels of interaction is invaluable to design effective communication between an adaptive social composition system and human participants. A method for identifying levels of engagement of participants within different structures or collaborative processes allows mediating software to match current patterns of behaviour with previous behaviours. New distinct patterns can be used to create new system responses. Encoding and decoding behaviours within a mediated environment allows a range of interactions to be identified, learnt and modified.

Design Strategy 4: Descriptive Studies & Empathic Design

Construct questions that reveal whether a participant made a conscious choice for an identified action during controlled tests. In experiment design, provide layers of interaction to reward exploration, which can be identified by user responses using a *descriptive study*. Differentiate between *quantitative* and *qualitative* data but use both to identify interactions and behaviours through collaborative tests. *Comparative Analysis & Learning.*

3.1.3 Summary (section 3.1)

In this section a rationale for interaction design experiments to support applied research was presented. Strategies for integrating qualitative and quantitative data within a controlled test were developed and articulated. Sample questions showing how levels of *explore-ability* can be measured using a layered design approach to challenges for a 'Within Subjects' test. Specific findings in terms of sound location *accuracy identification and* sound location *recognition response times* were presented showing statistically significant differences across diffused sound material in different relative locations. An Analysis of Variance Analysis (ANOVA) was used to identify and present the significant data. Radar charts were used to visualise these spatial and perceptual relationships derived from the full experiment data, which is included in appendix 3 for reference. The process and results of the pilot study were used to establish Design Strategies for Adaptive Social Composition, using software developed specifically in Max/MSP for data collection of interaction within a controlled environment using spatialised audio. The wide range sound material used in perception tests and the documented



(fig. 17 Test environment)

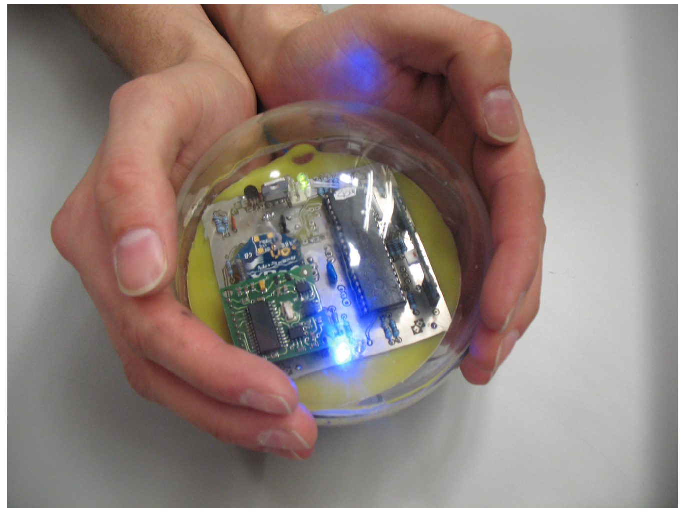
user responses can be used to design synthesis modules to recreate these characteristics in context. An ambisonic B-Format eight channel monitoring environment (shown above, figure 17) was established for undertaking controlled experiments. This is being used for further research implementing the design strategies presented in this thesis.

The next section further unpacks interaction design by considering *conscious* and *subconscious* interactions as significant factors for designing novel interfaces for individual or group interaction within a compositional framework. An integrated approach is proposed combining *interface design*, *collaborative environment,* and *mediating software.*

(fig. 18 Mbira or 'thumb piano')

3.2 Interface Design

Human factors and ergonomics are a significant factor in relation to the effectiveness of any interface, so the physical design of each unit is a careful balance of technology, materials and physical object design. In ergonomic terms, the design was inspired by the handheld Mbira (fig. 18) or thumb 'piano' a carved object or natural gourd which is cupped in both hands, metal strips or keys are affixed to a resonating board or over a hole in a hollow object. These metal strips can be played with either thumb, but variations on playing style are found where the gourd is held in one hand while the fingers are used for greater dexterity, more complex musical patterns are possible, for example using thumb and index finger (Tracey, 1963). For the Orb3 interface design a more sophisticated sensor array than a standard game controller was specified to provide both *direct* and *indirect* gestural acquisition. At the same time the intuitive simplicity of an object held in the palms face up in a gesture of reception or giving with no fixed buttons, switches or obvious controls was inspired by the Mbira. A simple way of combining these two requirements was identified, the simplest hand-held shape that would support a range of known gestures was chosen, a ball. This intuitive physical form suggesting known movements such as cupping, rolling or shaking was chosen to promote *explore-ability.* Wanderley used the term exploratory (Wanderley, 2002) to express how a player would discover new gestures whilst using a new or novel musical interface. *Explore-ability* is a qualitative measure used by Jorda (2003) to express this characteristic in relation to a new or novel interface with expressive potential. The Orb interface design motivates *explore-ability* by combining a natural holding position with a wide range of potential gestures and dynamic sensing. The opaque sphere provides a compact housing for the onboard electronics, it allows the interface to be weighted for stability during autonomous movement while supporting only one holding method enabling interpretation of passing, cupping and rolling actions as gestures. The ergonomic shape suggests familiar actions; lifting, passing, and a subsequent range of possible actions and gestures are designed around the constraints that the physical design places on the user. When held in both hands movement is restricted to small tilting and rotating actions or larger arm movements. This supports degrees of freedom and adaptivity of movement within known parameters. This is a significant design strategy as the physical constraints placed on the known movements of participants



(fig. 19 Orb Interface prototype 3)

allow classes of gestures to be created in relation to these actions. Direct actions fall into a *control model* where the interface is used in a literal way, affecting pitch, panning, and characteristics of the sound object that it controls. To go beyond a control model each Orb is designed to collect data to influence a simple sound object using environmental data such as *ambient temperature*, *light*, and *relative position*. This approach was explored in earlier work discussed in chapter 2. A further decision was for each interface to display autonomous behaviours as a strategy for introducing novices to a collaborative compositional framework. Conventionally interface feedback rewards a direct user action (*control model*). Part of the originality of the design is to communicate possible actions and configurations before the interface is used through autonomous behaviours. These behaviours were developed during the testing of mobile robots, detailed later in this chapter. An important decision was that the interface was designed to support collaborative composition between novices. A group of three participants was identified as the optimum number for *co-creation* using layered interaction models and supporting *user-role flexibility*. To motivate learning and demonstrate potential interaction between Orb interfaces autonomous behaviours were designed, drawing on the *evolving rule sets* evident in traditional board games, specifically Nine Men's Morris. The concept of emergent behaviour was drawn on to develop motion behaviours that participants could identify with. It was important that the interface indicated new ways to *play*, *intercept* and *contribute*. These autonomous motion behaviours, *lead, chase* and *avoid,* implement *proximal relationships* as a compositional parameter. They also represent the interaction models that each interface is designed to support; *sequential, organisational, relational.* The basic principle is that three mobile interfaces move around the floor area of the installation. Each Orb's movement is governed by one of three behaviours that are easily recognized by participants and are responsive to the performance environment or participant presence. The concept was to allow participants to influence interface autonomous behaviours by moving between them. The familiar physical design indicates that each interface can be picked and held, an intentional strategy to invite participation with an open compositional framework. Initially samples were chosen (Pilot Study: section 3.1) to test location perception and response times. Sound objects have been designed based on these findings, while basic process to transform this material using environmental parameters or identified gestures. The initial compositional aesthetic is minimalist using simple textures and frequencies as basic musical material, forming a *malleable music* accessible to novices but with the potential for developing engaging musical structures. An example of this aesthetic can be found in the work Matrix by Ryoji Ikeda (Ikeda, 2000) Simple frequencies, are separated and played on different channels, standing waves are created so that the relative listening position of moving listeners reveals new details or subtle shifts in the perceived sound. This body of work uses simple synthesis and precision sound design to create a sound environment that can be explored by listener movement. As a compositional strategy, this approach is engaging. Extracts from this work are presented on LastFM (Website reference 59). The apparent simplicity of the sonic elements is deceptive, *attentive listening* reveals new sonic structures. Over time, blips and beeps are added adding temporal structure motivating further movement. These works are presented as conventional recordings in stereo. Ikeda has also developed a series of installations exploring the concept. These works are highly resolved compositions and interaction is not invited beyond proximal listening. The approach outlined here identifies potential compositional applications and develops an interface that can adapt to different contexts. The concept of room as instrument, as explored in the work of Alvin Lucier informs environment design. The structured instructions or described gestures for groups of players, as found in the work of John Cage, inspired the concept of adaptive parameter mapping for novel interfaces designed to be used by novices and musicians. The minimalist structured sound design of Ryoji Ikeda's compositions *Matrix* was extended by implementing an ambisonic listening environment to enable sound placement in relation to listener and interface *proximal relationships*.

This integrated compositional approach is identified as a collaborative context where a new form of digital instrument design can engage novices and challenge musicians. A specific intention of the Orb interface design is that it can be composed for, the nature of gestures and collaborative interactions that have been design can be integrated with new musical structures. A *multiplicity of parameters* are provided but these are structured to enable new works to be developed. Technical details of the Orb interface design are presented in chapter 4: Proof of Concept.

Design Strategy 5: Interface Design: Ergonomics and Gestures

For hand held interfaces consider the physical constraints the holding or playing position will impose on the player. Identify potential gestures so that appropriate sensors can be selected to translate data into meaningful events. Consider interface modes to allow alternate parameter mapping to extend *expressive range.* Differentiate between *interaction models* to create gesture classes to provide explore-ability and support *user-role-flexibility*. *Empathic Design, Mode switching and Gesture Classes.*

3.2.1 Mobile Robotics

Two designs have been prototyped, to establish mobility as an interface feature and identify a viable method to reveal system interaction and behaviours in an accessible form. The first using a basic entertainment robot with limited capabilities, the second implementing a video based control system with remote control of motorized wheels.

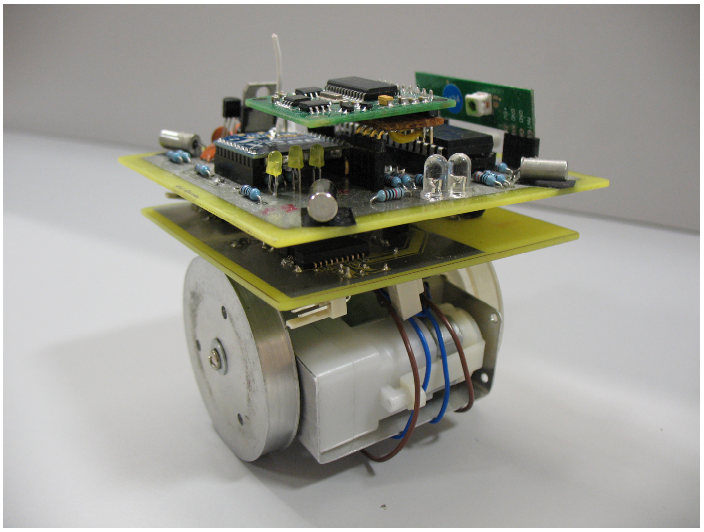
Line Tracker Robot

The initial test used a conventional line tracker robot. The basic line tracker (entertainment or toy) robot has several preset modes and movement patterns. Normally these modes are activated be pressing a switch on the ‘head’ of the robot causing the unit to move in a straight line until a line is detected. A sensor in the base of the robot detects the line causing a gear to be engaged which instantly redirects the robots trajectory at an angle greater than 90 degrees depending on surface traction and remaining battery power. This effectively allows the robot to be constrained to an area by drawing a large circle. A timer on the on-board circuit measures distance travelled or randomises the amount of time allowed before a mode change. If the robot does not detect a line within this timeframe, a second gearing is engaged reversing the motor direction and causing the robot to rotate on the spot. After a short pause, the robot will set off in a new random direction to locate another line. After observing these simple actions, an experiment was designed to explore the potential of these structured yet variable movements. The first experiment used video tracking of a group of three robots with each unit colour coded so the position, distance travelled and distance between could be captured. Each robot’s relative data was mapped to a synthesis parameter of a different virtual instrument so a distinct yet evolving sound could represent each robot's actions. As Woolf and Beck observed (2001) of their own robotic sound sculpture experiments, viewers were quick to attribute behavioural associations to apparent responses to human actions. Some initial tasks designed to motivate collaborative interaction invited participants of the early experiments to add or delete lines using non permanent marker pens on the floor to provide a scoring process for the robots to react to. Although a simple and playful intervention, this exploration revealed the potential of simple geometric shapes both for participants to make and for the tracking system to identify. At this point, the potential for a compositional system that could remap data and adapt to participant actions to influence their subsequent intervention was considered. Strategies for motivation (Bowers, 2001) form a significant part of collaborative sound installations for exploratory interaction and these initial structured tasks revealed new potential behaviours attributed by users to the robots. The potential for incorporating an evolving rule-set associated with each identified motion behaviour *lead, chase* and *avoid* was established. The next technical modification was to remove the onboard mode switch and line sensing allowing the system to activate and change modes remotely via overhead video sensing, using viewer proximity. This reinforced interaction between test subjects and robots as rather than a simple response based on line detection a tangible reaction to proximity and user movement could be mediated using tracking software written in Max/MSP.

The line-tracker robot test led to some useful discoveries indicating that both conscious and subconscious interactions were evident in participants and that playful interaction could be usefully mapped to musical parameters to extend interaction. Autonomous actions were effective although no direct system control was possible other than switching software modes between *Absorb* for environment exploration using the fixed mobility behaviours and onboard sensing to *Adapt* (sensor parameters calibrated for local or *direct* gesture acquisition).

Video Tracking for Radio Frequency Control

The second technical test used radio frequency or RF control. This prototype was designed to refine the interaction between three mobile robots by implementing behaviours mediated by comparing location data in real time with a virtual environment. For this test, a custom PCB was designed to convert the digital outputs from an I-CubeX interface to 'drive' two remote control cars. A vision system, again using a single overhead camera was used to locate each vehicle. The software could then assign a behaviour to each vehicle i.e.; *lead* and *chase* with a secondary behaviour *avoid.* This simple framework allowed the software to control each vehicle and constrain the vehicles to a virtual circuit running in either the same or opposite directions with appropriate actions/reactions. The second part of this test was to provide an open area where vehicles would *avoid* each other but either *chase* or *lead* in response to other moving objects. This approach refined the underpinning action and response models for a collaborative compositional framework based on wireless mobile interfaces exploring a restricted floor area. The potential of emergent behaviours within a collaborative sound space for non-musicians was identified. The main limitation of the second



(fig. 20 Interface prototype 2 autonomous motor behaviours)

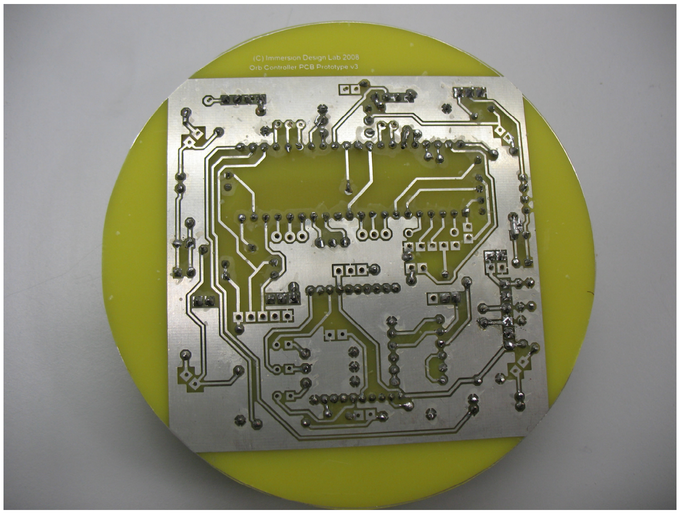
test was speed variability and battery performance dependent on surface traction. Video tracking was effective but control success was variable due to the limited turning circle of the units, which were not designed for this purpose; gear ratios could be modified to address this.

It was evident that a custom motorised wheel system with the positional motion properties of the toy robot and the finite degree of control of the radio control units was required. A design for a custom motorised unit with onboard default motion behaviour and external control by mediating software was essential if the system was to manifest adaptive properties and engaging motion behaviours (figure 20). A specification was developed with technical requirements and details on the forms of movement and direction control required.

(fig. 21 University of Plymouth Robot Football lab)

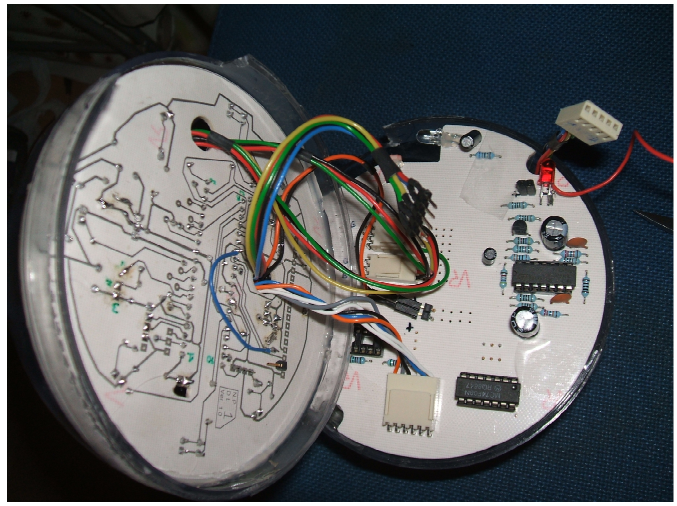
3.2.2 PCB Design

Three customs printed circuit boards (PCB's) were developed in circuit design software Proteus, these were implemented by technicians in the Faculty of Technology, University of Plymouth to develop the Orb interface prototype (figure 21) combining the findings from previous prototypes into a single mobile unit with onboard motion behaviours. The unit can accelerate or decelerate forwards and backwards. It can rotate on the spot supporting alignment with other Orbs. Onboard motor behaviours are stored on a PicAXE40XE chip, these can be over ridden when control data is received from the mediating software. Motion control development was undertaken in the University of Plymouth Robot Football lab, which provides a gantry, overhead camera for video tracking, areas for developing robot interactions and electronics implementation.



(fig. 22 Example printed circuit board)

During the development process several PCB (printed circuit boards) have been used to integrate different system elements (figure 22). These include a board to convert I-CubeX digital outputs to pulse width modulation for radio frequency control of remote vehicles during *lead, chase* and *avoid* behaviour testing. Individual boards for mounting dual axis accelerometers and a digital compass with a custom PCB used to mount specific sensors on a stripped down commercial 2.4Ghz wireless gamepad circuit board The customised unit allows for direct and indirect gestural acquisition by each Orb unit. Each unit is designed for *local* and *global* sensing of gestures and environment parameters respectively, software mediation sets sensing modes according to identified Orb states. After initial testing using modified game controllers, a new modular design integrating three custom designed PCB’s was specified.



(fig. 23 Orb custom PCBs mounted in Orb prototype)

These new circular PCB’s of reducing diameter provided sensing, motion behaviours and wireless communication, and motor control. These circular stacked PCBs fit neatly into a 110mm diameter ball (figure 27). Technical details are provided in Chapter 4: Proof of Concept.

3.2.3 Summary (section 3.2)

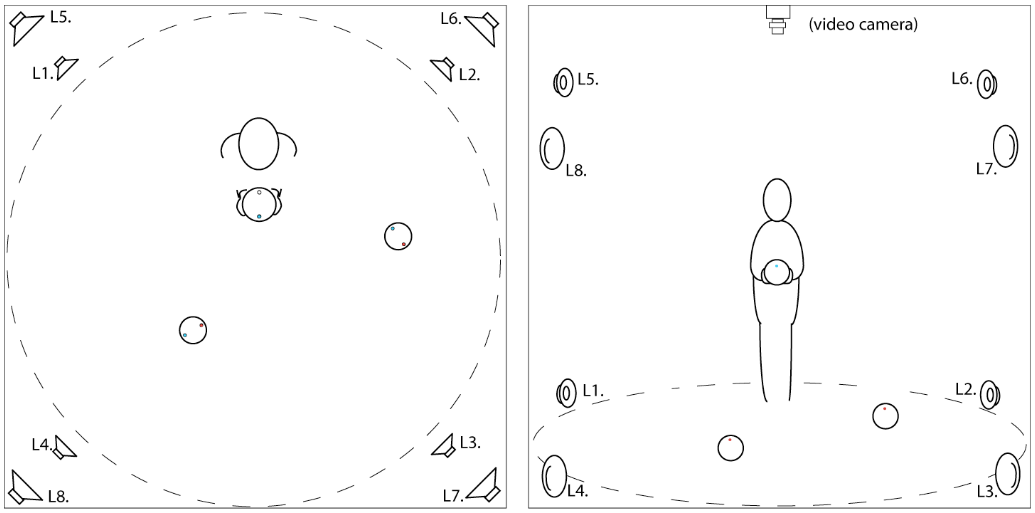
The three Orb interfaces are designed to be autonomous (*absorb* mode), with a set of movement behaviours constraining them to a monitored floor area. In *absorb* mode each Orb interface collects global environment data and is sensitive to changes in its location:

light, *temperature* and *relative proximity*. This simple principle enables the current sound object for each Orb to evolve in response to real-world parameters as perceived by participants, if a shadow falls across an *exploring* Orb interface its representative sound will be affected. The drop in light is registered locally and influences that Orb's sound object. Each Orb can be picked up, manipulated, or gestured with by participants, initially with direct controls. Gestures can be trained and repeated, these are mapped in relation to identified behaviours. An individual sound object is mediated by each Orb interface in *absorb* mode. A minimal approach to sound design is used providing sound material for participants to transform once gestures and actions have been explored. The sound design is intended to give each interface or Orb a tangible identity; a distinct sound object is broadcast relative to its position or motion. When an object is picked up or handled, it switches to *adapt* mode, in readiness for local gestures through participant actions. Sensing parameters are recalibrated and the existing sound object is directly affected by manipulating an Orb. Combined sensor data from each Orb is used to form a background ambience, a minimal soundscape on which each Orb controlled sound object is overlaid.

3.3 Collaborative Environment

The Orb3 environment design is conceived as an Adaptive Social Composition for between three and five contributors within a diffused real time soundscape, forming a scaleable sound installation. An eight-channel four-meter square cube configuration was specified for testing and system development using an m-audio Firewire410 multi-channel interface controlled dynamically by the mediating software framework. The speaker placement is not designed for passive listening as a 5.1 or 7.1 level plane array. Instead, participants are expected to move within the sound environment in response to audio cues, a B format ambisonic system is used. This movement or *proximal interaction* is sensed using overhead camera tracking, identifying 'contributor' positions, where data is sent to mediating software and sound is diffused in relation to participant and Orb locations. The cube layout allows dynamic movement of individual sound objects to be adapted relative to participant or Orb proximity and orientation in three dimensions, above, below, behind in front of individuals. The pilot study and controlled tests (section 3.1) were used to identify sound material users could locate and differentiate.

Within this top down view (figure 24) two Orb interfaces are shown in *absorb* mode. In this autonomous mode each Orb collects data to form a sound object; *light, heat* and *motion* parameters are used to set *frequency*, *equalisation* and *location*. Each Orb has three motor behaviours; *lead*, *chase* and *avoid*.



(fig. 24 Environment Diagram *absorb* mode)

The tracking system intervenes when an Orb could leave the arena or collide with another Orb or participant. Each Orb can be stopped and rotated on the spot to change direction. Each autonomous Orb is counted as a contributor; a participant interacting directly (holding) an Orb is also counted as one contributor.

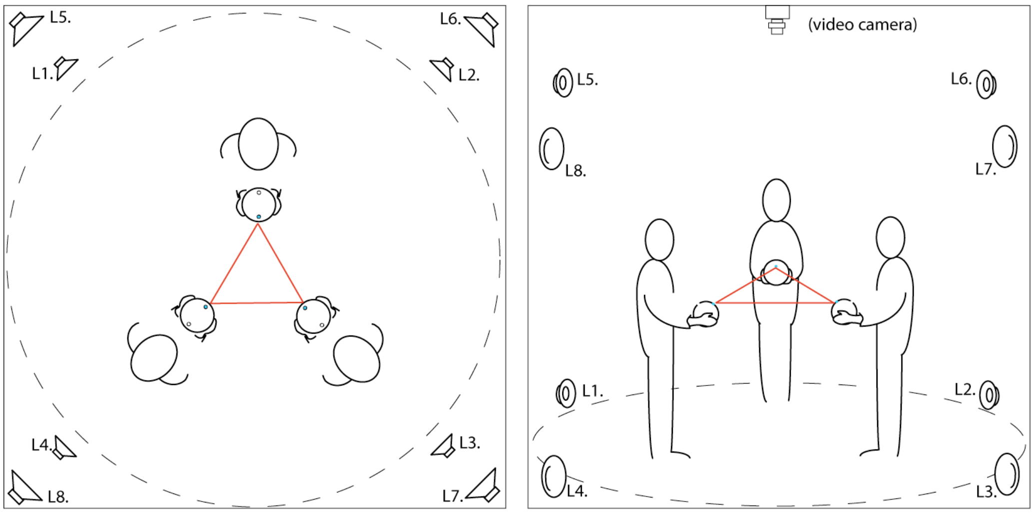
A participant interacting indirectly also counts as a single contributor, so the system design supports up to five concurrent human or robot 'contributors' with behaviours attributed depending on their interaction or current compositional dialogue. Basic Symbols are used to map the relationship between participants and/or Orb interfaces with compositional parameters remapped when a new symbol configuration is detected.

This environment diagram (figure 25) shows three Orb interfaces in *adapt* mode, when held by participants, sensor parameters are remapped to detect gestures. The overhead video camera is used to detect *proximal relationships* by comparing location data (overhead tracking of interface LED alignment) with Orb orientation (variable generated by onboard digital compass).

In this example, a triangle is created when each participant holds their Orb interface facing the centre of the tracked area. The two LEDs of each Orb are detected by the overhead video camera, two different colours, white and blue are used in adapt mode, white and red are used in absorb mode, allowing orientation and relative position to be plotted. A digital compass within each Orb also returns the current relative angle. This makes tracking of actions more reliable by combining both elements. Standard symbols used for proximal interaction in this design are: *Line, Triangle, Square* and *Circle.*

- *Line (trajectory)*
- *Triangle (envelope)*
- *Square (time and rhythm)*
- *Circle (memory/diffusion behaviour)*

Each symbol is allocated to a different synthesis parameter or diffusion process to create a dynamic sound environment.



(fig. 25 Environment Diagram *adapt* mode)

This approach is designed to motivate *sociability* by integrating participant collaborative behaviours that can be identified by software and mapped to different musical processes. Participants may discover these *co-creation* actions by exploring the influence they have over an Orb's sound object through Orb gestures, orientation and proximity to others. Just as in a card game, new combinations can be introduced and rewarded. Aural feedback is provided when a symbol is recognised, a *transformative* process is applied to the current background sound.

A *line* symbol initiates sound trajectory events. Sound object locations are relocated and aligned by the mediating system. Orb data parameters are mapped to logical control models. This sets all three Orbs to a *control model* of interaction. Each Orb directly controls the sound placement of an individual sound object. A *line* is detected when three contributors are identified in a row. This possibility is demonstrated when an Orb in autonomous mode manifests a *lead* motion behaviour and two Orbs manifest *follow* behaviours forming a visible line. Participants can observe Orbs in autonomous mode to learn potential compositional actions.

A *triangle* symbol changes the envelope for the current ambient background sound. This symbol is detected when each Orb is oriented towards a mutual centre

A *square* symbol influences time structures within the active sound objects. Contributor parameters stored in short-term memory are recalled and recalibrated based on time scaling algorithms. Relative distances between four contributors are measured.

The *circle* symbol is activated when five contributors form a circle, this is the maximum number of participants the system is designed to relate to. Each active sound object is stored in its current state and archived to memory. Each sound is given a rotating behaviour based on previous control interactions and contributor actions are remapped to transformative actions (synthesis parameters) in preference to previous control parameters until the related Orb interface is replaced on the floor.

The detection of proximal relationships over time forms a compositional structure within which participants engage through Conscious and Subconscious interactions. The conceptual framework is designed to provide accessibility while motivating group musical behaviours and collaboration with adaptive interfaces, which supports a *conversational* model.

For integrated systems designed for novices, the typical venue is a gallery, social venue or public event. Factors such as through traffic of people are significant as is the potential behaviour of participants, so the choice of environment is significant, although the tracking, data transformation and diffusion software can be calibrated to an appropriate venue. A symbol based approach allows individual and group behaviours to be incorporated into a meaningful compositional context that reveals potential interactions to participants by example.

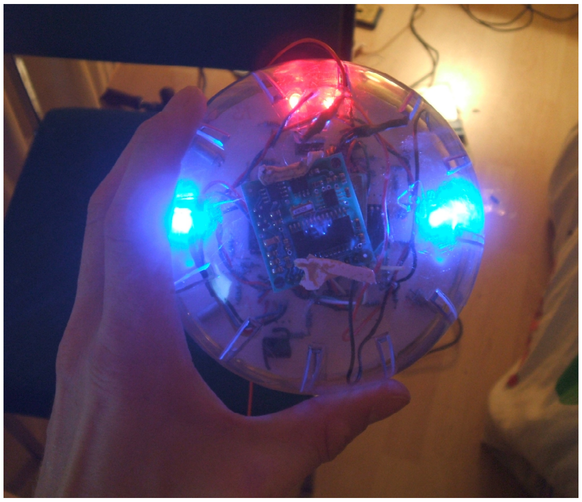
3.3.1 Conscious Interaction

Studying different interaction models, observed in related musical interfaces and performance practice, developed the Orb3 overall design approach. Groups of conscious behaviours are applied within mediating software that is structured to correlate participant actions with known or learnt musical behaviours, integrating concepts from related multi-user interfaces for musical collaboration. Conscious interactions can be interpreted by identifying patterns of *local* interface control in relation to current sound events; compositional processes are embedded in structures that can be adapted to suit either a *direct gestural control* (user interface manipulation) or encoded participant movement and location within described tracked relationships, a spatial score matching learnt symbols or proximal relationships. If the system identifies a known symbol, *line, triangle,* *square, circle* the relevant sound event or process is activated. Some of these patterns can be observed while Orb units are operating in autonomous absorb mode.

An interface's *explore-ability* can be measured by observing the levels of engagement through established interaction models. Mediating collected data can extend this: identifying a *control* behaviour in the context of concurrent musical events, embodying conscious interaction. Dynamic parameter mapping can extend *learn-ability* if the transformations are related to perceived system-mediated relationships. The term *adapt* was presented as a field-specific term for categorising identifiable exploratory control behaviours. Figure 26 shows an Orb interface prototype used in adapt mode. As contributors learn new local gestures in response to system cues their performance repertoire evolves. When participants become aware of global collaborative symbol-based proximal interactions, new performance strategies emerge, as contributors explore possible actions collaboratively.

3.3.2 Subconscious Interaction

Within the Orb3 system design, Subconscious Interactions are classes of behaviour or patterns of actions and responses, which can be identified through *indirect gestural acquisition* (user or autonomous interface proximity). These behaviours represent relationships that can tracked and compared to *global* events such as participant location and can be identified as a compositional element.



(fig. 26 Orb interface prototype in *adapt* mode)

In absorb mode, when an Orb is placed on the floor, it behaves autonomously. Motor behaviours are used to propel the Orb around the performance area. An example application of this is when an Orb in *lead* mode approaches a participant causing them to move, initiating *follow* mode as the participant attempts to move away. Because the interface behaviour is structured, the subconscious movement or relative position of participants can be recognised in software (participant relative position is monitored using video tracking explained later in this section). This participant presence causes the Orb interfaces to adopt a new behaviour and also affects global synthesis and diffusion parameters. Just as an Orb unit is absorbing environmental data (global) through autonomous behaviours the participant is subconsciously absorbing system influences by experiencing sonification of indirect actions. If a participant intersects an Orb's trajectory in *lead* mode the *avoid* behaviour is triggered. The term *Absorb* was proposed as a field-specific mode for describing these indirect behaviours that nevertheless have a significant influence over compositional processes within an adaptive framework. The point at which individual participant movement is transformed to a conscious interaction is either when they pick up an interface and interact with it directly, or when the move their position in relation to identifiable geometric positions forming patterns with other participants, a learnt behaviour in relation to observed system events. The following design strategy consolidates these principles:

Design Strategy 6: Evolving Rule Sets

A collaborative compositional environment combines interaction, interfaces, gestures and behaviour. Each of these elements is complex so an integrated approach is required. Proximal Interaction can be used as a compositional parameter. For this to be effective it is useful to introduce simple actions that are tangible to participants. Evolving rule sets are ideal for engaging novices by revealing possible interactions in context. Interface modes can be used to demonstrate different possible actions and motivate exploratory behaviours. *Integtegration, Context, Emergence.*

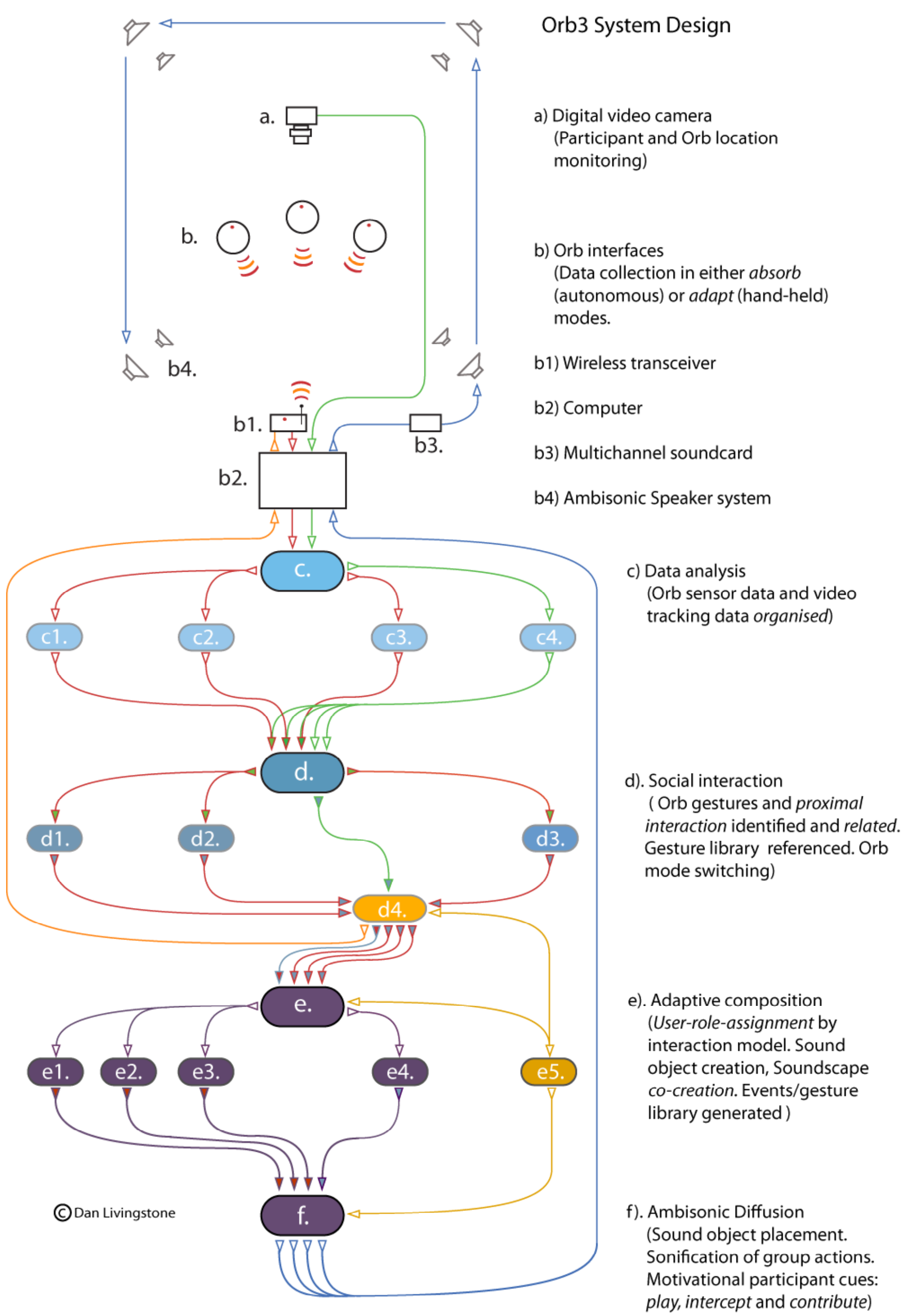
3.3 Mediating Software

Each Orb interface can hold a small program to format sensor data, control autonomous motion and switch between behaviour modes *absorb* and *adapt*. However to mediate interaction between Orb interfaces and to process participant gestures an integrated software framework is required. Software developed in Cycling74's Max/MSP 4.5 and Jitter 1.5 was used to prototype a mediating framework to identify gestures and assign roles to participants using Orb interfaces. The design also includes sound generation and diffusion.

A modular approach is essential when designing this type of system. it is very easy for any single element to become over complex. Software complexity is not a desirable feature if colleagues are to be invited to create new compositional structures. Each element and associated process should be broken into modules or 'patches' to provide efficient communication between them. A separate module or 'patch' is recommended for each interpretive or transformative process that contributes to the compositional framework. A significant challenge with this type of system is developing a low latency method of analysing real-time data streams from different sensors. In a real-time software-based system, identifying the start or end point of a gestural event (Nakra 2000) is not a simple task, even though the participant or viewer may recognise a change in action or control. By processing the data from each Orb interface separately, it is far easier to identify significant events in streams of potentially *meaningless numbers* (Zicarrelli 1991).

The following diagram (figure 27) presents a system design for adaptive social composition. The design is developed from the principles identified in chapter 2 and applies the design strategies discussed in this chapter and combines novel handheld physical interfaces with mediating software. The software design has been developed through modular testing of core features; *data analysis, social interaction, adaptive composition* and *ambisonic diffusion.* The software modules incorporate third party software open source externals for the MaxMSP authoring environment. A summary of each module is provided with the diagram. Each module is further unpacked in the subsequent text. Where a specific MaxMSP object or external is mentioned it is named in *italics*.

The intention of this section is to show a modular design process combining methods from different fields of Computer Music, and the relationship between each system element or software module. This is presented as a rationale for the design of new interfaces that support different classes of gesture through considered interaction design for collaboration. It demonstrates how participant gestures can be remapped to compositionally relevant parameters to motivate different musical behaviour. It applies the principles of *explore-ability* and *nourishment* by providing an adaptive context for collaborative actions using *user-role-flexibility*. Gestures and participant relative locations are integrated as compositional parameters that can be mediated by software within a compositional framework. *Co-creation* is embedded in the interaction design to engage novices and challenge musicians within an evolving rule set combining gesture, interaction model and group behaviour.



(fig. 27 Orb3 system design)

a) Digital video camera. Captures top down view of performance area. Live video feed is sent via firewire cable to data analysis module c. Used to monitor Orb interface and participant position and orientation.
b) Orb Interface. Supporting data collection using the following sensors: non-mercury tilt switches, light dependent resistors, dual axis accelerometer, digital compass, ultrasound. An XBEE transceiver sends/receives data wirelessly which is received by a fourth XBEE transceiver (b1), this is connected to computer (b2) running MaxMSP software (modules c to f). Each Orb includes a PicAXE40XE chip which is programmed to format live sensor data into lists and switches onboard motor behaviours. Data sent includes 5 digital signals and 8 analogue variables per Orb Interface.
c) Data analysis software (*organisational*) module (MaxMSP). Data is received wirelessly using the 2nd XBEE transceiver (b1) connected via universal serial bus (USB) port. Data received via serial object is processed in MaxMSP. Up to three sets of Orb sensor data (colour coded red) are processed in independent sub patches (c1,c2,c3). Each sub patch receives 5 digital signals and 8 analogue variables, this data is organised into separate lists using *spray*, *funnel*, *pack and mtr* MaxMSP objects for subsequent parameter mapping. A further sub patch (c4) performs motion analysis on the live video feed a. (colour coded green) to calculate position and proximity of each Orb interface. Third party externals from the *CV.jit* library (motion tracking, colour tracking, edge detection) are combined to extract XY coordinates for each Orb. Participant presence is also detected

using this method. Sensor data from each Orb interface is combined with positional data extracted from the camera feed. This combined data is grouped and labelled for each Orb interface. Up to two additional participant locations are also reported as XY coordinates (colour coded green). These 5 data groups are sent to d.

d) Social interaction software (*relational)* module. This module processes combined sensor and location data from each Orb interface (colour coded red/green), and location data for two additional participants (colour coded green/blue). Pattern recognition is used to identify different gestures from each Orb interface. For example; a rolling gesture of a held Orb will activate each tilt switch in turn, when a tilt switch is first activated a timer is activated, the time between subsequent tilt switch activations is used create a pattern, formatted as a list. Gestures are designed around interaction models; *control, sequential, organisational* and *relational*, parameters are remapped or filtered according to the identified interaction model. A separate group of sub-patches (d1,d2,d3) is provided for each Orb interface to identify and process gestures. As each gesture is designed around a known set of participant actions, each sensor is positioned on the physical interface so that initial patterns are known and can be trained and stored as lists. The third party external *nnlists* for MaxMSP is used to train new patterns or compare a current pattern with a previously learnt pattern stored in list form in archived external text files (e5). This external provides a simple to use method for comparing and training MaxMSP lists. It provides a feed-forward back-propagation artificial neural network implemented in JAVA. Multiple instances of this method are used to identify different patterns in each Orb interfaces' live data. (a more detailed breakdown of Orb gestures and parameter mapping is provided in chapter 4: Proof of concept). In addition to gesture identification this module also identifies *proximal relationships*. Each Orb interface onboard digital compass reports its' relative orientation as a variable. This variable is combined with the XY location coordinates for each Orb. This combined position and orientation data is sent to a further sub-patch (d4) to identify group behaviours. This is achieved by creating a new video data feed (colour coded green/blue) using *jitter* processing based on the known relative position and orientation of each Orb interface and the location of up to two participants. A grid based motion detection method using CV.jit externals for video analysis is used to identify *proximal relationships* (d4). This sub-patch also relays mode switching (colour coded orange) to mobile Orb units in absorb mode (when it is not held by a participant). The motor behaviours *lead, chase* and *avoid* can be triggered by the software for each Orb in relation to Orb or participant locations. Each behaviour is programmed into each Orb hosted by onboard PixAXE40XE chip. This has two objectives; generate new proximal relationships and engage participants by demonstrating behaviours in context. The design identifies and groups this positional data into compositional events when specific formations are identified (as in the examples provided earlier on pages 240,241.) These *proximal* events are encoded using symbols: *line, triangle, square, circle.* The social interaction module identifies the current gesture for each Orb interface. The interaction mode of each gesture is identified and data not relevant to that gesture is filtered. The filtered gesture data (colour coded red/blue) is routed to module e. If a group compositional or *proximal interaction* is identified the sensor data from each included Orb unit is combined and re-mapped to co-creation or *transformative* events. This new data (colour coded blue/red) is also forwarded to module e. for musical processing. Individual Orb gestures influence a sound object associated with that Orb interface. Group behaviours influence the background sound layer collaboratively and override current individual parameter mappings.

e) Adaptive composition (transformative) software module. This module receives Orb sensor data which has been filtered, formatted and categorised into individual gesture classes (colour coded red/blue), or group behaviours (colour coded blue/red). Gestures encoded as lists are assigned to different musical processes. Algorithms from Karl Heinz Essl's open source *Real time Composition library* or *RTC* for MaxMSP are used for processing these lists, providing pitch related functions (harmony, scaling, generation) time based functions (rhythm, pattern, brownian motion) and envelopes (dynamics, ramps). For example; the object *showchord* shows the pitch of a chord of midi note numbers. *Trans-pitch* transposes pitches by a specified interval. *Brown-rhythm* and *brown-melody* use Brownian motion to generate new material. *Ratio* is used to select list elements that are repeated, it uses the "series" selection method. R*andom-ramp* produces envelope shapes that fluctuate within a given dynamic scale. This module *transforms* the Individual and group gestures into sound material dependent on the identified interaction model: *control, sequential, organisational* and *relational* (these gesture classes are presented in chapter 4: Proof of concept). Each individual gesture is mapped to different parameters to either directly control a current sound object or to structure sound material or object placement. This assignment is mediated by comparing Orb gestures and the musical context of the sound objects produced. The compositional framework is designed so that new configurations can be added, sound synthesis methods changed and symbol based parameter mapping extended. A set of default mappings have been conceived to illustrate the potential of such systems. The Adaptive composition module passes the individual sound object created for each Orb (e1, e2 and e3) with new positional data to the diffusion module f (colour coded purple/red). The overall background soundscape mediated by group proximal interactions is also sent to f (colour coded purple/blue). This module can interject musical processes from previous sessions to motivate or reveal new interactions that have not been performed by participants. These previous events are stored in text files and form a simple gesture library of encoded events as lists (e5) a form of long term memory (colour coded fawn). This is activated when participants do not use or explore the available repertoire of gestures (identified in d4). For example, if a novice continues to repeat a control gesture such as sound object panning that gesture will be transformed, remapping the data generated to an alternate event, motivating further exploration.

f) Diffusion software (relational) module. This is used to position the sound object of each Orb relative to its' position and orientation within the physical performance area. Each sound is diffused relative to patterns or sequences generated by the Adaptive composition module in response to Orb proximity and user gestures. The background soundscape created from collaborative proximal interactions is diffused to reflect the identified *co-creation* event in context. For example different elements of the background sound can be diffused in different locations if positional data from module e5. is received. Stored sound events from e5. can be introduced in relation to participant *control* or *sequencing* gestures to motivate participant *play*, *intercept* or *contribution* behaviours.

The simplest method identify gestures is to compare a stored list with an incoming stream of numbers, there are numerous third party externals or objects available for Max/MSP that provide basic functions for analysing lists. A potential failing of this method is that even a known gesture is more likely to be identified after it has been performed. There are more sophisticated ways of analysing and interpreting live data streams, however many of these are processor intensive. A range of algorithmic objects and externals have been developed to support computer musicians creating new interactive works. Within the MAX/MSP community, one of the long established collections is the '*Real Time Composition Library*' provided as open source software by Karl Heinz Essl (Website reference 19). Olaf Mathews also provides a set of externals '*Artificial Tango Library*' for Max/MSP providing musical event analysis. The library supports generation, recognition and analysis of musical events. It uses algorithms for detecting chords, rhythm tonal sonority and other characteristics and is under continuous development. The '*Artificial Tango Library*' (Website reference 55) is based on the previous work of Hutchinson & Knopoff (Hutchinson & Knopoff 1978) and Jarno Seppänen (Seppänen 2001). As expressed previously, understanding the context of actions or gestures with a novel interface is extremely useful, and allows the composer or designer to group actions or gestures for effective mapping to sonic events. Even using the simple method of comparing a stored list with an incoming data stream can be used to enhance *explore-ability* and add layers of interaction to extend engagement. For example if the interface follows a control model with direct mapping to perceived parameters it is not necessary to identify a gesture algorithmically as it is sonified in real time. However, if list comparison is used to detect a repeated gesture it can be categorised as following a sequential model and the current parameter mapping to control a sound object can be superseded to one where gestures influence the sonic or temporal characteristics. To simplify the process of data analysis from the Orb interface design third party externals are used in the mediating software, to support an adaptive framework. Equally important is the use of interaction models to provide an evolving 'rule set' for contextually relevant dynamic parameter mapping. This approach provides a conceptual framework for adaptivity. It integrates simple methods for live data comparison with established algorithmic composition tools within identified interaction contexts.

An *adaptive* system that accommodates different interaction models and which *evolves* with a range of participant behaviours is more likely to motivate effective collaborative. For this reason the behavioural context of system, modules should also be considered. The Adaptivity of such a system can be implemented by providing a strategy for processing various data in context; within a simple board game, a set of known rules and potential strategies shapes the action. Similarly, in this design, the adaptive composition module includes strategies to identify and channel identified actions in context through *user-role assignment*. This involves grouping participant behaviours by interaction model through identified gestures. An evolving rule set is created by identifying actions or gestures in context, relating them to previous activities and enabling new characteristics to be revealed when gestures are layered or repeated.

As discussed in chapter 1, investigation of a range of novel or tangible interfaces revealed that novice participants actions will coincide with at least one of the following interaction models while using a tangible interface; *Exploratory, Organisational, Sequential, Relational.* For the Orb3 system software design each of these identified interaction models has a different group of associated activities or actions mediated by software in response to live data combining *direct* and *indirect* gesture acquisition. In this design anticipated gestures are grouped by interaction model with associated software patches for identification within the social composition module (d) (fig. 27) The structure of such a software environment can be divided into groups of modules with equally clear tasks; *Data analysis, Social Interaction, Adaptive composition and Ambisonic diffusion.* Within the example system design, (fig. 27) the adaptive composition module (e) presents the concept of *event groups* to associate actions and behaviours to assign parameter mapping dynamically in context. In principle, this also provides an open framework for others to create new compositional strategies for shared composition using collaborative interfaces.

Data analysis groups provide data capture from each wireless interface unpacking the data received from the USB port and scaling it to relevant variables or translated to system messages. Data capture from overhead camera provides two-dimensional positional tracking of Orb interfaces and participants (c4). Modules within these groups are designed to collect live data from *exploratory* actions. In the author's design the Orb interface modes *absorb* or *adapt* are a significant design strategy for dynamic parameter mapping. Data received is rerouted and calibrate data according to the detected mode. Data analysis groups follow an *organisational* model.

Social Interaction groups are for routing and parameter mapping the live data and channelling it to relevant synthesis and composition modules. These modules are structured so that incoming data can be formatted for different tasks depending on each Orb's identified gesture. One sub-patch (d4) module monitors the video data for known movements that emerge from different autonomous interface behaviour modes; *lead, chase, and avoid*. Social interaction groups follow a *relational* model.

Adaptive composition groups are designed to mediate musical events. These include sub-patches for sound synthesis and transformation of gestural control when group behaviours are detected. These sub-patches map identified gestures to musical processes. Algorithms are used to mediate musical relationships between participant created sound objects. These *relational* actions require a learning process, enabling the system to identify different behaviours through the actions and motion of participants in relation to the local gestural data they create. These behaviours can be evolved from simple algorithmic processes, just as the structured rules of traditional board games can motivate strategies of *attack, counter attack, subterfuge, distraction* and *anticipation* between players, and in the same way as structured interactions are framed by established rules. The system is provided with a basic repertoire of musical events or motivational cues that can be recalled and repeated using list comparison. In the design presented in this thesis formatted lists are used to represent, compare, store and recall gestures and musical sequences. These motivational cues are stored as text files, grouped by interaction model and can be triggered when a predefined event is identified. The principle of gamesmanship is identified as a valuable strategy for motivating learning. For example; an *attack* action would be to play a sound behind the participant to announce the presence of a virtual player, or to introduce a new sound not generated by participant actions. A *counter attack* is to implement a tangible response to a participant action, echoing an action but extending its complexity, presenting a challenge. *Subterfuge* is implemented by masking sound elements or removing tangible control of the interface, or by remapping parameters to a new function causing participants to adapt their behaviour (interaction model) from *control* to *relational*. *Distraction* can be utilised by introducing particular frequencies located spatially to motivate subconscious movement of participants (This aspect of system design was developed during observation of participants during controlled experiments (pilot study 3.1) investigating sound location perception accuracy and response times of individuals. Adaptive composition groups follow a *transformative* model.

Ambisonic Diffusion groups are used to position sound objects and co-created sound material within the performance environment. This supports a process of shared composition. An individual can directly control his or her sound object location (*control model*). Sounds can be placed in response to these actions (*relational model*) either by participant gestures or software mediation. The spatial relationship between sound objects can adapted when new behaviours are identified by other system modules. Motivational cues, stored as lists, can be recalled to motivate new behaviours or to respond t individual or group actions through tangible sound positioning (sequential model).

3.4 Integration

It is not the intention of this system design to prove that emergent behaviour has been manifested either by the system or by participants, however the approach is conceived to incorporate new gestures (*local, direct*) and identify recurring compositional events, assigning a behaviour to identified actions to mediate responses and trigger motivational cues. With such a complex combination of elements mapping digital and analogue data to a variety of interactions, synthesis, and contexts, it is essential to produce a modular system with considered transformation of data between logically structured system elements. For non-programmers an ideal environment for visual programming of this type is Cycling 74's commercial version of Max/MSP where each process can be implemented as a separate 'patch' which can be modified and refined independently. A significant problem with such an approach is the danger of implementing CPU (computer processing unit) intensive modules that when combined overload the system and reduce performance or generate errors. An effective solution to many specific tasks such as data processing, synthesis, musical event analysis, video tracking and diffusion can be to utilize some of the many external objects created by the max community. These objects and third-party patches often operate far more efficiently than extended patches of the standard externals and are used in the Max/MSP software environment to optimize performance and streamline development times. Many such systems are dependent on a range of third party externals hosted by the Max/MSP framework. The concept of resolution has been discussed in terms of system integration, and it has also been applied to software design. Rather than produce a system that collects a high resolution of data and therefore generates a far higher load on the CPU to process it, a number of lower resolution inputs are combined within an adaptive framework to capture a higher resolution of combined events. To plot a curve one only needs positions along that curve, or to observe it from different perspectives to remap its function. As the software is required to adapt to different inputs and varying data types it is useful to identify the least CPU-heavy methods for mapping this data and the sampling rate required for each element in order to maintain a reliable overall system. There is always room for improvement and with the continual development of externals and third party implementations in addition to related environments such as Processing, Pure Data and Eyesweb with associated gesture libraries, there is endless choice as to which software environment is most effective for a given system design.

For the purpose of designing this adaptive compositional framework, development has been restricted to Max/MSP 4.5 Jitter 1.5 with selected third party externals to perform specific tasks efficiently.

The primary method has been to reduce each system element to it's simplest form with clearly defined functions for each element, to observe the results of each separate element through prototyping and to document and publish initial findings for peer review at the key international conferences for the field. The overall system integration can best be explained using the model of a board game. We can visualise the participants and Orb interfaces as game pieces with their known and potential actions mapped as moves within the game space or sound environment. By providing an adaptive framework and incorporating a hierarchy of potential interactions (interaction models) and parameter mappings (data groups) it is possible to provide an evolving rule set. This provides the potential to create new data relationships to process actions during the current 'game' creating new compositional behaviours. In games design the concept of a 'combo' or 'special move' is well established. Simply put, this is the simultaneous performance of two known actions at a particular time during gameplay to achieve a higher-level goal. This strategy is extended within the adaptive framework presented. Gestures are grouped within different interaction models with associated actions, and data patterns are grouped with musical events. This allows the sensor mapping of each interface to be changed depending on the current context. If the software identifies a layered interaction (a gesture form one interaction model performed at the same time as a gesture from a different model) a new parameter mapping can be assigned. Musical event analysis can be used in real-time to identify the context of a gesture, and apply a different process or parameter mapping. This provides a depth of interaction beyond the basic control model used by many novel interfaces. It provides the potential to evolve new behaviours in participants by remapping sensor parameters within a tangible compositional exchange, and to identify repeated new actions in context to evolve new material. It is believed that this approach sustains *performative engagement* (Paine, 2004) by providing a compositional framework that evolves through exposure to new participants or collaborative events, allowing it to manifest new compositional relationships as it evolves new these behaviours.

The following design strategy sums up this design approach:

Design Strategy 7: Integrated System Design

An adaptive social composition system combines a collaborative sound environment with novel interfaces and mediating software. Any single system element could take considerable expertise to resolve. Break down the design issues using field specific terms and identify core *features*, *models, qualities* and *behaviours*. Use conceptual groupings to structure software modules and allocate tasks. Identify hardware and software requirements to focus development. Use a system diagram to design overall program structure and balance processing loads. Research third party resources to streamline software implementation. Use a modular development approach so each feature or function can be prototyped and tested independently. Principles of gamesmanship can be used to motivate novices and engage musicians, implement a *transformative* interaction model for mediating software. Combine *gesture classes*, *proximal interaction* and *interaction models* to provide a higher resolution of system events. Use conceptual tools to simplify complex interactions into manageable system events. Use software groups to structure and distribute processes within a shared context. *Acquisition, Analysis, Interaction, Composition, Diffusion.*

The following section presents seven design strategies for Adaptive Social Composition. These condense the principles established through the research presented in this thesis using terms drawn from the field specific vocabulary introduced in chapter 1 (page 105 Glossary of Terms)and contextualised in chapter 2.

3.5 Design Strategies (resolved)

Design Strategy 1: Interface platforms for prototyping novel interfaces

Consider user actions and potential gestures *indirect, direct, haptic, non haptic*. Identify appropriate sensors for data collection *accelerometer, ultra sound, digital compass, tilt switches, force sensing resistors, light dependent resistors* etc. Choose appropriate platform for testing, then implement standalone interface/novel controller once technologies and interactions are resolved. *Concept, Design and Implementation.*

Design Strategy 2: Interaction and Perception Testing

Design an experiment to identify interactions and participant responses with interactive content. Consider which study is appropriate; *Control, Correlation* or *Descriptive*. Implement data collection from a simplified version of a proposed interface or novel controller. Test anticipated *interaction models* and identify any new *emergent behaviours*. Use *Analysis of Variance* to demonstrate results to inform subsequent interaction design. *Observation, Testing and Objective Analysis.*

Design Strategy 3: Interaction Models – Resolution & Depth

Establish which interaction models are relevant to motivate interaction between participants to extend *instrument resolution* and *expressive depth*. If a Tangible interface is intended to have the expressive range of a traditional instrument, it must be designed considering a range of interaction models that support *gestures classes* or user actions in a shared compositional context: *Control, Sequential, Organisational, Relational.* This enables participants to *play*, *intercept* and *contribute*.

Design Strategy 4: Descriptive Studies & Experiment Design

Construct questions that reveal whether a participant made a conscious choice for an identified action during controlled tests. In experiment design, provide layers of interaction to reward exploration, which can be identified by user responses using a *descriptive study*. Differentiate between *quantitative* and *qualitative* data but use both to identify interactions and behaviours through collaborative tests. *Comparative Analysis & Learning.*

Design Strategy 5: Presence and Influence

Participants can be engaged by observing interactions in progress. *Proximal interaction* can be used to reveal musical relationships. Symbols can be used to encode *Direct* and *Indirect gesture acquisition* using overhead video analysis of participant locations data analysis of interface sensors. This can be used to motivate participant interaction and identify group behaviours. *Sociability.*

Design Strategy 5: Interface Design: Ergonomics and Gestures

For hand held interfaces consider the physical constraints the holding or playing position will impose on the player. Identify potential gestures so that appropriate sensors can be selected to translate data into meaningful events. Consider interface modes to allow alternate parameter mapping to extend *expressive range.* Differentiate between *interaction models* to create gesture classes to provide explore-ability and support *user-role-flexibility*. *Empathic Design, Mode switching and Gesture Classes.*

Design Strategy 6: Evolving Rule Sets

A collaborative compositional environment combines interaction, interfaces, gestures and behaviour. Each of these elements is complex so an integrated approach is required. Proximal Interaction can be used as a compositional parameter. For this to be effective it is useful to introduce simple actions that are tangible to participants. Evolving rule sets are ideal for engaging novices by revealing possible interactions in context. Interface modes can be used to demonstrate different possible actions and motivate exploratory behaviours. *Integration, Context, Emergence.*

Design Strategy 7: Integrated System Design

An adaptive social composition system combines a collaborative sound environment with novel interfaces and mediating software. Any single system element could take considerable expertise to resolve. Break down the design issues using field specific terms and identify core *features*, *models, qualities* and *behaviours*. Use conceptual groupings to structure software modules and allocate tasks. Identify hardware and software requirements to focus development. Use a system diagram to design overall program structure and balance processing loads. Research third party resources, to streamline software implementation. Use a modular development approach so each feature or function can be prototyped and tested independently. Principles of gamesmanship can be used to motivate novices and engage musicians, implement a *transformative* interaction model for mediating software. Combine *gesture classes*, *proximal interaction* and *interaction models* to provide a higher resolution of system events. Use conceptual tools to simplify complex interactions into manageable system events. Use software groups to structure and distribute processes within a shared context. *Acquisition, Analysis, Interaction, Composition, Diffusion.*

3.6 Summary (chapter 3)

A general background of typical approaches to prototyping novel controllers was discussed. An experiment was presented to show how interaction and response can be mediated using analysis of variance to identify new behaviours. In Psychology the ANOVA or Analysis of Variance experiment design is used to test specific dependent and independent variables in a controlled context. In this context, this proven method was developed to identify statistically significant data to reveal new patterns of behaviour or response (Pilot Study).

A summary of the three design elements: interface, environment and software, conceived as an integrated compositional framework, was presented. The importance of including interaction layers and *gesture classes* within novel interface design for mediating collaborative exchange and motivating emergent behaviour. The overall Orb3 Collaborative environment promotes compositional strategies identified in chapter 2: Adaptive Composition. This approach combines the features of a collaborative tangible interface for novices with key qualities normally only found in advanced customised instrument designs. A series of design strategies were presented to develop a critical approach to new digital instrument design. This demonstrates the principle that instrument design should be approached as creative act (Merlier, 2004) but that underpinning design strategies are required to go beyond the conventional interaction models found in many so called Novel controllers or Tangible interfaces. Participant interaction and behaviours were considered as musical parameters, just as virtual communities of interacting agents form musical processes with limited individual ‘experience’ learning through interaction and exchange, the Orb3 interface design approach adopts these principles to establish collaborative behaviours encoded as gestures and translated as symbols that can be learned, shared and transformed.

Seven design strategies were presented, condensing the principles developed by the author to design and describe an Adaptive Social Composition integrated system design. These strategies were used to develop the Orb3 interface, and demonstrate the core findings of this research in a condensed form that can be applied to the design of new digital instruments that combine social groups, novel interfaces and mediated sound environments

In the next chapter, the Orb3 interface design is presented as a proof of concept. This new digital instrument was developed using the design strategies identified in this chapter. Technical details are provided, including choice and model of sensors. Design decisions and rationale for ergonomics are presented. The technical specification, physical design and printed circuit boards are presented. Core features are explained. This new digital instrument is presented as a Proof of Concept.

Chapter 4

Proof of Concept

The Orb3 *interface design* considers the ergonomics of hand held musical interfaces, integrating sensors that transform typical control actions and gestures into identifiable behaviours. The Orb prototype weighs 400g including battery pack, the ball is 120mm in diameter with components housed in an opaque polycarbonate ball. The centre of weight is lowered so that the ball is stable whilst being held. The spherical nature of each unit does not reveal conventional instrument controls, strings to pluck, keys to press, holes to cover etc. However the shape weight and dimensions are designed to motivate a cupped, palms up holding position, a common gesture of giving or receiving which underpins the learning processes inherent in an adaptive social system. The design allows for small intimate gestures between palms such as rolling, shaking, tilting to be identified in real-time. These movements generated by wrist actions are considered to be *local* gestures. More dramatic movement is possible with arm movement, or full body articulation, for example, swooping motions or linear accelerations where height and orientation changes are dramatic. The design includes sensors for both subtle and dynamic movement. Large-scale gestures are considered to be *global*. In this context, *local* gestures are measured only by onboard sensors, while *global* gestures are mediated by combining sensor data with overhead video tracking of each Orb unit's relative position. The term *global* in this context denotes actions or gestures that are visibly expressive and can influence or shape the collaborative sound environment as co-created symbols or patterns that are identified to produce new sound material. Global actions go beyond a direct manipulation or control model of interaction. The *local* or intimate gestures affect the Orbs related sound object only. This ergonomic strategy is intended to shape the behaviour of participants while holding an interface. Just as different types of move in a board game invite a different response from an opponent, different types of gesture can invite or motivate a range of response. And similar to a board game, which has a rule set to group actions and possibilities and focus player behaviour, classes of gestures and interaction models are embedded in the mediating software. Within a gestural vocabulary the terms *direct* and *indirect* gesture acquisition are used to describe different classes of possible gestures. Within an adaptive interface both sensing modes are facilitated or recognized, the rationale for this is to extend the interface beyond literal parameter mapping of a controller, to the expressive potential of a higher-level instrument. Both the direct actions of participants and the proximal relationship between interfaces are sensed or tracked to support both direct and indirect gestural acquisition within a shared compositional context. A single digital video camera, connected via firewire to the mediating software is included to facilitate this. The camera is positioned centrally 3 meters overhead so a top down view of the 3-meter diameter floor area between speakers can be monitored. A wide range of video tracking method can be implemented within Max/MSP/Jitter. The Computer Music community has developed many additional externals that have sophisticated tracking and analysis built in to optimise the process of software design, leaving the designer or composer to concentrate on interaction and sonification issues. (Taptools by Timothy Place and CV.jit by Jean Marc Pelletier were introduced earlier (Website reference 11 and 12). These externals are combined with Max/MSP and are used to monitor global activities via a live overhead video feed. So, the concept behind the interface design is to provide an adaptive interface within an integrating software framework with an evolving rule-set. This rule-set can be visualized as the playing board of a conventional board game, where a range of known and anticipated moves are possible but an individual collaborator can interpret the context of these 'rules' to develop playing strategies. In this design however, the playing board is replaced by the mediating system of software, sound objects, and diffusion environment. As such, the playing arena and potential actions are experienced but not directly visible. Instead of physical counters playing pieces are replaced by the autonomous mobile interfaces or Orb interfaces and their users. The structured moves of playing pieces are replaced with different classes of gesture, mediated by each interface. The intention was to design a compositional framework that is accessible by novices but challenging for musicians and composers. Consider the board game chess, the novice may quickly learn some simple moves and begin to group them to form simple strategies, the advanced player may hold a complex sequence of moves, and associated responses in their head. These potential extended actions and responses are based on prior experience, previous games and opponents and a deeper understanding of the underpinning structure of the game arena, the apparently simple grid of 64 black and white squares. During this research into novel controllers and new digital instruments, different approaches were explored. The underpinning rules applied within these related designs could be better understood in terms of *interaction models* (chapter 1), *event groups,* and *motion behaviours* (chapter 3). In order to develop a range of strategies for adaptive social composition it was necessary to identify field specific language for evaluating the different characteristics and properties of these new devices (chapter 2). These principles have been combined to integrate *interaction models* (control, sequential, organizational, relational, transformative) with the concept of software mediated *event groups* (*Data analysis, Social Interaction, Adaptive composition and Ambisonic diffusion*) and *motion behaviours* (lead, follow, avoid) as an integrated design strategy for collaborative sound environments. These elements have been identified to develop new digital instruments that include a compositional framework and were used to conceive the Orb3 system design as a case study for Adaptive Social Composition.

A significant feature of this thesis is that it presents key terms to understand the context of interaction and behaviour with novel music interfaces or sound environments. To substantiate this, a broader evaluative vocabulary was collated from the available literature, responding to research questions presented to the field from Machover, Weinberg and Paine. In simple terms, those individuals recognized as pioneers in the field of novel instruments, sound environments or expressive interfaces have repeatedly questioned the value of these new developments, asserting that there is no evaluative language that adequately covers the diverse range of approaches. Furthermore, although well resolved technically, many of the most widely known or well documented examples follow either a *control model* of interaction with literal parameter mapping, or an instrument model with highly personalized and obscure features – a bespoke solo instrument. Those interfaces designed for group interaction tend to fall into the former category and are generally described as being designed for novices, with features reminiscent of simple percussion instruments. The Pilot study, presented later in this chapter was developed to establish a methodology for designing and testing system features for Adaptive Social Composition. This provided an effective way of providing simple tasks for test subjects to interact with diffused sound objects. The results of this process establish principles for designing the synthesis and diffusion elements of a collaborative sound environment with adaptive features.

Within the field of installation art or sound environments for public interaction two extremes of design can be seen, a black box approach where participants do not appreciate the tracking and control systems that are in place, or oversimplified grid based systems with limited *explore-ability*. The tangible interface community have developed a number of tabletop systems to encourage social interaction. In many cases, the handheld objects are reminiscent of board game pieces and encourage a range of control behaviours with a degree of exploration, some use symbol tracking to map different instances of the same object to different effects or synthesis parameters. Some of these tangible interfaces are simple implementations of grid based tracking or depend on limited symbol recognition unrelated to the actual shape or texture of the objects used. Other developments in handheld tangible interfaces exhibit a local control model, triggering, sampling, agitating, again with limited *explore-ability* and are thus easily learnt with reduced compositional flexibility and minimal expressive depth. The significant difference in the Orb3 interface design is that it combines interaction models from different domains. This addresses key concerns in the field by enabling extended exploration of a collaborative compositional framework that is intended to be *nourishing* (Machover, 2002) through a series of *continuous controls* that are designed to *stimulate rather than placate* (Machover, 2002). The compositional structure is designed around *analysis and interpretation of movement data* (Winkler, 1998) to extend immersion through layered interaction models or behaviours that are mediated by collaborative conscious and subconscious interactions. In terms of an aesthetic, the physical interface design allows users to see the internal components implying both the complexity and sensitive nature of the object. Imagine a simple snow globe, it illustrates a simple scene, however, when the globe is shaken, the literal illustration disappears and the contained environment is transformed by our physical influence over it.

4.1 Orb interface sensors

An array of 6 digital tilt switches circle the Orb on the outer edge, measuring direct angular rotation off the vertical axis, a simple tilting motion echoes the thumb depress on a key of the Mbira, triggering a sound event within the currently playing sound object. Each interface features a dual-axis accelerometer combined with tilt sensors to capture gestural movements. Dual axis accelerometers are used to capture movement velocity and influence synthesis parameters. A digital compass is used to map relative position and orientation of each interface to mediate sound-object diffusion trajectory. Ultrasound sensing has been tested for close proximity detection of other mobile interfaces, and has also been specified for measuring interface height from floor for direct pitch mapping in relation to Orb height. Bead thermistors are used to monitor ambient temperature and palm surface contact temperature when handheld. Skin temperature is used as a subconscious factor producing relatively constant variable that the user is not aware of, or able to easily control. LDR's (Light Dependent Resistors) are used to monitor ambient light levels, providing base parameters for sound object synthesis for each Orb. Basic robotics are used to continuously reposition the interface when in the autonomous mode '*absorb*', coordinated through the overhead camera tracking system using the three motion behaviours: *lead, follow, avoid.* The resulting motion creates identifiable patterns that are mapped to compositional variables. For example, a distinct sound object is generated for each Orb and the musical interplay between them forms the composition, whilst revealing some of the evolving rules to motivate participant interaction. Simple biofeedback to measure a participant's pulse is used to influence temporal characteristics of the related sound object, generated by the Orb when picked up and held by a participant. The combined data from groups of sensors can be mapped to identify known gestures and triggers corresponding events. Combined data patterns can also be monitored to identify repeated new gestures to evolve new behaviours.

Each Orb interface is designed to use the following sensors:

- Tilt switches (angle)
- Dual Axis Accelerometer (pitch and yaw)
- Ultrasound (distance)
- Digital Compass (orientation)
- Pressure (air movement)
- Bead Thermistors (temperature change)
- Light Dependent Resistors (ambient light)
- Biofeedback (pulse & galvanic skin response)

Each interface or Orb has two states: *Absorb* and *Adapt*. In *Absorb* mode the interface is autonomous. A pair of motorized wheels in the base is used to propel the Orb around the floor area while monitoring global data and interacting with other interfaces. Three motion behaviours; *chase follow and avoid* provide a behavioural framework for this state. The *Adapt* mode is activated when an interface is picked up, local movement or *direct* gestures shape the Orbs' existing sound object while the skin temperature of the participant is detected. Orb orientation and direction initially controls sound placement, although these parameter controls are remapped by the mediating software when certain behaviours are detected, extending *learn-ability* and motivating alternate interactions. (*Learnability* is a term used by Jorda (2003) in relation to a novel interface for musical expression. *Learn-ability* can be understood as a quantifiable measure of the extent to which musically significant gestures can be learned and performed successfully with a new interface.

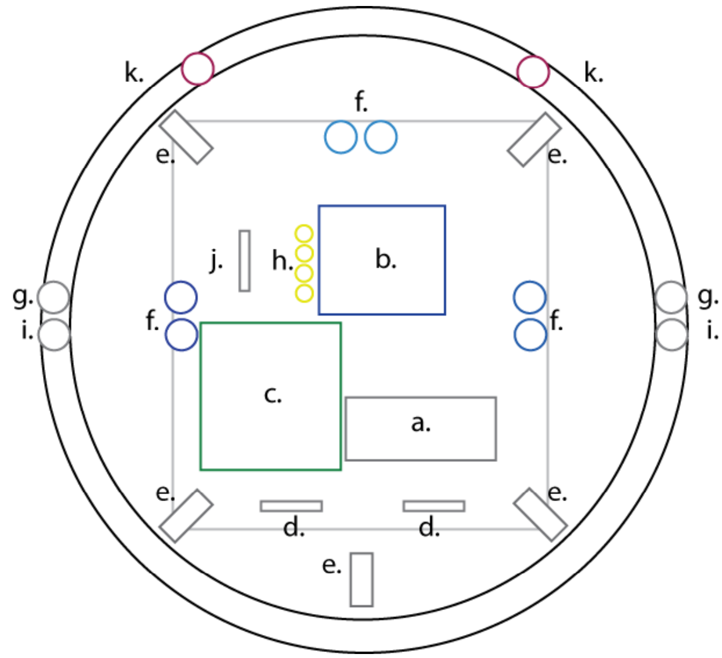
<u>Parameter Mapping</u>

Sensing parameters are mapped to physical motion and presence attributes, the specific mapping of the output of these integrated behaviours to synthesis or sampling is not specified in the interface design and is discussed later in this chapter: Software Design.

In *Absorb* mode the interface sensors are calibrated for global data collection (increased sensitivity) the variables used to generate an individual identity or sound object and relative diffusion controls for each Orb:

• Tilt sensors (mode switching between *absorb/adapt*)
• Dual Axis Accelerometer (linear motion and velocity – sound synthesis)
• Ultrasound (not used)
• Digital Compass (directional orientation - diffusion)
• Bead Thermistors (Ambient temperature - synthesis variables)
• Light sensors (ambient light - synthesis variables)
• Biofeedback (not in use)

The following diagram figures 28 and 29 show the main sensor positions as located on the custom designed PCB. The position of the XBEE transceiver and programmable PicAXE40XE chip is also shown. Sensor positioning is decided by identifying potential participant gestures. The ergonomic design of the interface is also a key factor.



(fig 28 sensors and core components)

a) PicAxe40XE chip (programmable, formats data, motor behaviours)

b) XBEE transceiver (wireless communications to computer)

c) Digital compass (orientation, relative angle)

d) Mounting for dual axis accelerometer (motion)

e) Non-mercury tilt switches (rotation)

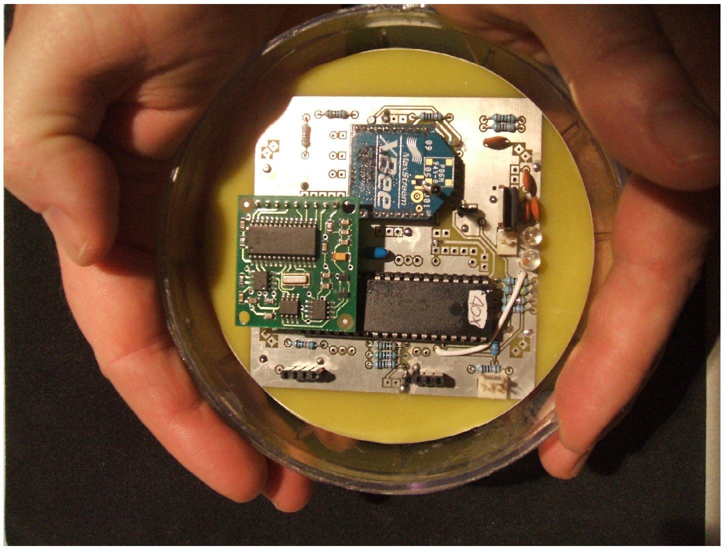
f) Light emitting diodes (aids video tracking/identifies mode)

g) Light emitting diode and light dependent resistor (pulse detection)

h) Status light emitting diodes (communications)

i) Contact for skin conductance (used for GSR)

k) Bead thermistor (temperature sensing)



(fig. 29 Orb prototype showing PCB)

In *Adapt* mode, the interface sensors are used to measure local parameters and interface orientation:

- Tilt sensors (direct triggering gestures)
- Dual Axis Accelerometer (detects *direct* gestures x, y, motion)
- Ultrasound (Orb distance z from floor while held – i.e. pitch)
- Digital Compass (participant direction)
- Bead Thermistors (skin temperature)
- Light sensors (ambient light and direct manipulation)
- Biofeedback (pulse – temporal characteristics, GSR – influences equalization of sound object (subconscious))

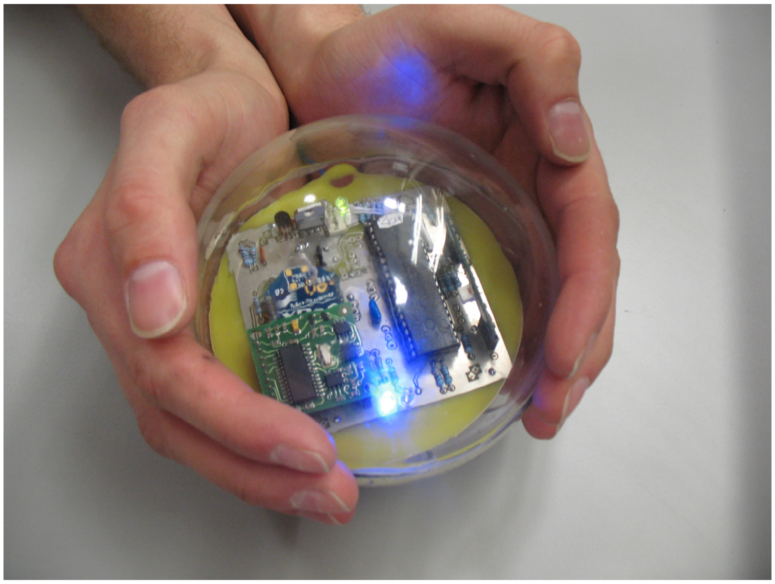
The combined parameters generate a distinct sound object for each Orb interface; the collected variables are broadcast wirelessly to computer where a separate synthesis module is provided for each Orb. These parameters can be mapped to different properties and characteristics, to provide a distinct framework for sound object evolution for each orb. This allows different composers to experiment with the system and apply their own musical context to the behaviours and gestures the system mediates.

In *Adapt* mode the collected variables can be assigned to different roles when received by the software allowing for different compositional approaches to be explored, the default settings follow a tangible set of parameter mappings, so the player can perceive the consequence of a gesture or action. This allows for identification of combined data patterns to reveal a tangible gesture. For example, to create and detect a swooping motion, the relative direction (digital compass) height variance (ultrasound) and acceleration (accelerometer) data are combined. The pattern that these three variables create can be recognized and stored in software and mapped to different interaction models. This principle allows new gestures to be designed and integrated and mapped to compositional events, following the principle of an evolving rule set.
Sound design within the NIME community (New Interfaces for Musical Expression) tends to explore percussive structures, with parameter control over dynamic synthesis. There is also a range of sub genres in reaction to mainstream Computer Music software which reinforces popular linear loop and sequence based composition models; Propellerhead's *Reason* (website reference 51), Steinberg's

Cubase & Nuendo (website reference 52), Apple's *Logic* (website reference 53), etc. where assembly of parts and layers of loops and presets are well understood in the public domain. The deconstruction of commercial interfaces and modification of circuits (circuit bending) and live programming all place the compositional emphasis back on a live process. The Orb3 system design presents a new compositional framework where each Orb creates an individual sound object from *global* data mediated by *local* interaction. This embeds the compositional approach in the way each interface element functions and reinforces a learning process that is not prescribed by visual interface elements or conventional music paradigms. Many multi user interfaces claim to be collaborative, but interaction that mirror commercial software models. Extended and custom instruments offer new expressive potential for individual performers to learn based on prior knowledge of the instrument from which the new hybrid is derived (*instrument model*). The *explore-ability* delivered by these systems offers access to a new musical domain for established performers. Within the Orb3 design the motivation to explore gesture, movement or spatial relationships through proximity is provided by audio feedback. This shifts the perception of the participant from a control model to an exploratory model of interaction, which is rewarded by matching *direct gestural acquisition* to synthesis and diffusion. So, a design strategy for this adaptive social composition framework was to integrate the tangible interaction principles of a table-based collaborative 'instrument' within a larger mediated system by treating the physical interfaces and participants as system elements. However, unlike the tangible objects in other systems the Orb interfaces are not passive objects or symbols. They integrate *direct* and *indirect gestural acquisition* normally found on bespoke hybrid instruments. As such, it is a progressive development adapting significant features into one system. Autonomous and evolutionary music systems are historically software-based and interacted with through a screen-based interface. Current developments in this field integrate evolutionary and behavioural approaches with physical environments. Documented approaches include systems that integrate human and robot players in a collaborative framework for live performance (Singer et al., 2005). The Orb3 design (figure 30) includes autonomous modes for interface behaviours and integrates human actions as compositional material, within an evolving discipline, this is a relatively new development and further research is needed to refine these processes.

4.2 Orb Gesture Classes

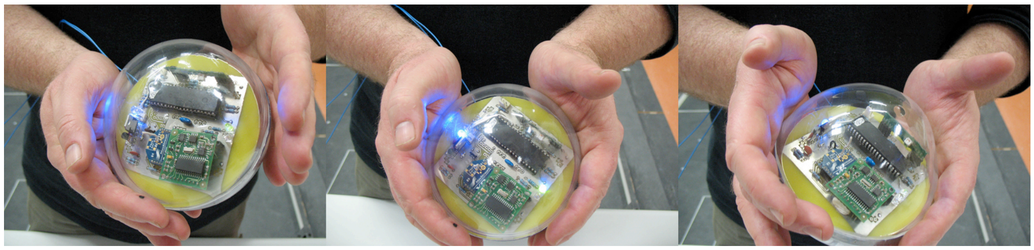
Gesture classes and interaction models have been integrated to develop the Orb interface design, combining sensor data and user actions in context. This accomodates a repertoire of actions intended to motivate a novice and to engage a musician, developing the priciple easy to learn, difficult to master. Gesture design includes *direct* and *indirect* data acquisition to provide a higher level interpretation of possible actions in context. This is achieved by considering the range of physical actions possible when holding an Orb interface, and positioning sensors to identify these actions as



(fig. 30 Orb interface in *adapt* mode)

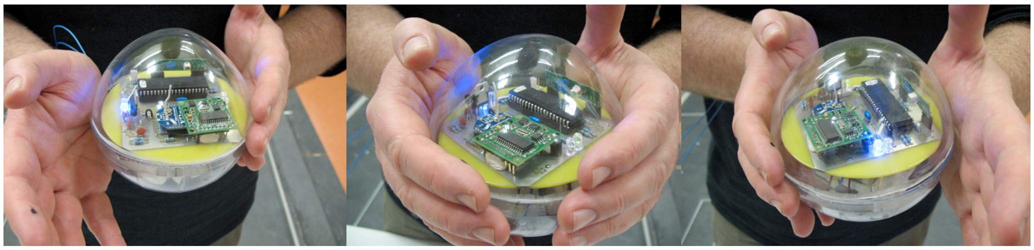
gestures. These gestures fall into different groups supporting a range of interaction models, beyond the commonly used intuitive *control* actions, through *relational* gestures to *transformative* events. The principle of an evolving rule-set is applied to structure different interaction models. This is significant as different levels of user interaction can be identified, data can be treated in relation to group actions with different user roles supported.

Gestures are identified by monitoring sensor data. Different combinations of actions generate different patterns of data. By observing user actions during testing and recording data patterns, principles for identifying different classes of gesture were defined. Each sensor was selected to identify a different characteristic. Tilt sensors are simple switches which activate when a pre-set angle,



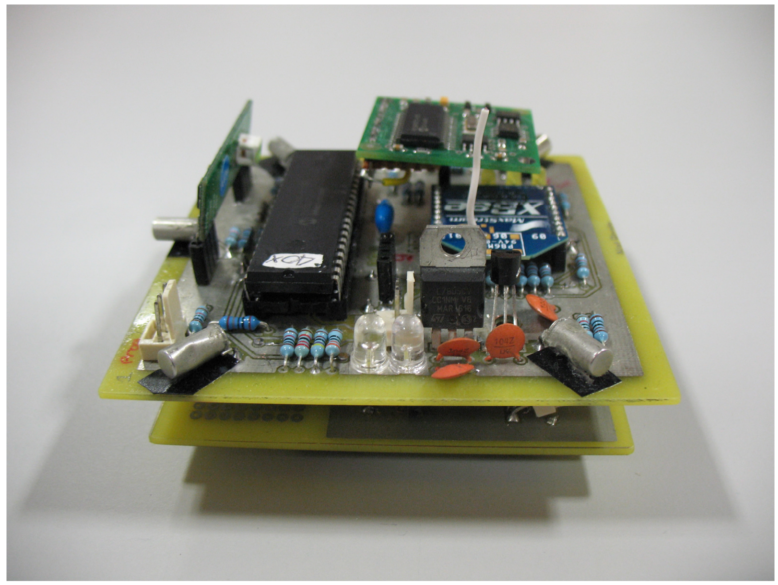
(fig. 31 Orb prototype rolling/tilting gesture)

normally 30 degrees from horizontal, is reached. By combining the on or off sequence of sensors placed around the outside circumference of an Orb interface simple rolling patterns can be identified. In figure 31 simple rolling actions activate the tilt sensors in sequence. The sensors themselves simply indicate wether an Orb has been tilted and in which relative direction, but by understanding the ergonmics of the interface shape and the actions that are possible when it is held, a higher resolution gesture can be identified in context. This principle is followed through to the more complex sensing that is part of the interface design; combine sensor data to identify known patterns that represent different user interactions or behaviours. Basic *control* gestures map sensor data directly to a logical parameter such as *pitch*, *velocity* and *duration* of a note or *panning*, *volume* and *equalisation* of a sound. Higher level gestures can be identified by increasing the resolution of an action adding a compositional context. This is acheived by storing an action and the resulting numerical data as a list using software created in MaxMSP. These lists are grouped into known gestures or compositional events, allowing comparison of a current action with a



(fig. 32 Orb interface lateral rotation gesture)

known gesture or event using simple pattern recognition (this is described in chapter 3 Mediating Software). This approach allows new gestures to be added and new contexts for interactions to be adapted, either by a composer prior to group interaction or during social interaction by repetition of a new combination, for example one user mimics the gesture of another demonstrating *sociability* by learning and exchanging repertoire in figure 32 a lateral rotation gesture is captured by the digital compass. The relative facing position of the Orb is represented as a variable. The interface design had to accomodate these potential behaviours, this was a key factor in the selection of sensors and the associated actions. Data from each Orb interface is sent via serial packets using an XBEE transciever, mode switching and motor control data can be sent direct to each Orb. Data is recieved on the host computer via Universal Serial Bus or USB using an additional XBEE transciever. A programmable PicAXE 4TXE chip is used on the main circuit board of each Orb to organise collected data. A program is downloaded to each Orb providing autonomous motor controls, which are designed to exhibit *lead*, *chase* and *avoid* behaviours.



(fig. 33 Stacked PCB prototype with sensor integration)

The PicAXE 4TXE also handles mode switching, isolating sensors that are not used in a given mode and controlling the onboard LED's (light emitting diodes) to indicate current status. In figure 33 a prototype with motor control and on-board behaviours is shown. Sockets for an accelerometer, digital compass, PicAxe40XE chip and XBEE transceiver are provided on the main PCB. LED's to display status and provide navigation lights to aid video tracking can also be seen. Non-mercury tilt switches are mounted on each corner. The square format prototype shown in figure 33 integrates all the following sensors on two stacked PCB's

Sensors included in the Orb interface design:

- *Non-mercury tilt switches*: activate when oriented above 30 degrees from horizontal

  Sited on the outside edge of the PCB to detect relative tilting actions away or towards the user to left or right when interface is held. Also used to identify rolling patterns.

- *Light dependent resistor* or *LDR*: produces a variable signal between 0 and 5volts in response to light.

  Used to switch off autonomous motor behaviours when Orb interface lifted from ground, activated byincrease in ambient light.
  Used to measure light variation in fingertip using blood flow method to identify user pulse or heartrate. Parameter mapped to timing events such as modulation of current sound object

- *12k bead thermistor:* Produces variable in response to temperature.

  Used to measure ambient temperature when Orb interface is autonomous. Used to measure user palm temperature when Orb is held. Variable mapped to sound object characteristic such as brightness, a lower temperature reduces the higher frequencies of a sound, a higher temperature increases them. Principle of improved clarity of definition of sound objects when a user is interacting by holding an interface.

- *Digital compass:* Produces a variable in response to magnetic north, also vulnerable to magnetic field of metal building structures and unshielded speaker systems. Requires calibration.

  Used to measure interface orientation or rotation in horizontal axis. Also used to detect proximal relationships through facing other Orbs or identified participants. Relative angle or orientation of each Orb is combined with Orb relative location as detected by overhead camera using simple motion tracking.

- *Dual axis accelerometer:* Produces two variables in response to linear motion or acceleration in horizontal axis.

  Used to measure direct actions on Orb interface. Values collected measure acceleration and decceleration over time

- *Ultrasound:* Produces a variable measuring distance to an object or surface with range up to 3 meters.

  Used to map Orb height from floor mapped to sound object *pitch* in early prototypes.

Digital Compass: relative horizontal orientation of interface as a single variable Z.

Direct Acquisition: used to identify simple rolling actions between facing palms

Dual Axis Accelerometer: XY motion in horizontal plane, two variables generated by linear motion.

Gesture classes:

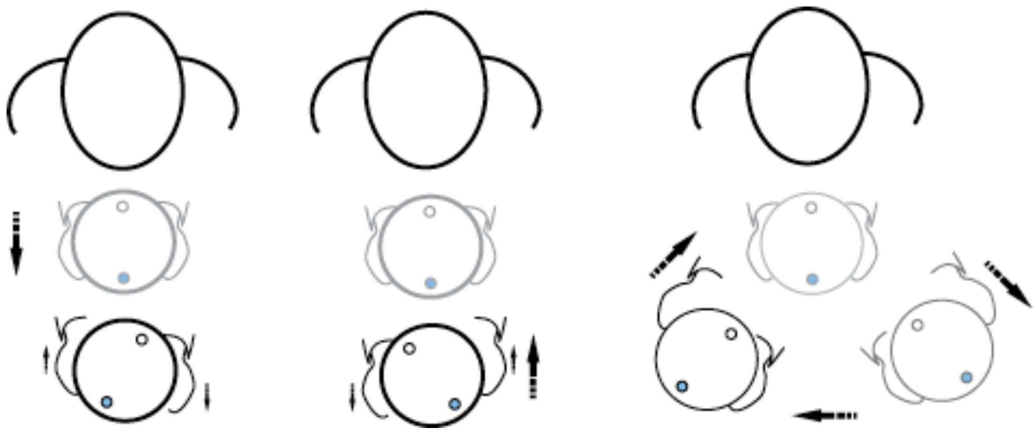
Direct acquisition: used to identify agitation, shaking, and linear push or pull actions.

A *control* gesture such as shaking filters the current sound object, a *relational* gesture such as linear motion changes filter parameters of current sound object in context (default choice of filter/parameter range for each gesture)

Parameters are scaled and mapped directly to events.

Indirect acquisition: User can describe symbols*, angle, triangle, square*

Indirect Acquisition supports *transformative* model of interaction.
Pattern analysis is used to identify basic symbols by identifying acceleration deceleration sequences of to or more linear action and combining orientation data from digital compass.



(fig. 34 Additional Orb gestures diagram)

Linear actions include push away at one speed, change of angle, and pull back at alternate speed. For example, a user can describe a triangle symbol by moving the interface in a linear acceleration to their left, as the interface is pulled back from the end point of this action a numerical change in the accelerometer data is evident. As the user moves the interface to the right, the compass orientation changes, the value increases, this pattern of one value accelerating then stopping followed by a gentle increase in detected angle allows a higher resolution of gesture to be detected. If the action is repeated by continued increase in angle followed by deceleration as the interface is moved across the users body to the right, the second side of a triangle is completed. The third side is created as the user pulls the interface back to the start position decelerating to a stop with the compass variable within range of the original value. This method allows angle, triangles or squares actions to be identified by combining sensor data in context. If such a gesture is identified a *transformative* event is triggered, for example changing the envelope of the current sound object.

4.3 Chapter Summary

In this chapter the design and development process for an adaptive novel interface was presented. Technical details were given to provide an overview of key design decisions showing position and application of sensors. Example gestures were discussed and the significance of design for collaboration through a shared repertoire that can be shared and developed during performance was explained. The significance of the design is that it is designed to motivate collaboration through a mediated process of *co-creation*. It follows the principle *easy to learn* by using ergonomic design to engage novices. It also demonstrates the principle *difficult to master,* by including mode switching, dynamic parameter mapping and collaborative behaviours. It demonstrates that new digital instruments can be designed for a range of participants, by following the design strategies presented in chapter 3. The core features of *explore-ability*, *user role flexibility* and *instrument resolution* have been delivered in a small wireless package with intuitive interaction. Individual, collaborative, and autonomous gestures are incorporated into the design. Interaction models are supported. Participants are enabled to *play, intercept* and *contribute*. *Proximal interactions* are supported to motivate *sociability*. The principle of an evolving rule set is integrated through mode switching and mediated behaviours.

Chapter 5

Conclusions

The development of complex interactive systems that integrate multi user interfaces is an interdisciplinary enterprise. The previous work in novel controllers or instrument design (Tanaka, 2000) enters a new dimension when multi-user interfaces (Jorda, 2005) are developed for group interaction. Individual practitioners who have initiated such projects have had to undertake significant research that has contributed to the field of Computer Music where these approaches are formally presented and defended through peer review. Often the individuals concerned initiate project teams to develop an initial prototype establishing an applied framework for collaborative research. Musicology, human computer interaction, interface design, electronic and mechanical engineering, computer mediated composition and performance and sonic arts practice are all disciplines evident within these groups. This illustrates that an iterative approach is required, developing and refining each element of such systems with publication of stages of development in related conferences and Journals. Equally, these integrated works have contributed to the development of interactive and performing arts practice with new knowledge and techniques enabling creators to achieve more embodied interaction between people and computers. These works are regularly exhibited on the international circuit: Ars Electronica, Siggraph, ISEA and so on. Dedicated Centres have also been established that offer specialist research opportunities, developing new resources and documenting academic, technical and creative outputs, many of the researchers referenced in this thesis are currently involved with these centres or have completed a scholarship, contributed to or led new research in related fields.

The Orb3 system design was conceived as a framework to develop the design strategies that have been resolved through this research. As such, it has allowed a new approach to multi-user interfaces to be articulated, referencing major developments in this area. Similarly, the performance technologies for tracking and sensing diverse interactions have been evaluated through direct observation of and interaction with related systems. Current developments in hybrid instruments and expressive controllers have also been investigated first hand at ICMC and NIME through conference performances, workshop demonstrations and open improvisation sessions. Original speculative development revealed new potential for integrating behaviour as an identifiable parameter, informed by the observation of participant actions with a wide range of new interfaces for musical expression. Developments in the field of evolutionary composition reinforce the value of research that reveals new understanding of our perceptual system, and how biologically inspired systems can demonstrate new musical behaviours. Adaptive social composition offers further opportunities for research within mediated frameworks that identify and motivate new collaborative behaviours. Interaction design is not limited to graphical interfaces and physical controllers. Multi-user systems have challenged our preconceptions of what collaboration is and how virtual entities can co-create significant compositional material that mirrors complex human behaviours. Artificial intelligence, emergent behaviour and neural networks have each inspired new compositional approaches, as have adaptive software-based systems and complex modeling of biological processes. The Orb3 system design integrates some of these approaches, drawing on established techniques or concepts to inform the overall interaction design within a mediated system through an iterative process of prototyping. Each system element can be further improved; robotics, electronics, tracking and gesture capture, bio-feedback, synthesis and diffusion are all technical elements that can be extended and offer opportunities for further research. The lack of *common standards* (Paradiso, 2002) in this field makes evaluation of specific interface approaches hard to resolve. Although new technical standards emerge to facilitate robust communication between interactive systems, there is a perceived lack of field-specific vocabulary for the non-technical elements of system design. However through this research a number of valuable strategies have been developed, by considering the core issues, related to the field. The established *design principles* (Cook, P. 2001) the core concepts of *efficiency* and *learn-ability* (Jorda, 2004) the significance of expressive range in terms of *instrument resolution* and *expressive depth* (Settel and Lippe 2003) and extending *performative engagement* (Paine, G 2002) through *exploratory* (Wanderly, 2001) behaviour. In addition, the transformative strategies identifying conscious and subconscious interaction through system modes *Absorb* and *Adapt* allow new collaborative musical behaviours to be integrated. The terms *Absorb* and *Adapt* have been used to provide a system framework that motivates new *compositional models* (Lippe

1996) that apply transformations to interaction behaviours to motivate new interaction within a shared compositional framework. The strategy of grouping participant actions and behaviours into *interaction models* has been useful for comparing different tangible interfaces. And more significantly for this strategy has been extended to incorporate event groups for dynamic parameter mapping within a shared compositional context. This new approach integrates software mediation to dynamically remap sensor data or identified gestures to *local* and *global* compositional contexts. A set of core features was established in chapter two: Adaptive Social Composition. Forms of identifiable emergent behaviour, both in the system and evident in the actions of participants offer the potential to increase human computer interaction, to adapt learning processes to cybrid environments and to create new musical interactions. The previous research (Mack and Rock, 1998) into *"innatentional blindness"* may prove significant in relation to designing new transformative behaviours that mediate between visual and auditory interactions using encoded distraction, offering further research opportunities for Adaptive Social Composition systems.

In Chapter one, the lack of vocabulary for evaluating or designing new digital instruments was identified and the need for a field-specific vocabulary was established. At the outset of the research presented in this thesis, a number of terms were collected and explained in context. These terms have been grouped into *qualities*, *models*, *interactions*, *behaviours* and *system features* to establish a design framework for collaborative musical interfaces and novel controllers. Field specific terms representative of the range of properties considered desirable in these new systems have been identified in the context of reviewing these related works. A definition and grouping of these terms was included in chapter 2. The primary problem in trying to establish common standards is that these new instruments are often individual or groups of interfaces integrated within software-mediated frameworks, so a table listing technical features does not provide an objective measurement of the qualitative contribution of an individual instrument or system. Nor does it deal with any design elements that are frequently embedded as rules or structures within electronic circuits or coded in software that affects player interaction and behaviour. However, by studying the available literature and reviewing such systems first hand it is possible to identify complimentary features from the diverse range of approaches and produce a coherent framework for evaluating these new and emerging systems with established field specific terms. It is also possible to use such a framework in an educational context, motivating the fast prototyping of effective novel controllers and hybrid instruments. In a research context, this collection of terms grouped into identifiable categories, can be used effectively to develop new digital instruments that combine software, hardware and human machine interaction.

5.1 Contributions to Knowledge

The first contribution to knowledge of this thesis is to introduce a glossary of field specific language; identifying and extending this field-specific vocabulary using established and recently published terms. This framework is presented below with a table (appendix 1) showing core features of related representative works (discussed in chapter 1.). A scale of 1 to 5 is used to indicate the balance of features. It is recommended that the date of comparison is included as many of these systems are developed iteratively. A feature missing or scoring low may be in development for the next version or revision of the system. The intended context or *end user* is highly significant so this category is also included. The framework can be used either for comparing features and qualities of current examples, or it can be used in the design process for new digital instruments or related systems.

An example table showing features from Overholt's *Overtone* Violin (Instrument Model), Jorda's *Reactable* (Collaborative Tangible Model), Weinbergs' *Beatbugs* (Sequential Model) and the Orb3 design (Adaptive controller/environment hybrid) is presented in appendix 1. (The reader may wish to refer to the definition of terms provided at the end of chapter 1)

This first contribution addresses the initial research question posed:

Can a musical interface be designed that engages the novice but has the expressive qualities and personalisation of a traditional instrument?

An evaluative framework is provided for comparative analysis of existing systems while providing the additional conceptual tools for designing new musical interactions between new digital instruments and participants/performers. An adaptive interface prototype titled 'Orb3' was designed and implemented (chapter 4) using the principles embedded in the collated terms. The Interface can function autonomously with structured movement behaviours, it can be held and used as a music controller with direct mapping of orientation to sound diffusion but it also features a range of sensors for sensing expressive gestures and biofeedback.

The second contribution was to establish and document a design process that incorporates statistical data analysis of interaction events within a 'B format' ambisonic diffusion environment using software written in Max/MSP. It addresses the second research question:

Can the principles found in turn based board games be used to develop social composition frameworks that are intuitive to use in a collaborative musical context?

The principles of turn taking, gamesmanship and evolving rule-sets observed in traditional board games were adapted to provide simple challenges for subjects to interact with discrete sound objects. This approach uses the established method of Analysis of Variance or ANOVA (King & Minium 2003) with an experiment design based on two dependent and two independent variables. This methodology integrates perception and interaction tests to identify subject behaviours in response to diffused sound material. The software developed to support the Pilot Study is designed for further research by supporting interaction and perception experiments within an ambisonic environment. It supports the following features:

1) Sound event triggering for a b-format ambisonic diffusion environment, using samples or musical sequences.
2) Interface monitoring and data collection for multiple users is included allowing new experiments to be designed around the core processes.
3) Each session is fully automated and repeatable for multiple subjects. Data is filtered into lists, exported as text files of dependent and independent variables for subsequent analysis using ANOVA.

A number of experiments have been designed, one of which was run under controlled conditions as a pilot study (presented in chapter 3) to prototype core elements of an Adaptive Social Composition system-design. The details of this pilot study using ANOVA (analysis of variance) to identify patterns in user perception through interaction and response described in chapter 3.1 Supporting data for the Pilot Study is fully documented in appendix 2.

The third and perhaps most challenging question in terms of establishing effective strategies for system design was:

How would one describe such a system, what qualities and characteristics would it manifest and how would one design it.

The term Adaptive Social Composition was proposed to describe an integrated design process where the sound environment, interface, and participant behaviours are combined within a shared compositional framework, mediated by software. This thesis presents design strategies for Adaptive Social Composition. These explain how dynamic parameter mapping of novel interfaces can identify proximal relationships in groups of participants within mediated spaces for collaborative composition. A prototype system design integrating, *Data analysis, Social Interaction, Adaptive composition and Ambisonic diffusion* entitled Orb3 was presented (presented in chapter three). The core design strategies established for this Adaptive Social Composition system were applied to develop a new adaptive interface design. A technical overview of this novel interface design was presented in chapter 4: Proof of Concept.

Overall, this integrated approach combines novel adaptive interfaces within a collaborative context, described as Adaptive Social Composition. It follows the model of an evolving rule-set as found in 'Nine Men's Morris' (as discussed in chapter 2.2) to motivate and engage novices. The strategic complexity of 'Chess' is embedded in a structure of known constraints, providing extension and challenge to experienced players. The mediating software design presents *interaction models*, *event groups* and *participant behaviours* instead of different moves from a number of clearly identified playing pieces, but the principle of *easy to learn, difficult to master* is carried forward as a strategy for collaborative mediated interaction within a sound environment. A novel interface with its own set of movement behaviours replaces the simple representational objects used in board games or Tangible interfaces for novices, with a sophisticated digital instrument. The playing board and rules used to score the sequence of actions in a game of draughts, just as the score presented to an orchestra mediates the actions, roles and gestures of performers.

Stages of development were published in independently peer reviewed international conference publications; these papers are included in appendix 3.

References

NB: Some references refer to papers published in online proceedings where page numbers are not provided.

ALSOP, R. 2003. Compositional Processes in Developing Poly-Media Performance Works. *In: Proceedings of the Australasian Computer Music Conference, Perth 2003.* Australasian Computer Music Association: Fitzroy, pp.5-12

ARFIB, D. VERFAILLE, V. 2003. Driving Pitch Shifting and Time Scaling Algorithms with Adaptive Gestural Techniques. *In: Proceedings of the 6th Int. Conference on Digital Audio Effects (DAFx-03), 8/12 September 2003 London, UK*, pp.1-6

BANDT, R. 2004. Sounding Public Space, Sound Artists in the Public Domain *In: Proceedings of the International Conference on Auditory Display, 6/9 July 2004 Sidney Australia.* [Accessed 25 May 2005]. Available from World Wide Web: http://www.informatik.uni-trier.de/~ley/db/indices/a-tree/b/Bandt:Ros.html

BARBOSA, A. 2003. A Survey of Network Systems for Music and Sonic Art Creation. *Leonardo Music Journal, Volume 13*, MIT Press, pp. 53–59, 2003

BENJAMIN, J. 2002. Introducing Gesture Journal Gesture 1:1 *John Benjamins Publishing Company* pp.1–7 [Accessed 23 June 2005]. Available from World Wide Web: http://www.benjamins.com/jbp/journals/Gest/1-1/0001a.pdf

BENTLEY, P. J. CORNE, D. W. (eds.) 2001. *Creative Evolutionary Systems.* San Francisco, CA: Morgann Kauffman

BEYLS, P. 2005. Evolving Adaptive Sensors in a Synthetic Listener. *In: Proceedings International Computer Music Conference*, 5/9 *September, 2005, Barcelona,* Spain: International Computer Music Association pp. 563-567

BLACKWELL, T. M. 2004. Self-organised Music. *Organised Sound: Volume 9 no. 2.* Cambridge: Cambridge University Press pp.137-150

BLAINE, T. 2005. The Convergence of Alternate Controllers and Musical Interfaces in Interactive Entertainment. *In: Proceedings of the International Conference New Interfaces for Musical Expression Conference, 26/28 May 2005, Vancouver,* Canada. pp.27-33

BLAINE, T. FELS, S. 2003. Collaborative Musical Experiences for Novices. *Journal of New Music Research, 2003 Volume 32 no. 4* pp.411-428

BLAINE, T. FELS, S. 2003. Contexts of Collaborative Musical Experiences. *In: Proceedings of the International Conference on New Instruments for Musical Expression, 22/24 May 2003*, *Montreal, Quebec,* Canada. pp.129-134

BONGERS, B. 1998. An Interview with Sensorband. *Computer Music Journal 1998, Volume 22:1*, pp.13–24

BONGERS. B. 2000. Physical Interfaces in the Electronic Arts: Interaction Theory and Interfacing Techniques for Real-time Performance. *In:*

Wanderley, M Battier, M. (eds) Trends in Gestural Control of Music. Paris Ircam Centre Pompidou. pp.41-70

BOWERS, J. 2001. Tonetable: A Multi-User, Mixed-Media, Interactive Installation. *In: Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01), 6/8 December 2001, Limerick,* Ireland*.* pp.2-5

BUCHLA, D. 2005. A History of Buchla's Musical Instruments. *In: Proceedings of the International Conference on New Interfaces for Musical Expression (NIME-05), 26/28 May 2005 Vancouver,* Canada. pp.1-2

CAMURRI, A. 1995. Interactive Dance/Music Systems. *In: Proceedings of the International Computer Music Conference*, *September 1995*, *Banff,* Canada: International Computer Music Association pp.245-252

CAMURRI, A. 2000. Artificial Intelligence Architectures for composition and performance environment*. In: Miranda, E. R. (ed) Readings in Music and Artificial Intelligence (Contemporary Music Studies), Amsterdam.* Harwood

CAMURRI, A. 2004. Multimodal Interfaces for Expressive Sound Control *In: Proceedings of the 7th International Conference on Digital Effects (DAFx'04) 5/8 October 2004, Naples,* Italy. pp.41-4

CAMURRI, A. MAZZARINO, B. VOLPE, G. 2004. Analysis of Expressive Gesture: The Eyesweb Expressive Gesture Processing Library. *In: Gesture-based Communication in Human-Computer Interaction,* Springer Verlag

CAMURRI, A. DE POLI, G. LEMAN, M. VOLPE, G. 2001. A multi-layered conceptual framework for expressive gesture applications. *In proceedings of the International Mosart Workshop, Barcelona,* Spain*. 2001.* Pompeu Fabra

CAMURRI, A. HASHIMOTO, S. RICCHETTI, M. RICCI, A. SUZUKI, K. TROCCA, R. VOLPE, G. 2000. Eyesweb: Toward Gesture and Affect Recognition in Interactive Dance and Music Systems. *In: Computer Music Journal,* volume 24:1, pp.57-69

CAMURRI, A. MAZZARINO, B. VOLPE, G. 2004. Expressive gestural control of sound and visual output in multimodal interactive systems. *In: Proceedings of the International Conference on Sound and Music Computing 2004, Paris,* France

CAMURRI, A., S.HASHIMOTO, K.SUZUKI, R.TROCCA. 1999 KANSEI analysis of dance performance. *In: Proceedings of the International Conference Man and Cybernetics 1999, Tokyo.* Japan: IEEE

COOK, P. 2001. Principles for Designing Computer Music Controllers. *In: Proceedings of the 1st Workshop on New Interfaces for Musical Expression (NIME01) 2001 Seattle, Washington,* ACMSIGCHI, [Accessed 05 May 2004]. Available from World Wide Web:
http://portal.acm.org/citation.cfm?id=1085154

DAHLSTEDT, P. 2004. Sounds Unheard of: Evolutionary algorithms as creative tools for the contemporary composer. *Doctoral thesis 2004 Goteborg Sweden:* Chalmers Unviversit of Technology

DAWKINS, R. 1986 *The Blind Watchmaker*. London, Harlow: Longman 1986

DEPALLE, P. TASSERT, S. WANDERLY, M. 1997. Instruments Virtuels
Résonance, No. 12. Paris: IRCAM/Centre Georges-Pompidou: pp.5-8. [Accessed 21 June 2004]. Available from World Wide Web: www.ircam.fr/equipes/analyse-synthese/
listePublications/articlesRodet/ICMC01/WanderleyIcmc01.pdf.gz

DOBRIAN, C. BEVILACQUA, F. 2003. Gestural Control of Music Using the Vicon 8 Motion Capture. *In: Proceedings of the 2003 International Conference on New Interfaces for Musical Expression 22/24 May 2003, Montreal, Canada.* [Accessed 27 June 2004]. Available from World Wide Web: *http://www.nime.org/2003/onlineproceedings/TOC.html*

FOCKEN, D. STIEFELHAGEN, R. 2002. Towards Vision-based 3-D People Tracking in a Smart Room. *IEEE International Conference on Multimodal Interfaces, October 14/16 2002 Pittsburgh, PA, USA, ,* pp.400-405

GAYE, L MAZE, R. HOLMQUIST, L. 2003. Sonic City: The Urban Environment as a Musical Interface. *In: Proceedings of the International Conference on New Instruments for Musical Expression22/24 May 2003, Montreal, Canada.* [Accessed 27 June 2004]. Available from World Wide Web: *http://www.nime.org/2003/onlineproceedings/TOC.html*

GETZELS, J. W. CSIKSZENTMIHALYI, M. 1976. *The Creative Vision: a longitudinal study of problem-finding in Art*. New York: Wiley

GOGINS, M. 1995. Gabor Synthesis of Recurrent Iterated Function Systems *In: Proceedings of the International Computer Music Conference, December 1995, San Francisco,* California: International Computer Music Association pp. 349-348

HARTSON, W. R. WATSON, P. C. 1983. *The Psychology of Chess*, London.

HERMANN, T. KRAUSE, J. AND RITTER, H. 2002. Real Time Control of Sonification Models with a Haptric Interface. *In: Proceedings of the International Conference on Auditory Display, 2/5 July, 2002*, *Kyoto,* Japan

HEWITT, D. STEVENSON, I. 2003. Emic-Extended Mic-stand Interface Controller. *In: Proceedings of the 2003 Conference on New Interfaces for Musical Expression 22/24 May 2003, Montreal,* Canada. [Accessed 27 June 2004]. Available from World Wide Web: http://www.nime.org/2003/onlineproceedings/TOC.html

HOLLAN, J. HUTCHINS, E, KIRSCH, D. 2000. Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research *In: ACM Transactions on Computer-Human Interaction, Vol. 7, No.2, June 2000.* pp.174-196

HSU, W. 2005. Using Timbre in a Computer Based Improvisation System *In: Conference Proceedings International Computer Music Conference, 5/9 September 2005, Barcelona,* Spain: International Computer Music Association pp. 691-694

HUNT, A. WANDERLEY, M. AND KIRK, R. 2000 Towards a Model of Mapping Strategies for Instrumental Performance. *In: Proceedings of the International Computer Music Conference, September 2000, San Francisco, California:* International Computer Music Association, pp.209–212

IMPETT, J. 1999. Real times: implementing a temporal phenomenology in an interactive system, P. Walmsley and P.Wolfe (eds.) *In: Proceedings, Cambridge Music Processing.* Colloquium

IMPETT, J. A META-TRUMPET(ER). 1994. *In: Proceedings of the International Computer Music Conference, 12/17 September 1994, San Francisco,* California: International Computer Music Association, pp.147–150

JONG, J. 2005. Freesound Project. *In: proceedings of the International Computer Music Conference 5/ 9 September 2005, Barcelona,* Spain. [Accessed 6 September 2005]. Available from World Wide Web: http://freesound.iua.upf.edu/

JORDA, S 2004. Digital Instruments and Players: Part I – Efficiency and Apprenticeship. *In: Proceedings of the International Conference on New Instruments for Musical Expression, 3/5 June 2004, Hamamatsu, Shizuoka,* Japan. pp.59-63

JORDA, S. 2005. Multi-user Instruments: Models, Examples and Promises *In: Proceedings of the International Conference on New Instruments for Musical Expression, 26/28 May, Vancouver,* Canada. pp23-26

JORDA, S. 2003. Interactive Music Systems for Everyone: Exploring Visual Feedback as a way of creating more Intuitive, Efficient and Learnable Instruments. *In: Proceedings of the Stockholm Music Acoustics Conference, 6/9 August 2003, Stockholm,* Sweden. [Accessed 8 May 2004]. Available from World Wide Web: http://mtg.upf.edu/publications/smac03jorda.pdf

KAZUHIKO, T. TATSUMI, S. JUN, O. 1999. Real-time Estimation of Human Body Postures using Kalman Filter. *In: proceedings of the International workshop on Robot and Human Interaction RO-MAN'99, 27/29 September 1999, Pisa,* Italy. [Accessed 22 February 2004]. Available from World Wide Web:http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/7227/19488/00900338.pdf

KING, B. MINIUM, E. 2003. *Statistical Reasoning in Psychology and Education*, Fourth Edition. Hoboken, New Jersey: John Wiley & Sons, Inc.

KNAPP, R. B. LUSTED, H. S. 1990. A Bioelectric Controller for Computer Music Applications. *Computer Music Journal Volume 14 no.1,* pp.42–47

LEVITIN, D.J. 2002 Control Parameters for Musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound Volume 7 no.2,* pp.171-189

LIPPE, C. 2002. Real time Interaction Among Composers, Performers, and Computer Systems. *Information Processing Society of Japan SIG Notes, Volume 2002, Number 123,* pp.1-6

LIPPE, C. 1996. A Look at Performer/Machine Interaction Using Real-time Systems. *In: Conference Proceedings International Computer Music Conference, 19/24 August,* Hong Kong. International Computer Music Association. pp.116-117

LIVINGSTONE, D. 2001. The Space Between the Assumed Real and The Digital Virtual *In:* Ascott, R. *Reframing Consciousness Art, Mind and Technology* Intellect Books 2001 ISBN: 1841500135 pp.138-143

LIVINGSTONE, D. 2003. Emergent Behaviour in the context of Reactive Compositional Environments *In: Proceedings of the IX Brazilian Symposium on Computer Music (IX SBCM) Music as Emergent Behaviour, 2/8 August 2003, University of Campinas,* Brazil. pp.235-240

LIVINGSTONE, D. MIRANDA, E. 2004 Composition for Ubiquitous Responsive Environments *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami, Florida: International Computer Music Association.* pp.321-325

LIVINGSTONE, D. MIRANDA, E. 2005. Orb3 - Adaptive Interface Design for Real time Sound Synthesis and Diffusion within Socially Mediated Spaces *In: proceedings of the International Conference on New Instruments for Musical Expression, 26/28 May 2005, Vancouver,* Canada. pp.65-69

LIVINGSTONE, D. O'SHEA C. 2005. Tactile Composition Systems for Collaborative Free Sound *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami, Florida: International Computer Music Association.* pp.687-690

LIVINGSTONE, D. MIRANDA, E. 2005. Orb3 - Musical Robots within an Adaptive Social Composition System *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami,* Florida: International Computer Music Association. pp.543-546

LUCIER, A. 1995. Every room has its own melody, I am sitting in a room. *Reflections,* Köln: MusikTexte 1995. pp.94-102

MACHOVER, T . 2002. Instruments, Interactivity, and Inevitability. *In: Proceedings of the International Conference on New Instruments for Musical Expression, 24/26 May 2002, Dublin,* Ireland. pp.3 . [Accessed 10 February 2004]. Available from World Wide Web: http://portal.acm.org/citation.cfm?id=1085227

MACHOVER, T. J.CHUNG. 1989. Hyperinstruments: Musically intelligent and interactive performance and creativity systems. *In: proceedings of the International Computer Music Conference, 1989, San Francisco,* California: International Computer Music Association, pp.186-190

MAGNANENSI, G. ROLFE, C. 2005. Ligature 2003: for circuit bent speech synthesizer and educational toys. (Conference Performance)

In: proceedings of the International Conference on New Instruments for Musical Expression, 26/28 May 2005, Vancouver, Canada*.* pp.275-281

MAIA JR. A. MANZOLLI, J. DO VALLE, R. PEREIRA, N.S. L. 1999. A Computer Environment for Polymodal Music.*Organised Sound,* Cambridge University Press pp.32-40

MATHEWS, M. 1984. Electronic Violin: A Research Tool. *Journal of the Violin Society of America Volume 8 no.1* pp.71-88

MELO, A. DREVER, J. WIGGINS, G. 2005. Electroacoustic Performance Interfaces that Learn from their Users. *In: Proceedings of the International Computer Music Conference, 5/9 September 2005, Barcelona,* Spain. pp.567-570

MERLIER, B. 2004. Three-Dimensional Gestural Controller based on Eyecon motion capture system. *In: proceedings of the International Conference Sound and Music Computing, 20/22 October 2004,* IRCAM, Paris

MIRANDA, E. R. 2001. Evolving Cellular Automata Music: From Sound Synthesis to Composition. *In: proceedings of the sixth European Conference on Artificial Life. Workshop on Artificial Life Models for Musical Applications, 9 September 2001* Prague: Editoriale Bios

MIRANDA, E. R. 2000. On the Music of Emergent Behaviour: What can Evolutionary Computation bring to the Musician? *In: Annie S. Wu(ed) Genetic and Evolutionary Computation Conference, July 2000, Las Vegas, Nevada.* University of Central Florida GECCO. pp.140-142

MIRANDA, E. ZHANG, Q. 2005. Composition as Game Strategy: Making Music by playing board games against evolved artificial Neural Networks *In: proceedings of the International Computer Music Conference, 5/9 September, Barcelona,* Spain. pp.471-474

MIRANDA, E.M. WANDERLY, M. 2006 *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard.* The Computer Music and Digital Audio Series volume 21 A –R Editions, Inc. Wisconsin. ISBN 978-0-89579-585-4

MODDLER, P. MYATT, T. SAUP, M. 2003. An Experimental Set of Hand Gestures for Expressive Control of Musical Parameters in Realtime. *In: proceedings of the International Conference on New Interfaces for Musical Expression, 22/24 May 2003, Montreal,* Canada pp.146-151

MORONI, A. MANZOLLI, J. VON ZUBEN, F. GUDWIN, R. 2000. Vox Populi: An Interactive Evolutionary System for Algorithmic Music Composition *Leonardo Music Journal, Vol. 10:* MIT Press pp.49–54

MORONI, A. MANZOLLI, J. VON ZUBEN, F. GUDWIN, R. 1999. Evolutionary Computation applied to Algorithmic Composition. *In P. Angeline (ed.), Proceedings of the IEEE Congress on Evolutionary Computation 1999*, Piscataway, NJ: IEEE Press. pp.807-811

NAKRA, T. M. 2000. Ph.D. Thesis Inside the Conductors Jacket: Analysis, Interpretation and Musical Synthesis of Expressive Gesture M.I.T. Media Laboratory Perceptual Computing Section Technical Report No. 518 . [Accessed 25 March 2004]. Available from World Wide Web: http://dspace.mit.edu/handle/1721.1/9165

NAKRA, T. M. 2000. Searching for meaning in gestural data. *In M. Wanderly and M.Battier (Eds), Trends in Gestural Control of Music.* Paris: IRCAM-Centre Pompidou . [Accessed 22 March 2004]. Available from World Wide Web: http://www.ircam.fr/equipes/analyse-synthese/listePublications/articlesRodet/ICMC01/WanderleyIcmc01.pdf.gz

NELSON, P. 1997. The UPIC System as an Instrument of Learning. *Organised Sound Volume 2 no.1* pp.35-42

NEWTON-DUNN H. NAKANON H. GIBSON J. 2003. Block Jam: A Tangible Interface for Interactive Music. *In: proceedings of the Conference on New Interfaces for Musical Expression, 2003, Montreal,* Canada. [Accessed 14 April 2004]. Available from World Wide Web: http://citeseer.ist.psu.edu/577479.html

OVERHOLT, D. 2005. CREATE / STEIM The Overtone Violin *In: proceedings of the International Conference on New Interfaces for Musical Expression, 26/28 May 2005, Vancouver,* Canada. pp.34-37

PAINE, G. 2002. Interactivity, Where to from here? *Organised Sound Volume 7 no.3,* pp.295-304

PARADISO, J. 1997. Electronic music interfaces: new ways to play. *IEEE SpectrumMagazine, Vol. 34 no.12,* pp.18-30

PARADISO, J. 2002. The Edge of NIME: from Cult to Conference. *In: proceedings of the 2002 Conference on New Instruments for Musical Expression, 24/26 May, Dublin,* Ireland

PATTEN J., RECHT B., ISHII H. 2002 Audiopad: A Tag-based Interface for Musical Performance. *In: proceedings of the 2002 Conference on New*

Instruments for Musical Expression, 24/26 May, Dublin, Ireland. [Accessed 6 May 2004]. Available from World Wide Web: http://tangible.media.mit.edu/content/papers/pdf/audiopad_NIME02.pdf

ROKEBY, B. 1998. The construction of experience: Interface as content. *Digital Illusions: Entertaining the future with high technology, ed. Association for Computing Machinery,* Reading. Addison-Wesley

ROKEBY, DAVID 1995. Transforming Mirrors: Subjectivity and Control in Interactive Media. *In: Penny, Simon (ed.), Critical Issues in Electronic Media 1995,* USA: State University of New York

ROVAN, B. AND WECHSLER, R. 2000. The multi-dimensional mapping of movement-to-sound within an integrated Eyecon and Max Environment. Performance of Seine Hohle Form 10e *Symposium International des Arts Électroniques, December 2000*. *Paris,* France

ROVAN, B. WEISS, F. AND WECHSLER, R. 2001. Seine hohle Form: a project report. Artistic Collaboration in an Interactive Dance and Music Performance Environment. *In proceedings of the International Conference on Computational Semiotics in Games and New Media, 10/12 September 2001*, Amsterdam

ROWE, R. 1993. *Interactive Music Systems: Machine Listening and Composing.* The MIT Press, Cambridge, Massachusetts

ROWE, R. 2004. *Machine Musicianship.* The MIT Press, Cambridge, Massachusetts

SCHEIMER, G. SABIR, K. HAVRYLIV, M. 2004. Pocket Gamelan: A j2me Environment for Just Intonation *In: Proceedings of the International*

Computer Music Conference, 1/6 November 2004, Miami, Florida: pp.654-657

SELIGMANN, D. 2004. Enhancing Engagement in Multimodality Environments by Sound Movements in a Virtual Space. *Artful Media IEEE Computer Society.* [Accessed 17 June 2005]. Available from World Wide Web: *http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/93/28736/01289034.pdf*

SETTEL, Z. LIPPE, C. 2003. Convolution Brother's Instrument Design *In: Proceedings of the International Conference on New Interfaces for Musical Expression*, 22/24 May 2003, Montreal, Canada. pp.197-200 [Accessed 14 April 2004]. Available from World Wide Web: http://www.music.mcgill.ca/musictech/nime/onlineproceedings/Papers/NIME03_Settel.pdf

SINGER, E. FEDDERSON, J. BOWEN, B. 2005. A large Scale Networked Robotic Musical Instrument Installation. *In: proceedings of the International Conference on New Instruments for Musical Expression, 26/28 May 2005, Vancouver,* Canada. pp.50-55

SINGER, E. LARKE, K. BIANCIARDI , D. 2003. LEMUR GuitarBot: MIDI Robotic String Instrument. *In: Proceedings of the International Conference on New Interfaces for Musical Expression*, 22/24 May 2003, Montreal, Canada [Accessed 22 April 2004]. Available from World Wide Web: http://www.music.mcgill.ca/musictech/nime/onlineproceedings/Papers/NIME03_SingerLarke.pdf

SPARACINO, F. DAVENPORT, G. PENTLAND, A. 2000. Media in performance: Interactive spaces for dance, theater, circus, and museum exhibits

IBM Systems Journal Vol. 39, Nos. 3 and 4, 2000. p.479

SPARACINO, F. WREN, C. DAVENPORT, G. PENTLAND, A. 1999. Augmented Performance in Dance and Theater MIT Media Lab [Accessed 3 June 2004]. Available from World Wide Web: http://xenia.media.mit.edu/~flavia/Papers/flavia_augmented_performance.pdf

SPICER, M. 2004. AALIVENET: An agent based distributed interactive composition environment. *In: proceedings of the International Computer Music Conference, 1/6 2004,* Miami: International Computer Music Association. pp.211-214

TANAKA, A. 2000. Musical Performance Practice on Sensor-Based Instruments. *In M. Wanderly and M. Battier (eds), Trends in Gestural control of Music.* Paris: Ircam Centre Pompidou.
TANAKA, A. 2004. Mobile Music Making. *In: proceedings of the International Conference on New Instruments for Musical Expression, 3/5 June 2004,* Hamamatsu, Shizuoka, Japan. pp.154-156

TANAKA, A. AUTHIER, A. VALADON, G. VIOLETTE, L. 2004. Enhancing Musical Experience through Proximal Interaction. *In: proceedings of the International Computer Music Conference, 1/6 2004,* Miami: International Computer Music Association. pp.205-209

TAYLOR, R. TORRES, D. BOULANGER, PIERRE 2005. Using Music to Interact with a virtual Character. *In: Proceedings of the International Conference, New Interfaces for Musical Expression 5/9 September, Vancouver,* Canada. pp.220-223

THIEBAUT, J. A 2004. A Graphical Interface for Trajectory Design and Musical Purposes. *In: Proceedings of the International Computer Music Conference, 1/6 November, Miami,* Florida: International Computer Music Association

TIMMERMANS, H. WOLFEREN, G.NEWLAND, P. KUNATH, S. 2000. MEDIATE: Key Sonic Developments in an Interactive Installation for Children with Autism Conference. *In: proceedings of the International Computer Music Conference, 1/6 November 2004, Miami,* Florida: International Computer Music Association

TRACEY, A. 1963. Three tunes on the mbira dzaVadzimu. *African Music, Volume 3 no. 2.* pp.23-26

VARELA, FRANCISCO, J. 1991. *The Embodied Mind* Cambridge, Massachusetts: MIT Press

WANDERLEY M. AND BATTIER M. 2000. Trends in Gestural Control of Music. Ircam - Centre Pompidou, 2000. [Accessed 12 May 2004]. Available from World Wide Web: http://recherche.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/

WANDERLY, M. ORIO,N 2002. Evaluating Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal Vol. 26, Issue 3: New Performance Interfaces 2002.* MIT press pp.62-67

WASSERMAN, K., M. BLANCHARD, U. BERNARDET, J. MANZOLLI, P. VERSCHURE 2000. ROBOSER: An Automous Interactive Musical Composition System. *In: proceedings of the International Computer Music Conference, September 2000*, Berlin: International Computer Music Association. pp.531-534

WATERS, P. ROWE, A. 2004. Alt-Space: Audio-visual interactive software for developing narrative environments. *In: proceedings of the conference on Computers in Art and Design Education CADE 04 Copenhagen, Denmark* [Accessed 11 September 2005]. Available from World Wide Web: http://www.cade2004.cbs.dk/proceedings/fullpapers/15_waters_final_fullpaper.pdf

WECHSLER, R. WEISS, F. 2004. Motion Sensing for Interactive Dance, *IEEE-Pervasive Computing, Mobile and Ubiquitous Systems,* Jan-March 2004. [Accessed 16 June 2005]. Available from World Wide Web: http://csdl2.computer.org/persagen/DLAbsToc.jsp?resourcePath=/dl/mags/pc/&toc=comp/mags/pc/2004/01/b1toc.xml&DOI=10.1109/MPRV.2004.1269129

WEINBERG G. 2002. Interconnected Musical Networks: Bringing Expression and Thoughtfulness to Collaborative Music Making. Ph.D. Thesis. MIT Media Lab, Cambridge Massachusetts. [Accessed 10 October 2004]. Available from World Wide Web: http://www.media.mit.edu/hyperins/gil_weinberg_phdthesis.pdf

WEINBERG, G. AIMI, R. JENNINGS, K. 2002. The Beatbug Network: A Rhythmic System for Independent Group Collaboration. *In: proceedings of the International Conference on New Interfaces for Musical Expression, 2002, Dublin,* Ireland. [Accessed 15 October 2004]. Available from World Wide Web: http://nime.org/2002/proceedings/paper/weinberg.pdf

WEINBERG, G. DRISCOLL, S 2005. Connecting Novices and Experts Through Collaborative Improvisation. *In: proceedings of the International Conference on New Instruments for Musical Expression, 26/28 May 2005, Vancouver,* Canada. pp.17-22

WESSEL, D. 1991. Improvisation with Highly Interactive Real-Time Performance System. *In: proceedings of the International Computer Music Conference,* San Francisco: Computer Music Association, pp.344-347

WESSEL, D. WRIGHT, M. SCHOTT, J. 2002. Intimate Musical Control of Computers with a Variety of Controllers and Gesture Mapping Metaphors. *In: proceedings of the International Conference on New Interfaces for Musical Expression, 24/26 May 2002,* Dublin. http://nime.org/2002/proceedings/paper/wessel.pdf

WHALLEY, I. 2004. Adding Machine Cognition to a Web-based interactive composition. *In: proceedings International Computer Music Conference, 1/9 November, Miami 2004,* pp.197-200

WHALLEY, I. 2005. Software Agents and Creating Music/Sound Art: Farmes, Directions and where to from here. *In: proceedings of the International Computer Music Conference, Barcelona 2005,* International Computer Music Association, pp. 691-694

WINKLER T. 2000. Audience Participation and Response in Movement-Sensing Installations. *In: proceedings of the International Computer Music Conference, 2000, Berlin,* Germany. [Accessed 18 October 2004]. Available from World Wide Web: http://www.brown.edu/Departments/Music/sites/winkler//papers/Audience_Participation_2000.pdf

WINKLER, T. 1995. Making Motion Musical: Gestural Mapping Strategies for Interactive Computer Music. *In: proceedings of the International Computer Music Conference, 1995,* Hong Kong: International Computer Music Association, pp.261–264

WINKLER, TODD 1998. *Composing Interactive Music – Techniques and Ideas Using Max.* Cambridge, Massachusetts: MIT Press

WOOLF S. AND BECK T. 2002. Experiments with Reactive Robotic Sound Sculptures. ALife VIII: Workshop Proceedings. [Accessed 27 October 2004]. Available from World Wide Web:
http://www.alife.org/alife8/workshops/6.pdf

WRIGHT M. FREED A. MOMENI A. 2003. OpenSoundControl: State of the Art. *In: proceedings of the Conference on New Interfaces for Musical Expression, 22/24 May, Montreal,* Canada. [Accessed 27 October 2004]. Available from World Wide Web:http://www.music.mcgill.ca/musictech/nime/onlineproceedings/Papers/NIME03_Wright.pdf

XIN WEI SHA, GIOVANNI IACHELLO, STEVEN DOW, YOICHIRO SERITA, TAZAMA ST. JULIEN, JULIEN FISTRE 2003. Continuous Sensing of Gesture for Control of Audio-Visual Media. *In: proceedings of the Seventh IEEE International Symposium on Wearable Computers.* IEEE Computer Society

YEO, W. BERGER, J. LEE, Z. 2004. Sonart: A Framework for data Sonification, Visualisation and Networked Multimedia Applications. *In: proceedings of the International Computer Music Conference, 1/6 November 2004, Miami,* Florida: International Computer Music Association, pp.180-184

ZHU, Y. REN, H. XU, G. LIN, X. 2000. Toward Real-time Human-computer Interaction with Continuous Dynamic Hand Gestures. *In: proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 26-30 March 2000, Grenoble*, France. IEEE Computer Society pp.544

ZICARELLI, D. 1991. Communicating with meaningless numbers. *Computer Music Journal, Volume 15, no. 4.* MIT Press. pp. 74–77

Website References:

1. Max/MSP (software) http://www.cycling74.com
[Accessed 10/08/2005

2. LISP (software) Common Lisp documentation & resources
http://clisp.cons.org/resources.html
accessed 3/08/2005

3. Pure Data (software) Community site http://puredata.info/
accessed 3/07/2004

4. Zero G (software) Lola Leon Miriam (Vocaloid series)) http://www.zero-g.co.uk/index.cfm?Articleid=805
accessed 11/08/2005

5. Ableton Live (software) http://www.ableton.com/
accessed 11/03/2004
6. Big eye (software) http://www.steim.org/steim/bigeye.html
accessed 15/03/2003

7. Eyesweb (software) Resource and Research community pages
http://www.infomus.dist.unige.it/eywindex.html
accessed 17/04/2004

8. Eyecon (software) by Frieder-Weiss
http://www.frieder-weiss.de/eyecon/index.html
accessed 17/04/2004

9. Processing (software) http://processing.org/
accessed 5/09/2006

10. Max Objects Database (software)
http://www.maxobjects.com/
accessed 10/08/2004

11. Tap Tools (software) http://www.electrotap.com/taptoolsmax/
accessed 13/03/2004

12. CV.jit externals for Max/MSP by Pelletier, J. M.
http://www.iamas.ac.jp/~jovan02/cv/
accessed 7/04/05

13. Macromedia (now Adobe) http://www.adobe.com/
accessed 7/04/06

14. Lemur Interface http://www.jazzmutant.com/
accessed 3/03/05

15. Ambisonics Overview http://www.ambisonic.net/
accessed 3/04/05

16. Ambisonics externals for Max/MSP http://www.icst.net/downloads/
accessed 3/04/05

17. M-Audio Firewire410 interface
http://www.maudio.co.uk/products/en_gb/FireWire410-main.html
accessed 23/03/04

18. Simon Game http://en.wikipedia.org/wiki/Simon_(game)
accessed 14/05/06

19. RTC Lib (Real Time Composition Library) for Max/MSP by Karl Heinz Essl
http://www.essl.at/software.html
accessed 10/08/2006

20. *Metrosonics* Paine, G.
http://www.meterosonics.com/
accessed 12/09/2006

21 . *Stanza* - Sound and Video Artist
http://www.stanza.co.uk/i_spy/index.htm
accessed 04/09/2006

22. *Forest Walk* Cardiff, J.
http://www.abeymdeia.com/Janweb/forest.htm
accessed 12/09/2005

23. *Java 2 Micro Edition*
http://java.sun.com/javame/index.jsp
accessed 22/09/2006

24. *Bopit* by Hasbro
http://www.hasbro.com/bopit/
accessed 23/09/06

25. *Alt Zero* installation and Online collaborative sound environment.
http://www.squidsoup.org/altzero/
accessed 22/10/05

26. *Machinima*: further information and links from Wikipedia:
http://en.wikipedia.org/wiki/Machinima
accessed 14/07/2005

27. *EyeToy*:Video Tracking System used in novel game product
http://www.eyetoy.com
accessed 25/04/05

28. *Gametrak* by in2games: novel gesture capture for action/sports games
http://www.in2games.uk.com/corporate/index.htm
accessed 22/03/04

29. *Harmonix* Game Developer noted for music related games and novel interfacing. http://www.harmonixmusic.com/
accessed 25/06/06

30. *Revolution* gesture based game controller by Nintendo
http://www.nintendo.com/newsarticle?articleid=TfyOgnUc7zB-ZlzdQvMhKJOgS5LsR2nK accessed 11/07/05

31. *Sound Garten* by Wolf, 2003
http://www.soundgarten.com/contact.html
accessed 21/09/04

32. *ISS Cube* by Quarta, M. 2003
http://www.somethingonline.org/txt/d_iss.html
accessed 20/06/04

33. Lucier, A. Archive, interview and documentary.
http://www.ubu.com/sound/lucier.html
accessed 03/04/06

34. Jong, J. 2005
Freesound Project International Computer Music Conference 2005 September 5th to 9th http://freesound.iua.upf.edu/
accessed 04/09/05

35. *Lemur* touch screen interface
http://www.jazzmutant.com/
accessed 08/04/05

36. *Wild Divine* Bio Feedback game system
http://www.wilddivine.com
accessed 10/01/04

37. *VisCog* Visual Cognition Laboratory at the University of Illinois. Change Detection Research
http://viscog.beckman.uiuc.edu/djs_lab/demos.html
accessed 23/05/04

38. *Phatboy* MIDI interface/controller from Keyfax
http://www.keyfax.com/keyfax/phatboy/phatboy1.html
accessed 20/09/06

39. *Spin Doctor* MIDI interface/controller from Kenton electronics
http://www.kentonuk.com/
accessed 20/09/06

40. *Pocketfader* MIDI interface/controller from Doepfer

http://www.doepfer.de/home.htm
accessed 29/09/06

41. *UC16, UC33e* USB MIDI interface/controllers from Evolution
http://www.MIDI-store.com/shop/detail.cgi?id=122381
accessed 20/09/05

42. *Trigger Finger* (USB/MIDI interface/controller) m audio
http://www.m-audio.com/products/en_us/TriggerFinger-main.html
accessed 21/03/05

43. *i- cube* (CV to MIDI interface) Infusion Systems
http://infusionsystems.com/catalog/product_info.php/cPath/21/products_id/28
accessed 22/02/05

44. *Miditron* (CV to MIDI interface) by Eric Singer Eroktronix
http://eroktronix.com/
accessed 13/03/04

45. *Arduino* (I/O board for sensors/lights motors)
http://www.arduino.cc/
accessed 13/03/04

46. *Kroonde* (rack mounted sensor interface)
http://www.la-kitchen.fr/kitchenlab/kroonde-en.html
accessed 12/03/04

47. *Circuit Bending*
http://www.anti-theory.com/soundart/circuitbend/
accessed 21/09/06

48. *Controller Band*
http://www.controller-band.de/
accessed 21/09/06

49. *Junxion* MIDI mapping software from Steim
http://www.steim.org/steim/products.html
accessed 18/09/05

50. *Wireless Micron* game controller by Mad Catz
http://www.madcatz.com/Default.asp?Page=28
accessed 11/02/05

51. *Reason* sequencing/mixing software by Propellerheads
http://www.propellerheads.se/
accessed 22/05/06

52. *Cubase* sequencing/mixing software by Steinberg
http://www.steinberg.net/
accessed 22/05/06

53. *Logic* sequencing/mixing software by Apple
http://www.apple.com/logicpro/
accessed 23/09/05

54. *Arch-OS* Portland Square Plymouth
http://www.arch-os.com/
accessed 27/05/07

55. *Artificial Tango Library* (Audio analysis algorithms for MaxMSP by Olaf Mathews)
http://artificial-tango.nullmedium.de/
accessed 10/05/06

56. *Sonic Visualiser* developed at the Centre for Digital Music,
Queen Mary, University of London.
http://www.elec.qmul.ac.uk/digitalmusic/

57. *MJX nnLists* (Artificial neural network externals for MaxMSP using Java
http://miajo.co.uk/?MaxMSP
accessed 21/04/06

58. *RTClib* (Real-time composition library for MaxMSP by Karl Heinz Essl
http://www.essl.at/software.html
accessed 23/05/06

Sound Recordings:

Lucier, A. 1990. *I am Sitting in a Room* (1980 sound recording). New York, Lovely Music.

Lucier, A. 1995. *Reflections / Reflexionen*. Köln, MusikTexte.

Lucier, A. 1997. *I am Sitting in a Room* (1970 sound recording). With a minimum of means. Målsryd, Content.

Referenced Developmental Works and Events:

PoundtheSound (2006)
Adaptive Games prototype Crimp, M. Ford, C. supervisor Livingstone, D. for MSc Intelligent Interactive Systems, University of Plymouth

Livingstone, D. Swain, E. (1997)
Soundspace: Software prototype. Technical Implementation: Emile Swain, Concept and Design: Dan Livingstone. University of Plymouth.

ICCMR (2004)
Interdisciplinary Center for Computer Music Research, Computer Music Seminar Portland Square, February 2004, University of Plymouth.

Appendix 1: Glossary of Terms

Qualities, System Models, Interactions Models, Behaviours and *Features*.

Features

- *User number*
  Individual, group or distributed network.
- *User role flexibility*
  Collaborative interface supporting different interaction models for participants supporting learning and efficiency (Jorda 2004))
- *User mapping*
  Parameter mapping by user
- *Adaptive mapping*
  System/interface re assigns parameters during live interaction (Livingstone 2005)
- *Direct motion tracking*
  Performer/participant proximal interaction/body gesture (Tanaka 2004)
- *Indirect motion tracking*
  Relative or proximal tracking of performers/participants.
- *Direct gesture acquisition*
  Data collection through integrated sensors. Typically, these are mounted on an instrument or embedded in a novel controller.
- *Indirect gesture acquisition*

Analysis of sound output to determine gesture based on knowledge of instrument properties or playing context.

- *Live synthesis and/or diffusion (Transformative interactions)*

  Encourages participants to *play intercept, change or make contributions* (Bandt, 2004)

- *Intended audience/end User*

  General Public

  Performer/Musician

  Novice (individual)

  Novices (group)

Qualities

- *Gestural Vocabulary* (Mathews 1984)

  Interface supports wide range of established and new musical gestures.

- *Nourishing* (Machover 2002)

  System/interface motivates continued discovery and creativity through audiovisual feedback

- *Dynamic Context* (Winkler 1998)

  System/interface motivates player/participant spontaneity within an evolving musical framework.

- *Player Paradigm* (Rowe 1993)

  System exhibits player like behaviour or machine musicianship.

- *Self-Organising* (Blackwell 2004) (Whalley 2005)

An open or conversational generative framework; providing sophisticated interactions between people and improvisational systems.

<u>System Models</u>

- *Performance* model (Rowe 1993)

  Assigned interpretative role players/participants with reactive music.
- *Instrument* model (Wanderly 2000)

  Extended conventional instrument with additional sensing for solo performer – *explore-ability*
- *Unencumbered* model (Cammuri 1995)

  System uses video analysis for data acquisition (*direct* and *indirect* gestures within controlled environment performance/installation).
- *Compositional* model (Winkler 1998)

  System designer/composer provides compositional framework, typically mapping performance data to interpretative musical parameters for performer/participant interaction.
- *Distributed* model (Weinberg 2002)

  Novel interface forms part of a collaborative network manipulating shared content.
- *Collaborative* model (Jorda 2000)

Typically, a tracking system monitors generic *Tangible* objects manipulated on a table surface with symbol encoding for parameter mapping of group interaction.

Interaction models

- *Control* Supports direct manipulation of musical output parameters.
- *Sequential* Supports linear ordering of defined musical elements, typically samples or events.
- *Organisational* Supports non-linear restructuring of defined elements or events.
- *Relational* Motivates musical relationships between objects or symbols through manipulation and structuring of elements.
- *Conversational* Motivates and sustains a musical dialogue between system and performer/participant.
- *Transformative* Extends musical dialogue by adapting content and processes in response to identified behaviour.

Behaviours

- *Exploratory* (Wanderley 2002)
  Performer/participant discovers new gestures/interactions during use.

- *Interpretative* Performer/participant assigns system actions/events to personal context/goal.

- *Transformative* (Modler et al., 2003, Thiebaut 2004) Performer/participant or system mediated abstraction of musical material.
- *Sociable* (Miranda 2001) Learning and memory modelled in software in relation to evolving repertoire. Either a system or performer/participant attribute.

This glossary of terms is established in chapter 1 and contextualised in chapter 2. This structured collection of field specific vocabulary consolidates these terms into identifiable groups.

An example table is shown comparing features of different digital instruments. A scale 1 to 5 is used to indicate the number or level, for example an interface with an *exploreability* level of 5 indicates the highest level of influence over sounds created and gestures recognized by the system. Items left blank do not feature the referenced element. Of course a system may score low in a given area because it is not designed offer that function to the intended user or audience. This table collates the evaluative vocabulary from current literature while adding new categories and behaviours to provide an evaluative framework for New Digital Instruments.

System	*Overtone*	*Reactable*	*Beatbugs*	*Orb3*
Date	9/05	9/05	5/05	06
Features				
• *user number*	1	1-5	5-10	1-5
• *user role flexibility*	1	4	3	4
• *user mappin g*	5	4	0	0
• *adaptive mapping*	0	0	0	3
• *direct motion trackin g*	5	0	0	4
• *indirect motion tracking*	0	3	1	4
• *direct gesture acquisition*	5	0	2	4
• *indirect gesture acquisition*	2	0	0	0
• *live synthesis*	3	3	1	3
• *live diffusion*	2	2	1	4
• *intended audience/user*	musician	mixed	novice	mixed
Qualities				
• *Gestural Vocabula r y*	5	2	1	3
• *Nourishing*	5	3	1	3
• *Dynamic Context*	2	4	2	4
• *Player Paradig m*	1	1	0	3
• *Self Organising*	0	3	1	4
System Models				
• *Performanc e*		3	1	
• *Instrumen t*	5			
• *Unencumbered*		2		4
• *Compositional*		3		2
• *Distribute d*		2	2	
• *Collaborati v e*		4	3	4
Interaction model				
• *control*	5	4	2	3
• *sequential*	5	2	4	2
• *organisational*	3	3		5
• *relational*		5		3
• *conversational*				2
• *adapti v e*				3
Behaviours				
• *explorator y*	5	3	2	4
• *interpretative*	5	3		3
• *transformative*	5	2	1	3
• *sociabilit y*	4			3

Examples referenced: *Overtone* Violin (Overholt 2005) *Reactable* (Jorda 2005*)* *Beatbug* (Weinberg 2004) *Orb3 system design/prototype* (Livingstone 2006).

Appendix 2: Pilot Study Data

Appendix 3: Published Papers

LIVINGSTONE, D. 2001 The Space Between the Assumed Real and The Digital Virtual *In:* Ascott, R. *Reframing Consciousness Art, Mind and Technology* Intellect Books 2001 ISBN: 1841500135 pp.138-143

LIVINGSTONE, D. 2003 Emergent Behaviour in the context of Reactive Compositional Environments *In: Proceedings of the IX Brazilian Symposium on Computer Music (IX SBCM) Music as Emergent Behaviour, 2/8 August 2003, University of Campinas,* Brazil. pp.235-240

LIVINGSTONE, D. MIRANDA, E. 2004 Composition for Ubiquitous Responsive Environments *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami, Florida: International Computer Music Association.*
pp.321-325

LIVINGSTONE, D. MIRANDA, E. 2005 Orb3 - Adaptive Interface Design for Real time Sound Synthesis and Diffusion within Socially Mediated Spaces *In: proceedings of the International Conference on New Instruments for Musical Expression, 26/28 May 2005, Vancouver,* Canada. pp.65-69

LIVINGSTONE, D. O'SHEA C. 2005 Tactile Composition Systems for Collaborative Free Sound *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami, Florida: International Computer Music Association.*
pp.687-690

LIVINGSTONE, D. MIRANDA, E. 2005 Orb3 - Musical Robots within an Adaptive Social Composition System *In: Proceedings of the International Computer Music Conference, 1/6 November 2004, Miami,* Florida: International Computer Music Association.
pp.543-546

The Space between the Assumed Real and the Digital Virtual

Dan Livingstone

In writing this paper my first task is to engage the reader in the conceptual framework proposed in order to identify the space I envisage. Some clarification of the terms I use will be necessary in order to focus our attention within this framework, as owing to its interdisciplinary nature, more avenues for exploration and debate arise than are resolved.

I use the *assumed real* to identify the physical space our bodies inhabit; reality is too definitive a term and presupposes that we all are in agreement as to the space that this term encompasses. Our individual realities are informed by our experience of the world, more frequently mediated through technology. An obvious but nevertheless compelling example of this is the exponential increase in mainstream broadcast television framing the output of a network of live CCTV and surveillance systems as infotainment where the recursive spectacle of us consuming our various levels of antisocial behaviour is as much a part of the content as the digital reprocessing which allows us to view it.

> Although we know the imagery from video cameras in banks and supermarkets is relayed to a central control room, although we can guess the presence of security officers, eyes glued to control monitors with computer aided perceptions – visionics - it is actually impossible to imagine the pattern, to guess the interpretation produced by sightless vision.
>
> (Virilio 1994).

By *digital virtual* I encompass screen-based representation and simulation systems that provide interfaces for either individual or group experiences that are manifested within digitally generated spaces. This would include gaming environments, collaborative virtual workspaces and other forms of networked virtual reality.

If we were to attempt to reach agreement about how one of the sensorimotor systems that informs our perception of physical space operates, we would become sidetracked, as these biological and psychological mechanisms do not operate independently. Francisco Varela et al. provide an excellent illustration of this potential debate, discussing visual perception by asking what came first, the world or the image and suggesting that researchers from both cognitivist and connectionist persuasions would support the "chicken position".

> The world out there has pregiven properties. These exist prior to the image that is cast on the cognitive system, whose task is to recover them appropriately (whether through symbols or global subsymbolic states
>
> (Varela et al. 1991).

The alternative to this very convincing statement is given as the egg position:

> The cognitive system projects its own world, and the apparent reality of this world is merely a reflection of internal laws of the system."
>
> (Varela et al. 1991).

These inner and outer extremes focus on the visual as the central issue. This is also a recurrent theme in the discussion of the perception of physical space, whether from a philosophical or scientific perspective.

> Inside and outside form a dialectic of division, the obvious geometry of which blinds us as soon as we bring it into metaphorical domains
>
> (Gaston Bachelard *Poetics of Space*).

The space I envisage not only falls between these divisions but in a sense is manifested through the friction that is generated between them. It can only exist because we are prepared to engage with it, simultaneously authoring and documenting our own experience through interaction with both real and virtual space in the same moment. This space could be seen as augmented reality but this term is misleading. Conventional manifestations of augmented reality are usually dependent on the subject having additional information processing power integrated with his existing sensorimotor capacities; a pilot with digital overlays of objective navigational and strategic data has an increased view of the physical world through the system. The emphasis is on augmenting what is already there. The space between the assumed real and the digital virtual is not 'already there'. We are complicit in generating it through interacting with a system that enables us to occupy this space both on a physical and conceptual level.

Sorcerers Apprentice (Livingstone 1997) is a current project that enables us to occupy this space. It is a tri-part system that traverses perceived space, conceived space and lived space (Lefebvre 1991) simultaneously. This space could be considered in terms of Soja's 'Thirdspace' although in its current form the social and collective implications of the mechanism have not been resolved.

> Everything comes together in Thirdspace: subjectivity and objectivity, the abstract and the concrete, the real and the imagined, the knowable and the unimaginable, the repetitive and the differential, structure and agency, mind and body, consciousness and the unconscious, the disciplined and the transdisciplinary, everyday life and unending history.
>
> (Soya 1996) .

The system is described in phases in order to identify the various processes through which this integrated space is realised. This system depends initially on collaboration between a subjective human agent and an objective artificial agent, although as it evolves, the boundaries between agents blur.

Phase One: Perceptual Shifts

The first goal is to move the human agent from a position of spectator/user to that of performer/instigator. It is necessary to strip away the critical and physical limitations of a screenspace and to relocate the human agent's perspective from a position of viewer of graphical virtual representation to conceiver of spatial relationships through interaction with a limited but focused digital mechanism. The mechanism is not intended to augment our perception of physical reality but to refocus our cognitive processes on the physical space we occupy, thereby enabling us to overlay our own individual subjective responses to that space. An environment is suggested and extended through our interaction with it; sound is manipulated to describe our perception of this space. As we drag objects from the screenspace into the actual space our bodies inhabit, the transient perceptual shifts that occur allow us to operate in the grey area between representation and interpretation.

> Tight linkage between visual, kinesthetic, and auditory modalities is the key to the sense of immersion that is created by many computer games, simulations and virtual- reality systems"
>
> (Laurel 1993).

In order to break our preconceptions of what interacting with a computer mediated experience might be and anchor this experience in the *assumed real,* a graphical metaphor on a screen is used. Objects on a screen can be manipulated via infra red. We are both part of and separate from this interaction, emphasising our participatory complicitness in the generation of subjective spatial referencing. The design of each object references the compass, our tool for navigating physical space. As we drag it across the screen its readout changes, numerical coordinates update. Even as we drag the objects from this screenspace, the readout continues to relay its position in space. As the visual representation is limited to the screenspace, three dimensional spatialised sound is used to sustain our perception of the object that we have moved from the digital virtual into the physical space we currently occupy. We are engaged, we perceive the object to be there and if we doubt it's presence, a glance back to the screen will affirm its location in degrees. To further reinforce our suspension of disbelief we can continue to manipulate the object, drawing it closer to us or moving it away. In harnessing the physical movement of the human agent as a means to manipulate perceptually generated objects within physical space, we are using a form of embodied interaction as discussed by Lakoff and Johnson's 'experientialist approach to cognition'.

> Meaningful conceptual structures arise from two sources: (1) from the structured nature of bodily and social experience and (2) from our innate capacity to imaginatively project from certain well structured aspects of bodily and interactual experience to abstract conceptual structures. Rational thought is the application of very general cognitive processes -focusing, scanning, superimposition, figure-ground reversal"
>
> (Lakoff 1988).

As we continue to explore this mechanism by dragging additional objects into being, the conceptual relationships we make between the *assumed real* and the *digital virtual* form the foundations of a new space. Through perceiving the objects we have co-created as coordinates within physical space,we have produced a perceptual construct, that is, a series of relationships that can be documented and explored.

> The business of interaction between an individual and a machine (as opposed to interaction with an installation and in public) is to reinvent passion, which I would understand as the flash of fire that sparks from the crunching of familiar gears in unfamiliar combinations"
>
> (Cubitt 1995).

Phase Two: Translation

As we engage with the perceptual construct, the artificial agent translates our activity and broadcasts it in a remote location and our interactions are recreated. This 'remote location' can either be a physical space such as a gallery, using physical projection such as a laser system where transition between points is reminiscent of a child drawing in space using a sparkler, or within a virtual space created dynamically using VRML. In either case, the perceptual construct, the broadcast of this visual mapping of our auditory interactions has been disembodied from its physical context. But as the artificial agent learns through participation in the recreation of the perceptual construct, it's own ability to collaborate in the realisation of further constructs emerges. This process of translation is significant as it provides documentation of the perceptual constructs generated by different subject's, thereby enabling us to build comparative data in order to explore the nature of each subjects response to this speculative mechanism.

Overlaying a series of perceptual constructs within one virtual space would allow a far deeper exploration of the conceptual relationships that each individual brings to the system. For example, if we provide constraints to our subject's interaction within the system, the perceptual differences between constructs will become more acute. We will use a very simple example: Visualise, if you will, eight points in space; now compose them within the physical space you currently occupy. We will impose constraints: the points must be positioned so that a cube can be created by joining them; no point can fall outside the room you occupy. Each manifestation of the cube will be different; each individual's construct will be different. The more we engage with the concept and react to the constraints, the more creative the solutions will become. The key here is to underline the fact that however you compose the cube mentally, it has a relationship with the physical world, or more accurately, our experience of this world. The constraints have sharpened our awareness of the perceptual constructs context. If we could now retrieve all the diverse instances of the cube that had been previously generated within this space, by overlaying our perceptually constructed cube with diagrammatic references to previous cubes, how would this level of experience affect our strategies for co-forming more sophisticated perceptual constructs. I use the example of the cube to illustrate how the system could interpret and reinforce this perceptual process; but within this critical space there are no such constraints. How an individual interprets these points is not defined:

> A point in space seems perfectly objective. But how are we to define the points of our everyday world? Points can be taken either as primitive elements, as intersecting lines, as certain triples of intersecting planes, or as certain classes of nesting volumes. These definitions are equally adequate, and yet they are incompatible: what a point is will vary with each form of description"
>
> (Varela et al. 1991).

Phase Three: Integration.

A human agent within this world could attempt to describe the physical space occupied to the artificial agent within the system; this process of communication would not be a one-way dialogue, as the system would have the capacity to 'question' the perceptual construct being generated by integrating it with the physical environment, either by real-time projected diagrammatic overlays through a robotics or augmented vision system, or more simply; by modelling the space through VRML. It would be equally valid for the artificial agent to generate space by storing, reconstructing and extending previous perceptual constructs in order to question our methods of communication.

The third layer of this mechanism for spatial generation and critical exploration is the integration of the two previous phases, providing a critical space that is formed through the collaboration of both human and artificial agents. Interacting with this system has the potential for refocusing the way we perceive and engage with both physical and digitally generated spaces. We have the potential to bring compositional and spatial aesthetics into these relationships. Will the system also develop this potential? And by interacting with this tri-part system, can we evolve a more objective or critical approach to the way we interact with technologically mediated experiences?

References

Bachelard, G. 1994 *The Poetics of Space*. Boston: Beacon Press.

Cubitt, S. 1995. From an essay on interaction for VideoPositive.

Lakoff, G. 1988 "Cognitive Semantics". In *Meaning and Mental Representations*, Umberto Eco et al. Bloomington: Indiana University Press.

Laurel, B. 1993. *Computers as Theatre*. Addison-Wesley. p.161.

Lefebvre, H. 1991. *The Production of Space*. Oxford: Blackwell. p38-50.

Livingstone, D. 1997. 'The Sorcerers Apprentice'. In 'Mediaspace', contained within. *Digital Creativity*. Intellect Ltd.

Varela, Francisco, J. 1991. The Embodied Mind. Cambridge, Massachusetts: MIT Press. pp172,.232.

Virilio, P. 1994. *The Vision Machine*. London: BFI Publishing, p62.

Dan Livingstone is a Senior Lecturer in Interactive Media with the School of Computing at the University of Plymouth. His current role within the School is that of programme manager for the BSc MediaLab Arts course. He is a member of STAR (Centre for Science Technology and Arts Research) His current research activity deals with the design of interactive environments and augmented realities that investigate new forms of perception and consciousness. In 1996 he was an active contributor to the 'Mediaspace' satellite broadcasts, a series of interactive programmes subscribed to by a number of European universities. Other collaborative research activity includes a project with contemporary artist John Skinner and professional musician and composer Chas Dickie (of 'Startled Insects' fame). In 1995 a series of collaborative experiments were held at the Abbotsbury Studio in West Dorset, investigating the critical aspects of interactive performance and composition. Livingstone moved to Plymouth in 1994 after completing a Masters Degree in Sequential Design, specialising in Interactive Media, at the University of Brighton. During his MA studies he won the Rediffusion Award for Multimedia and was sponsored by the Rediffusion Simulation Research Centre, to continue his research into human facial expression. 'Darwins Expressions' was exhibited in September 1993.

Emergent Behaviour in the context of Reactive Compositional Environments

Dan Livingstone

Computer Music Research Group, Institute of Digital Art & Technology, School of Computing, University of Plymouth, Plymouth PL148AA England
d.livingstone@plymouth.ac.uk

Abstract. *This discussion paper identifies a number of fields of practice that consider emergent behaviour to be a key element in realising new creative forms. Creators across these fields manifest compositional processes, immersive environments, interface design & tracking systems, human computer interaction, interactive and generative music, collaborative soundscapes and are becoming increasingly engaged by the possibilities of emergent behaviour. The potential for new interdisciplinary forms integrating gesture capture, motion tracking, sound synthesis and collaborative forms between people/performers/composers and their environments is a developing field of research that investigates process driven collaboration to inform the design of reactive compositional spaces.*

1. Introduction

Process driven collaboration can be described as an embedded strategy that instigates a shared goal to stimulate interaction or participation, either between performers and their instruments, composers and sound or participants and technologically mediated experiences. Increasingly these embedded strategies can be found at the software layer of interactive or compositional systems, for example an Algorithmic approach extending the potential for both the generation of new sound relationships where the dynamics of the environment or performance are directly affected by participants of the system and the system is perceived to be responsive, indeed across a range of fields of practice this 'responsiveness' has been identified and extended, leading to a number of ways of describing emergent behavior. Where once we would have described interaction between users and systems with a clear hierarchy implicit in the language used, we now find these relationships have evolved, in part due to the increased use of embedded strategies to facilitate real time compositional processes in response to interaction. These forms of collaboration between participants and systems in many cases lead to new forms of behavior being realised as an extension of the creative potential of both parties, this language of behavior is playing a key role in the development of new interdisciplinary collaborative processes.

2. Resolution

As practitioners and researchers from differing fields exchange expertise and approaches new possibilities come into focus and a deeper understanding of the language of interdisciplinary work is reached, technologically mediated relationships can be very effective across a range of resolutions, for example an interaction as simple as moving or clicking a mouse forms the primary act of interaction most people have with computer technology, but clearly the resolution of this act is determined by the sophistication of the interpretation of the act in relation to context and intent, both on the part of the user and in terms of 'expectation' or 'anticipation' of the system or software design – ie, the resolution of the act is multiplied by the understanding of the range of anticipated or implied behavior, so any system that multiplies the resolution to extend the language of reciprocal engagement with a context or process embedded within the work has the potential to manifest emergent behavior. Koert van Mensvoort of the Eindhoven University of Technology has been developing an 'active cursor' method for simulating haptic feedback:

> *"The position of the cursor channel is normally used for input only. We developed a cursor interface in which the system manipulates the cursor position to give feedback to the user. The user still has main control over the cursor movements, but the system is allowed to apply tiny displacements to the cursor position. This system has a lot in common with existing force-feedback systems, except for the fact that in force-feedback systems the location of the cursor is manipulated as a result of the force sent to the haptic display, whereas in our system the cursor location is directly manipulated."*
> *[Koert van Mensvoort 2002]*

The key point here from the perspective of interdisciplinary practice is the increase in resolution of information possible from one human computer exchange - a well understood process driven act. As researchers in the field of HCI increase the possible range of reciprocal interaction with feedback processes simulating tactile sensations through visual stimulus, these methods can be added to the possible language of behaviors that can drawn on in the design of interactive environments.

> *"Once interface designers can count on its presence, haptic feedback can become a standard communication channel with the user. Our method was developed for use with standard mouse, but should work on any cursor-controlled interface. "*
> *[Koert van Mensvoort 2002]*

There is a clear potential here when we begin to describe mouse movement as gestures, as nonverbal language but there are also significant implications on how a system is programmed to react or 'learn' from this, establishing a process of collaboration or dialogue, we will consider work in this area later in this paper.

In many areas of practice direct manipulation of the media or processes inherent in a system is not a key requirement, the system or piece has been resolved and an increase in resolution of the reciprocal cycle is achieved by a number of means. A low-tech but nevertheless engaging approach can be seen in the recent collaboration between Sam Woolf and Tine Bech whose approach integrates sound sculpture within 'reactive' robots that appear to display a range of autonomous behavior. Simple analogue sensors and control circuits are used to extend the interface of a system to allow intuitive interaction to take place.

> *"Despite its simplicity, Echidna exhibited a large range of interesting sonic behaviors. This behavior reflects not the sophistication of the underlying electronics, but the complexity of the environment in which the sculpture is situated."*
> *[Woolf & Beck 2002]*

Their paper initially describes the sound sculptures themselves and goes on to 'defend the use of simple reactive robotics in interactive art' but they also make some significant observations not only on the apparent autonomous behavior within the systems but also between viewers who play a significant role in the process.

> *" we should not forget that humble reactive robotic systems capable of sensing and reflecting the complexity of their environments have the capacity for unpredictable and life like behavior that encourages playful somatic interaction."*
> *[Woolf & Beck 2002]*

It is an intriguing approach that leads to reflection not only on the emergent behavior manifested by the system but also how the reactive nature of the work instigates this process driven activity within the participants, a clear example of the dialogue or relationship that is formed is given and again it is only a matter of resolution to establish and articulate more complex compositional interactions with such a system.

> *"…despite the simplicity of its control circuitry, Boundless appears to display complex autonomous behavior. If approached by an observer it will attempt to withdraw, as if trying to flee from a perceived predator. If approached by several people from more than one side, Boundless jitters indecisively, as if unsure of which way to turn.*
> *[Woolf & Beck 2002]*

It is perfectly reasonable to counter this observation by suggesting that participants ascribe interpretation to perceived actions and react accordingly but if these non manipulatory modes of interaction are more clearly understood then the potential for sophisticated compositional and collaborative processes within reactive environments becomes a realistic proposition. Just as the designers of screen based interaction are developing subtle but sophisticated visual feedback systems to enhance immersion through representation of tactile, physical properties within a software environment, creators of computer mediated hybrid environments or cybrids are increasingly looking at gesture capture and motion tracking to enhance the sytems reactive properties to both participants and environmental parameters alike.

3. Behavioural Semantics

A sophisticated area of research that integrates both an evolutionary approach and compositional gestural interaction can be found in the work of Fels and Manzolli where the semantics of spatial relationships and biological cycles are integrated to provide a new compositional process, interaction is mapped between two participants and their gestural interaction influences the genetic make up of the compositional textures generated. Two approaches are discussed in detail in their paper ‘Interactive, Evolutionary Textured Sound Composition’ the second method uses direct tracking of two participants within a physical environment and they have successfully mapped performer presence and what could be described as compositionally driven semantic behavioral activities to provide a new form of compositional space.

> *“In the second technique the two objects are people. The position of the people are tracked using a local positioning system (LPS) developed in-house at the University of British Columbia. The LPS system uses infrared-based active badges and camera modules for tracking the position of moving objects. The idea behind using the interaction of two people to manipulate the genetic algorithm comes from thinking about the semantics of how two people interact with each other and their environment.”*
> *[Fels & Manzolli 2002]*

The emphasis on integrating the way we perceive and respond to spatial relationships in physical space as an extension of compositional process is another effective form of process driven collaboration, participants have a physical context for their interaction and a mental model of how interactions or movements through this environment in relation to the other performer will affect the music mediated by the system,

> *“…we have developed a system that allows a performer(s) to control an underlying evolutionary process which in turn creates music. We have encoded melodic structure as a genome and have defined a number of genetic operations that can be applied to a population of melodies. We have mapped some of the relationship semantics between two objects to control semantically related operations in the evolutionary cycle.”*
> *[Fels & Manzolli 2002]*

A range of disciplines are now using tracking of participants to add resolution and identify behavioral reactions whilst interacting with these systems and this will inevitably allow us to develop more responsive systems that facilitate forms of collaboration not only mediated by technology but with technological entities be they virtual or integrated into physical systems. In the area of interactive music there are many examples where composers and designers have extrapolated from modes of improvisation and collaborative processes to extend compositional possibilities, a key area of study for those of us engaged with reactive cybrid sound environments is again at the behavioral level where either we need to anticipate forms of behavior for

our systems to react to and learn from or we need our systems to facilitate responses or reactions that in turn lead to forms of emergent behavior.

4. Strategies for Participation

When designing interactive sound environments or systems for public spaces or for direct participation, an understanding of the forms of dialogue that are engendered by such systems and those interacting with and through them is a significant factor, again the development of these works is adding to the language of interaction in productive ways. In a paper discussing the interactive music system 'Tonetable' Bowers underlines some of the successful outcomes and implications of this approach, it is clear that the Author's strategy for the work is to engage the public collaboratively and also that the work is influenced sensorily by participants activity. The system is table top mounted with four control wheels and participants are invited to manipulate objects within the space which is simultaneously diffused as sound around them, in some ways this work is in the same interaction genre as Toshio Iwai's 'Resonance of 4' installation which has been successfully exhibited at a number of international locations, where four participants interact with a sequencer like shared grid via mouse interaction, in Bowers work a number of strategies for understanding the way people interact with and collaborate through the system are discussed.

> *"ToneTable manifests a variety of sonic and graphical behaviours which can be progressively revealed through engagement (both individually and collectively) with it. This can give a 'structure of motivation' to its use. That is, we intended to provide an 'in-built' incentive to explore the table and its varied behaviours and image-sound relations. Indeed, in detail, the dynamical behaviours of ToneTable were defined and calibrated with various non-linearities."*
> *[Bowers, J. 2001]*

This notion of structure of motivation and incentive to explore allows participants to explore sound image relationships but also invites a range of behaviours or actions from participants, resolution is multiplied not only be the number of interactors manipulating the system via the visual feedback but tacit compositional agreements or shared journeys are embarked upon as participants actively listen to the output.

> *"interruption in object-behaviour is intended to add interest to the graphics as well as being an outcome that is easier to achieve through concerted collaborative activity between participants. Thus, the threshold for the occurrence of orbiting behaviour is set so that it will tend to be exceeded by a local force produced by two or more proximal wavefronts. That is, two or more participants need to align their perturbations of the surface to produce the orbiting effect."*
> *[Bowers, J. 2001]*

Bowers also reflects on the range of strategies explored in the development of this approach, this articulation can be considered both in terms of the design of interaction and in terms of collaborative compositional processes but is clearly worth further exploration and definition to inform the design of such systems.

> *"we have tried a number of design strategies for addressing such settings. We have explored notions of 'collaboration through a virtual medium', 'collaborative added value', 'layers of noticeability', 'structures of motivation'. These are all concepts intended to suggest ways for orienting design for variable participation."*
> *[Bowers, J. 2001]*

5. Conclusion

The approaches I have discussed all have significant contributions to offer to the area of research I am engaged in; Reactive Compositional Cybrid Environments, I am currently experimenting with a portable system that I have developed. This system comprises original software developed with Max/MSP/Jitter running on a G4 Apple Laptop, the software 'listens' to the chosen environment through audio analysis via MSP while analogue sensors capture live interactions that inform compositional decisions initiated by the software, the compositional process is mediated by the physical or acoustic properties of the space and the presence or interaction of participants, real world data is integrated in the synthesis process of the system. An external Yamaha rack Synthesiser allows Formant shaping and FM synthesis and is also controlled by the software and reduces CPU overhead. The system uses an Icube for general data collection via midi, an additional midi input is available for other interaction or control surfaces to be integrated while composing or improvising with the system. Gesture capture and positional data is currently facilitated by video input, two Digital Video cameras are used to correlate simple 3 dimensional positional data, for example the orientation of a gesture can be related to a specific spatialised sound output. The software also mediates the compositional output and co-ordinates the eight channel sound diffusion in real time; sounds can be positioned and moved throughout the environment in response to the original compositional framework, which subsequently evolves through live performance and interaction.

A key goal in the development of this approach is to enable and record the emergent behavior that occurs between software, people and live spaces as an integral part of the compositional process. Future areas for further investigation include a more detailed analysis of reactive or responsive compositional spaces, observation of emergent behavior to inform design of interface elements and listener or composer objects and field testing of the system including interfacing with live data from a building management system.

References

Ir. Mfa. Koert van Mensvoort
What you see is what you feel -Exploiting the dominance of the visual over the haptic domain to simulate force-feedback with cursor displacements P2,5.
*Proceedings Designing Interactive Systems 200*2, ACM Press, 345-348.

Woolf S. & Beck T.
Experiments with Reactive Robotic Sound Sculptures. P2, 3.
ALife VIII: Workshop proceedings 2002

Fels, S. & Manzolli J.
Interactive, Evolutionary Textured Sound Composition P7, 9
EGMM01: EG Multimedia Workshop 2001

Bowers, J.
TONETABLE: A Multi-User, Mixed-Media, Interactive Installation P2,4,5.
Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01), Limerick, Ireland, December 6-8, 2001

Composition for Ubiquitous Responsive Sound Environments

Dan Livingstone and Eduardo Miranda

Computer Music Research

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth
PL148AA United Kingdom
dlivingstone@plymouth.ac.uk eduardo.miranda@plymouth.ac.uk

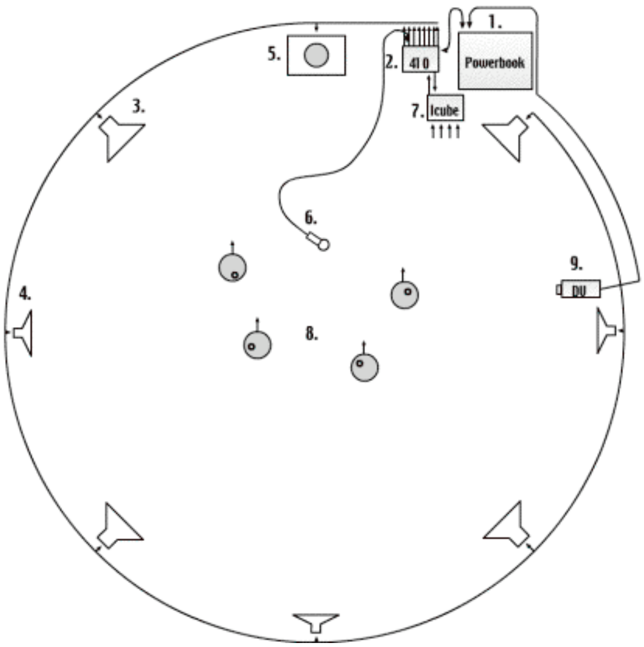
Abstract

As new forms of social interaction with sound are developed through hardware, software and ubiquitous technologies it follows that emergent behavior, gesture capture and motion tracking will increasingly play a compositional role within generative and reactive sound environments. This case study defines an adaptive system, which enables participants within these spaces to have a tangible influence on the compositional process. Both individual and collaborative interaction modes are considered in the context of generative and real time systems, which are dynamically affected by user presence.

1 Introduction

Interactive music systems are often designed to provide engaging gestural control, enable new forms of musical expression, and are generally accepted to include three classes of compositional algorithms; sequencing, generation and transformation [Rowe 1993]. This case study establishes an integrative model for process driven collaboration [Livingstone 2003] within responsive compositional environments, by detailing the flow of interaction between composer/participants, a responsive sound environment and an adaptive compositional process. The system regenerates a soundscape dynamically by mapping 'known' gestures to influence diffusion and spatialization of sound objects created from evolving data, degrees of control are determined by clarity and scale of gesture, and the system is designed to adapt to these interactions by initial 'call and response' feedback within the structure of the composition. This is seen as a beneficial extension of the performer/performer relationship [Lippe 2002]. The structure of each 'response' is stored in memory and compared to previous instances via mapping of the properties of each sound event ie; gestural trigger, related to concurrent sound object and its diffusion properties. This sequence of properties that encapsulate or encode each response instance can be visualized using additive synthesis as a model, each element (timing, diffusion, gesture mapping, synthesis and concurrent sound objects can be stored in a series of envelopes or partials that codify a discrete sound event. The interplay of sound events evolves as a dialogue is established between users and the system.
The sound objects themselves are designed to initiate this dialogue, as 'psychoacoustic triggers' to interaction and are transformed through interaction. The system is given a collection of 'gestures' or patterns and building blocks for a range of sounds. Initial 'response' patterns are based on relating intentional gesture/movement to sound diffusion, and in(attention)al [Mack & Rock 2000] gesture/movement to re-synthesis of sound objects in memory. Compositional parameters are designed through continued observation of interaction through the vision system combined with live environmental sensor data. The system includes listener objects which can 'live sample' ambient sound material or intentional sound input with environmental data, enabling an adaptive approach for capturing sound events stored in short term memory as data parameters only; the system transforms this combined data and live audio which includes the acoustic properties of the physical environment to new compositional material. The integrative model discussed has been prototyped on a small scale using an apple laptop, icube, m-audio firewire410, 7.1 sound output, environmental sensors input through I-cube midi interface and video tracking for gesture capture. Software has been written in cycling74's version of Max/MSP/Jitter for OSX. This system (fig1 responsive system) provides an intuitive interface where simple gestures are used to interact with a 'real world composition' [Costa, Manzolli, Verschure 2003]

that is; an adaptive sound environment that is designed to be responsive to new interactions, as opposed to a pre – defined reactive systems approach.



(Fig1 responsive system)

1. G4 Apple laptop
2. M-audio Firewire 410 audio/midi interface
3. Set of eight active speakers positioned in either 7.1 surround format or two tiers simulating multiple height speaker system in atria, Portland Square.
4. Indicates offset speaker location of 2nd higher tier of four speakers
5. Additional active sub bass speaker
6. Microphone (AKG CB300) for live sampling
7. Infusion systems Icube interface.
8. Wireless 'composer & listener objects – clusters of sensors provide either environmental data or direct interaction) – can also use Bluetooth enabled mobile phone for diffusion control.

The system illustrated provides an adaptive framework that can be used for both specific compositional installations or for controlled experiments in sound perception/interaction/reaction to develop new compositional methods.

2 Interaction

Primary interaction with the system is via gesture capture using a real time feed through two fixed cameras, one overhead for general movement and orientation relative to physical space and one in front of the user localized to capture left or right hand movement. The vision system data is captured with Cycling74's Jitter software using matrix objects to track, map and compare a limited palette of symbols, *circle*, *line*, *triangle*, *square* and *cross* to user gestures drawn in the air with a finger/hand. Resolution of this action over time is increased using CV.jit externals by Jean-Marc Pelletier (http://www.iamas.ac.jp/~jovan02/cv/) to track depth, speed and direction of gesture, the base information or algorithm for each symbol is stored in a matrix so the sequence of numbers that correlates to a 'known' gesture can have levels of accuracy and relative scale, allowing subtle variations in captured gestures to be identified. Max **mtr** (*multi track sequencer*) **capture** (*stores number streams*) and **env** (*Script-configurable envelope editor*) objects are used to store and compare these sequences with real time input from the vision system. This approach creates short term memory, allowing the system to match 'known patterns' (reactive system) but also to identify repeated unknown patterns which can then be added to long term memory as new symbols (responsive system), for example drawing an 's' several times will add this as a new symbol, previously unknown. This gestural composition process of interaction enables small scale gesture (individual – see fig 2) but can also be mapped to larger scale (group –see fig 4) behavior, for example social groups viewed from above can intentionally recreate symbols collaboratively by forming patterns or 'known' symbols tracked with similar methods as an individual hand gesture, this is achieved by designing an adaptive composition system and applying a methodology that allows for adaptive resolution, this approach can be considered as diachronic emergentism as consideration of the acoustic perception of sound objects is a key factor in sustaining the collaborative real time compositional process through effective sound design and spatialization to influence participants behavior and establish a musical 'dialogue'. Both the system and users are sensitive to the environmental properties of the composition environment; the localized portable version (fig 1) includes sensors for ambient temperature, light, and air movement, 7.1 sound diffusion, gesture capture and 'composer objects'. (A large-scale implementation of this system using additional data from a building management system with extended social interaction is being developed for field-testing in the Portland Square building, University of Plymouth. UK. http://www.arch-os.com/)

2.1 Conscious interaction

During individual interaction a parallel is drawn between recognized/non-recognized data combinations and intended/unintended actions recognized data is mapped to specific processes, drawing a *circle* will trigger a rotary pan with speed and direction, straight *lines* give panning settings, a *triangle* will create a new envelope for current sound object, a *square* defines a measure of time and a *cross* fades current sound object/s. Of course superficially this symbol based approach leads to a limited palette of compositional possibilities, one must consider that this limited palette provides a clear framework for interaction in line with the 'call-response' model which also enables the software design to factor user reaction to current real time outputs. An agent based sound object can pan itself relative to user position, so any recognized intended action is subject to current system and environment parameters this approach can be described as multiple low resolution events combined to provide a more sophisticated higher resolution 'world view'

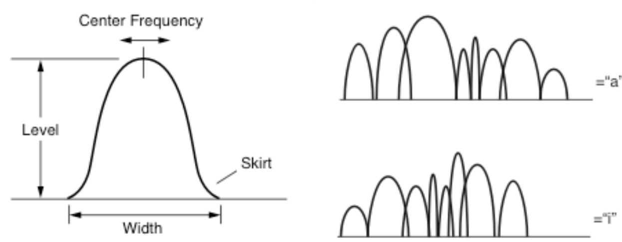
(*fig2 small scale interaction - Gesture is tracked in relation to live sampled sound (sonograph) in this instance a new symbol 's' is added to short term memory creating a new envelope from data associated with this pattern or number sequence for diffusing a sound object in real time)*

enabling the system to interject compositionally significant events [Camurri 2000] Composer objects; wireless interfaces capturing localized environmental data can also be manipulated by users to gain a higher level of influence over sound object creation/design. Both custom built sensor clusters and blue tooth mobile phone interaction are being prototyped, initial experiments show two distinctly different modes of intended interaction, users manipulating composer objects seek *dynamic control* over sound composition whereas mobile users have *positional control* of one element of the soundscape in relation to other participants for example. These modes of *intended* interaction all have the potential to establish a compositional dialogue with the system.

2.2 Subconscious interaction

When the system is not tracking known symbols or direct influence from composer objects, listening agents are used to mediate subconscious interaction. An example of subconscious interaction can be seen in figure three, an overhead motion tracking camera detects one slow moving and two fast moving social groups, connecting these groups relative co-ordinates creates a *triangle* the compositional system may recognize this or can be taught this variation based on the equilateral triangle pattern stored in long term memory, in this compositional system the *triangle* shape influences the envelope driving the Formant synthesis element of a current sound object A *square* symbol influences timing and is identified in relation to the overall tracked area this can be applied to the current envelope or if a specified time has passed since the last identified symbol it is applied to current sound diffusion patterns. For example an '*ah*' shaped formant is initiated on recognizing a *triangle*, environmental parameters excite resonators or anti resonators allowing the '*ah*' to shift from vocal to nasal, if a *square* is formed timing values are adapted, a *circle* symbol morphs the formant to an *'ai'* shape or other vowel shape by shifting the center frequency or manipulating the envelope generators through time based timbral shifts. A Yamaha FS1R rack synthesizer is used for real time control of Formant and FM synthesis, the system use a software interface created in MAX/MSP to control it. Of course users of the system may become attentive of this 'subconscious' process and choose to *intentionally* change the current sonic structure, collectively forming a uniform *square* for a longer duration to extend the temporal properties of a current sound object. Alternatively if no symbols are recognized live data from the composer/listener objects or selected parallel data from the building management system is used to reform base sound objects through FM synthesis.

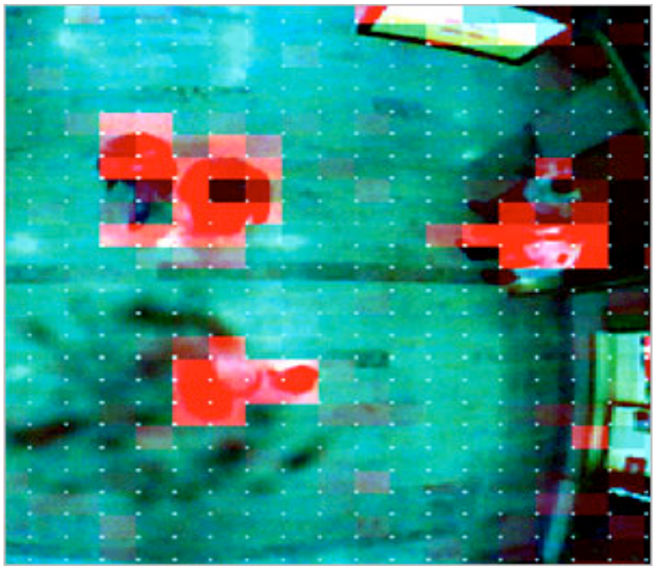
(Fig 3 basic formant shaping)



environmental data changes are usually slow in interior environments so these elements are mapped to the timbre and color of sounds created, providing an overall structure for the real time composition that is responsive to either the ambient light, temperature and air movement (test system) or to selected data real world data from the building management system.

2.3 Resolution

The portable system illustrated (fig1) is currently located in a small office the Portland Square Building, so while local gesture capture and synthesis methods are being tested and refined (small scale interaction) the full building management data and camera streams are monitored, allowing comparative analysis and continued prototyping for the real world system. The concept of 'resolution' has been a valuable tool in developing methods for tracking and synthesis that can be 'transposed' to a complex social environment with embedded ubiquitous technologies.



(Fig 4 large-scale interaction, Arch OS vision system, Portland Square – example of how subconscious interaction can be given meaningful compositional attributes)

3 Composition

The compositional approach is one of continued observation and refinement of the interaction process; the system has been given a subtle 'voice' through basic formants [Styger, Keller 1994], which are combined with more spatially specific fm synthesized sound objects. Output from the vision system tracking either small-scale gestures or large-scale movement is monitored to identify possible musical relationships that can be used to influence behavior. This is an approach of considered sound design which produces effective base sound material that is not overly complex, the system, site and participants influence these base materials either directly or indirectly to form new sonic structures that reflect the movement and physical properties of the compositional environment. This approach can be considered as an adaptive variation of spatial music. It is also a learning process, a number of tasks have been developed to field test these compositional processes.

4 Conclusion

A prototype responsive system has been developed which integrates a range of computer music techniques to provide a compositional approach to generative or interactive music. In this paper we described one prototype example/system … key findings are that through a person centered design approach to interaction on an intimate level (hand gestures) that in principle an effective compositional dialogue can be established; novel interactions can be recognized and 'learnt' both by the system and those interacting with it. A strategy for deployment of this system on a larger scale, using the atria connecting offices and teaching spaces, has been outlined for Portland Square building, University of Plymouth, UK.

Interaction design for larger social groups is being refined based on initial observation of motion of people through these spaces. (Fig 4) Ongoing research seeks to identify methods of recognizing social 'intended' interaction with the system by field-testing and refining the symbol based compositional process discussed (2.1). This research establishes process driven collaboration as a compositional methodology. Future applications could include the capture of 'perceptual constructs'[1] [Livingstone 1998] as sound signatures of those participating with these systems. The potential for large scale adaptive games systems integrating personal mobile technologies with large scale social 'learning' environments offers significant potential for interdisciplinary research.

References

Camurri, A. (2000) "Artificial Intelligence Architectures for composition and performance environment". *In Miranda, E. R. Readings in Music and Artificial Intelligence (Contemporary Music Studies)* Routeledge pp.163-188

Costa, Manzolli, Vershure (2003) "Constructing Sonic Structures with Visual Roboser". *In Proceedings of the IX Brazilian Symposium on Computer Music, SBC* pp131-136

Lippe, C (2002) "Real time Interaction Among Composers, Performers, and Computer Systems" *Information Processing Society of Japan SIG Notes, Volume 2002, Number* 123, pp. 1-6.

Livingstone D. (2003) "Emergent Behavior in the context of Reactive Compositional Environments". *In Proceedings of the IX Brazilian Symposium on Computer Music, SBC* pp235-240

Livingstone D. (1998) "The Space between the Assumed Real and The Digital Virtual". *In Ascott, R. Reframing Consciousness – art, mind & technology Intellect Books* pp138-143

Mack, A. Rock, R. (2000) *Inattentional Blindness* (Cognitive Psychology) MIT Press pp.215-255

Rowe, R (1993) *Interactive Music Systems: Machine Listening and Composing,* Cambridge, The MIT Press.

Styger, T. & Keller, E. (1994). "Formant Synthesis". *In E. Keller (ed.), Fundamentals of Speech Synthesis and Speech Recognition: Basic Concepts, State of the Art, and Future Challenges* Chichester: John Wiley. pp. 109-128

[1] *A perceptual construct in this context is a series of spatial relationships created by an individual manipulating sound objects as reference points in relation to physical space using the mental model of a wire frame cube to enhance spatial perception of sound objects, positional coordinates for each diffused sound are captured and compared with the users description or individual mental model of 'their' cube in relation to the physical environment and cubes of other people.*

ORB3 – MUSICAL ROBOTS WITHIN AN ADAPTIVE SOCIAL COMPOSITION SYSTEM

Dan Livingstone
Computer Music Research

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth PL148AA United Kingdom
dlivingstone@plymouth.ac.uk

Eduardo Miranda
Computer Music Research

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth PL148AA United Kingdom
eduardo.miranda@plymouth.ac.uk

ABSTRACT

Gesture capture, motion tracking and 3D visualisation technologies have generated many new musical forms, often extending the mannerisms or behaviours of a given performer or discipline, providing new compositional frameworks for real time synthesis in response to action. In many cases these approaches are presented within a single domain, a live stage performance, a site specific installation, a shared networked visualisation of collaborative composition. The reality is that these 'interactivating spaces' [1] whether haptic, [5] tactile [9] or ubiquitous [11] is that they manifest new forms of interaction, between people, systems and the medium of sound.

Free Sound can be understood to be an extension of the 'open work' where the base materials for a compositional process are created through a model of exchange, interaction and resynthesis. The resulting output of these activities can be broadcast and disseminated through a range of technologies to both social and private spaces. This research suggests that there are new interaction models and social compositional frameworks to be found in these cybrid spaces, a previously intangible location often dominated by the broadcast and publishing industry. A marketing model defined by revenue streams and a value chain. In the case of socially mediated composition or 'free sound' there is still a value chain, it's investors and beneficiaries are the open source community, the collaborators and participants within such mediated systems and the resulting free sound.

Keywords

Adaptive System, Sound Installation, Smart Interfaces, Music Robots, Spatial Music, Conscious Subconscious Interaction, Interaction models.

1. INTRODUCTION

In the design of new interfacing methods [1] for sound manipulation and control it is often the case that the primary focus is the point of tactile interaction, the exploration of new gestural controllers or methods for mapping and transforming data to create sound material [2]. This approach has led to the development of numerous novel and individual interfaces [3], in many cases the interaction mode is learnt by the user, in order to complete the feedback loop, thereby achieving dynamic results through an exploratory model of interaction.

With a modular adaptive systems approach the emphasis is on providing an interface framework for different types of interaction that can be initiated by both users and 'smart' interfaces, ie new interaction behaviors can be identified by the system independently, in response to users actions, whether direct tactile control or simple movement, location, gesture or position. Figure 1 illustrates the ORB3 system, an example of an adaptive systems approach.

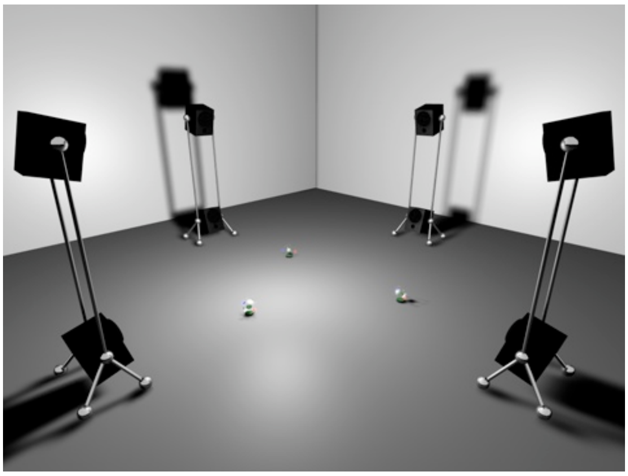


Figure 1. *Auditory Sphere. 8 active speakers angled to provide versatile software controlled diffusion. Diffusion and synthesis generated from environment/interaction data collected by each Orb. Software developed in MAX/MSP running on G4 Apple laptop with M-Audio 410 Firewire mobile multi - channel interface, custom built 'composer – listener' objects (wireless Orb3 interfaces)*

Orb3 is a compositional space that can operate autonomously through each interfaces interaction with its' environment, each Orb collects the base materials for sound composition, participants moving between Orbs influence their motion and data collection behaviours, Orbs can also be picked up and held as a controller for the sound material they have generated, each Orb has the potential, through software, to adapt to its current user, a number of models have been developed to motivate this interaction, based on observation of interaction with a number of novel interfaces and custom controllers:

"…to have a musical response accentuated by the player who sent the original call, to plant a musical "seed" that would be picked up by the group in various manners, etc." Weinburg [9] A significant observation during performances by children using the Beatbug system developed at MIT was described at the International Computer Music Conference in Miami 2004; the children made exaggerated swooping motions with the Beatbugs as they 'passed' sounds while interacting with the controller. At the time the Beatbugs were not equipped to react to this *emergent behavior*, although neither the audience or children were aware of this at the time. This observation led to the next refinement of this approach to a network of interfaces, potentially using blue tooth technologies and motion detection to refine and utilize this interaction, this anecdote reinforces the value of the adaptive systems approach we present which is a continually evolving field of applied research for novel interfaces and interactive music systems.

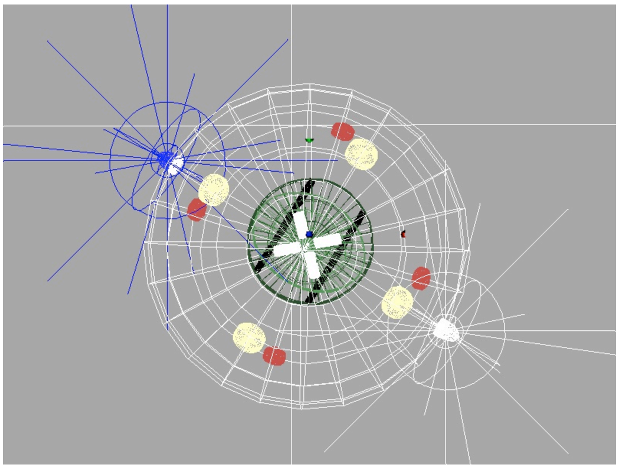


Figure 2. *Orb3 Design Constraints.*
Design aesthetic can be achieved by designing custom PCB's for sensor placement, careful selection of plastics for manufacture and careful consideration of tactile properties for intuitive interaction.

2. ORB3 DEVELOPMENT PROCESS

The Orb3 interface design was developed through observation of interaction with wired 'composer and listener' objects[i]. These original objects were static spheres housing a cluster of analogue sensors (Light dependant resistors, bead thermistors, vibration and tilt switches etc) for measuring ambient light, ambient heat, general motion and orientation. The original system comprised four such spheres which could be placed and relocated to generate and vary data used to synthesize sound material for 7.1 sound diffusion controlled by a gesture and motion based video tracking system[ii]

Each sphere contained a total of 8 sensors, wired to a control voltage to midi converter (Infusion Systems Icube) this method worked effectively for developing software and refining synthesis and sound design for prototyping a large-scale adaptive system. Local interaction was less successful due to restricted movement of wired objects and unexpected behaviors and reactions of participants. For example; using the prototype system the shadows cast between spheres as participants moved around the room were recorded by a drop in light values sensed by the sphere affected, causing subtle changes in base sound materials generated for the sound-scape, this was an intended compositional element of the system but on realising this process, many participants could not resist the temptation to explore further, initially cupping or shielding areas of the spheres and inevitably moving and repositioning them, anticipating a direct response. It was immediately apparent that the simplicity of the sphere encouraged a series of interactions that could further inform sound design for socially mediated sound spaces. It also led to the realization that the software techniques applied to the vision system for adding new data relationships based on symbol recognition could be migrated to the interface design for each sphere developing more expressive tactile control, and more significantly, using the relative position and orientation of each sphere as a compositional parameter that could be heard in the diffused sound-scape, that was also registered by visual or tactile feedback on the interface itself. Other observations were that often participants chose to work collaboratively, taking a sphere each, influencing a parameter passing it on, this worked particularly effectively in groups of three, where patterns of motion and exchange had the potential to create rhythm and flow, some general experiments were done with different numbers of spheres to see if this affected interaction modes, it is speculated that providing an odd number of interfaces provides more movement through transfer and exchange and encourages turn taking. It was also noted that during periods of inaction or when participants were more passive different listening modes were reported, this in turn has influenced the sound design of the refined system, incorporating different 'play states' or modes - some further controlled experiments are in data to support these assertions. The logical development of these passive and active modes mediated by participants is to add simple robotics to each sphere to allow each one to move and interact with other spheres independently.

2.1. Design for Collaboration

Having established some significant refinements from the initial prototypes a specification for a more robust adaptive interface was resolved. Primarily a wireless approach was required, high performance with reliable transfer of digital and analogue data from sensors, in addition a wireless microphone embedded in each unit for live sampling. Internal lighting was added to indicate interaction modes and force feedback in response to interaction. These features introduced new design challenges, as the revised design needed onboard power for wireless operation and ideally solar charging to extend session times. A final addition was the inclusion of lasers and proximity sensing to enable quick alignment and event triggering between spheres. A simple method for overhead positional video tracking

(max/msp Jitter) using a single fixed camera provides an effective method for documenting movement and behavior of each orb during a live session through time lapse imagining.
A mobile robotic element has been prototyped for each sphere, allowing them to move and reposition themselves autonomously or in 'collaboration' by integrating positional tracking (digital compass/distance and trajectory) and proximity triggers (ultrasound sensors). This dynamic motion provides a visual element that reveals the compositional potential of the system, while demonstrating some of the synthesis and diffusion properties that are influenced by the interaction between or with each sphere. When each Orb is collecting data to influence sound synthesis and diffusion, or being followed or manipulated by participants this collaborative process can be displayed from a top down perspective, using either projection or plasma screen display.

"Most of the systems that allow the creation of sound and image in real-time don't have the capability for organizing events at a global level. This is however, required if the aim is to allow the composition of a piece that involves feedback from events sonic and visual, in the construction of interactive audiovisual compositions." [3] (Franco et al 2004)

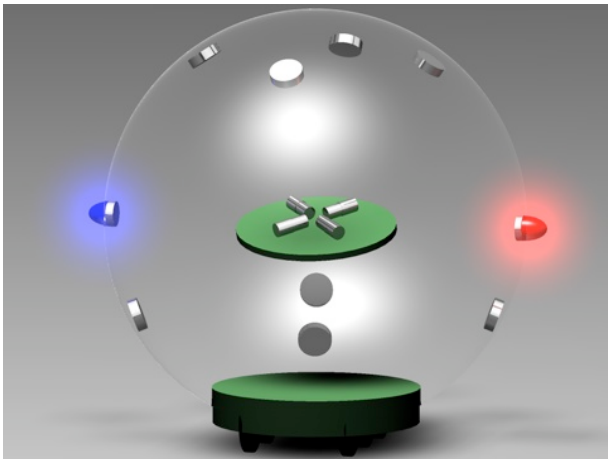


Figure 3. *Wireless Mobile Orb v2.0 in Absorb mode.*

(OrbV2.5 features data collection through light, temperature, orientation, motion sensing, laser alignment, microphone, mobility, rotation, Exploratory interaction transforming data into sound material)

For the purpose of this paper emphasis has been placed on the Orb3 interfaces, the key features are interaction modes and social composition, simply expressed as 'play states'. Sound is the primary medium but in order to make visible the transformative processes underpinning the compositional output ways are being sought to create a visual aesthetic from both the data and interaction of people, making visible behavior and interactions, effectively creating graphical transcription as real time feedback to participants.

Developers of collaborative musical interfaces with tactile, graphical and sensory feed back are developing new terminologies to describe the design process for these systems in terms that begin to articulate their compositional and social modalities. Collaborative interactive music systems, such as 'Block Jam' (Newton-Dunn, Nakano, Gibson 2002) where interconnecting blocks are collaboratively assembled to organise musical phrases and sequences begin to identify new musical forms enabling participants to create *'meaningful musical structures' through 'collaboration and exploration'* [6] (Newton Dunn et. al. 2002)
Other collaborative works such as ToneTable [2] (Bowers J. 2001) use interactive visual elements as an integral interface element, in this case participants manipulate 4 trackballs, 'disturbing' a projected fluid surface with associated textures and diffused sounds, again it is the observation of improvisation and collaboration with a real-time composition system that distinguishes this emerging musical form. The author discusses emergent behaviors and extended engagement as a development of the system design; *'structures of motivation' 'variable participation'* [2] (Bowers, J. 2001) A highly refined table top tactile control surface for two or more participants has been developed by Patten and Brecht, 'Audiopad' [8], which has been extensively exhibited. The system provides a graphically dynamic projected overlay oriented around electronically tagged tracked physical objects or 'pucks' for real-time control of preprogrammed electronic music, moved by hand with fingertip control. Key elements in terms of a compositional model are *'spontaneous reinterpretation'* and a combination of *'visual and tactile dialogue'[8](Patten J. Brecht B. 2003)*
The design and installation for the Orb3 system forms an auditory sphere (fig. 1.) using an 8 Channel sound diffusion through which participants move, view, listen and reconstruct the compositional process through social interaction within it. The audiovisual feedback in response to these varied interaction modalities is an active process, one of content driven collaboration.[iii]

2.2. Communications & Parameters

Each Orb sends data via a 2.4ghz wireless RF interface to a G4 laptop running Max/MSP, a combination of analogue and digital data can be sent and processed by the control software created in max. The software itself is not simply a parameter mapping utillity, it is designed to correlate different data against previous interactions, a form of compositional memory where environmental parameters of previous sessions are compared with current ones to identify repeated behaviors of the system and actions of participants. The software is designed to be adaptive, previously un-recorded or new data configurations are identified and used to compose new sound events or objects. The software sends data to each Orb to indicate it's state and trigger visual or tactile feedback, ie; activate laser/proximity sensing for positioning, activate status leds, activate force feedback.

Each Orb has two compositional states – *Absorb* and *Adapt*. In *Absorb* mode an Orb is autonomous and located on the floor, it's sensors are calibrated to collect environmental data, ambient light, ambient temperature, relative position and orientation, it can also live sample sound for processing - the software controls this calibration which is activated through Orb alignment – each Orb is fitted with proximity sensor, a laser and LDR - placing the three Orbs in a triangle and directing each Laser to the next Orbs locating LDR activates this mode, which is part of the initial setup process.

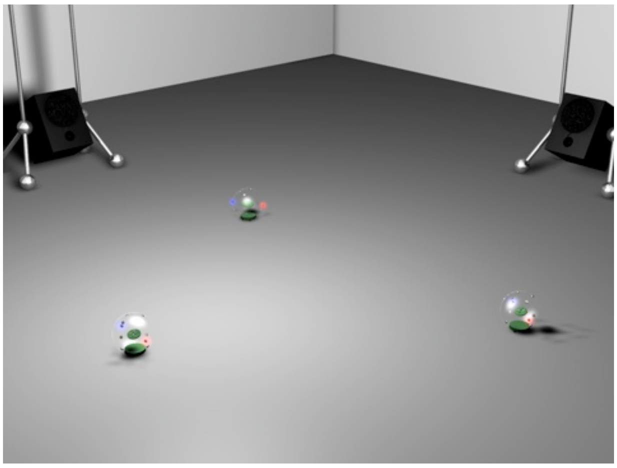


Figure 4. *Laser alignment - Triangulation.*

(Alignment –view of each Orb, lasers are activated, two Orbs are shown in listening 'Absorb' mode after calibration, one (lower right) is about to move out of alignment in response to parameter changes, autonomous – may move or rotate to attract participants, system responds by panning a sound object in relation to Orb location).

Environmental data changes are usually slow in interior environments so these elements are mapped to the timbre and color of sounds created with larger fluctuations affecting diffusion, thus providing an overall structure for the real time composition that is responsive to ambient light, temperature and general movement. *Adapt* mode is activated when the alignment of Orbs is disturbed, either by walking between them, interrupting the laser tracking or by picking them up which also activates vibration sensors and initiates orientation mapping - angle and orientation of each Orb in this state directly influences panning and diffusion rates of synthesized sounds. During *Adapt* mode the laser is deactivated and the ambient temperature measurement is recalibrated to respond to body heat through hand contacts on the Orbs lower surface. Bead thermistors with fast response times are used so as an Orb is passed from one hand to another, or between users, it registers and marks these changes. This data combined with orientation data allows for a range of subtle and dynamic sound events to be initiated by each participant in collaboration with both the system and with other people.

2.3. Emergent Behavior

As an adaptive portable system, the Orb3 environment creates an opportunity for observing and recording forms of emergent behavior in relation to spatial sound interaction, this provides researchers in this field with a structured framework to inform the design of mobile and autonomous interfaces, such as musical robots or adaptive social composition systems.
" we should not forget that humble reactive robotic systems capable of sensing and reflecting the complexity of their environments have the capacity for unpredictable and life like behavior that encourages playful somatic interaction." [10] (Woolf & Beck 2002)

The inclusion of play through collaboration is not a by product of this system, it has been developed explicitly to motivate different responses through consideration of ergonomics and human factors, developing from the considered observations of researchers and practitioners in related fields. The 'play states' or modes titled *Absorb* and *Adapt* have been designed with consideration of both composer/listener object interaction and the listening process or perceptual triggers to motivate participants.
In the 'play state' *Absorb* the Orbs are programmed to activate when certain parameters or sound events are captured, or when conditions match previously encountered sequences, the 'intention to listen' is shown through both the status LEDS and motion/rotation in response to stimulus. This modality can also be attributed to the behavior of participants, who move towards the 'Auditory sphere' of course initially their interest is more likely to be the spectacle of the technology or other participants behavior, however moving into the 'Auditory sphere' shows an intention to participate, *to listen.* Participants interaction at this stage can be described as Subconscious, they are not necessarily aware that their presence and orientation is influencing the system.

The 'play state' *Adapt* is active when the triangular alignment of the three Orbs is disturbed, the software reconfigures itself to a more sensitive state, ready to be interacted with, held, passed, moved in relation to sound synthesis & diffusion as perceived and manipulated by a participant. Applying a different interaction model to the parameter mapping of sensors is an effective way to initiate ORB states. In software terms this is achieved by switching the algorithms mediating data analysis, through patterns stored in short and long term memory (Max objects *capture, decode, funbuff, histo* and *spray* are integrated with *mtr* to record, store and replay streams of data, which are compared against previously collected and live data [short term memory], a form of score following). By picking up an Orb a participant is moving from the *Absorb* state, instead choosing to interact, to explore and through this action perceiving and identifying the source of broadcast sounds, through their manipulation of an Orb. This modality is further reinforced when direct control of sounds are influenced by the participant. Their behavior changes as they *Adapt* to the parameters they have influence over. This can be described as Conscious interaction, a heightened state of attention and engagement, [6] [Newton Dunn et. al. 2002] the intention to collaborate with the system and others using it, improvisation, not simply 'call and response' [4] [Lippe C. 2002] as there are no familiar, formal or structured elements in the form of musical patterns, note sequences or beats inherent in the open nature of this spatial sound environment. A key development with this system is that it continues to adapt while capturing, archiving and broadcasting new behaviors. A range of technologies have been explored a custom PCB with optocouplers was implemented to take multiple outputs from a midi to control voltage converter to send motion and trajectory motor control to each Orb using modified consumer radio control vehicle parts, modification to gearing, a new steerage mechanism and a form of pulse width modulation to control Orb trajectories provides more subtle movement. Multiple overhead color tracking has been tested, a fine balance between ORB speeds and processing of multiple tracked colours (LED combinations as figure 5.) in variable conditions for effective mediation was required. Several task-based experiments using perceptual constructs [12] to establish and refine interaction models have been used to refine compositional processes within the system. Interaction models have been tested by assigning participants simple compositional tasks based on establishing mental models for the relationship between sound objects and their perceived location in the auditory sphere. Quantitative data can be extracted from the short term memory of the system, noting the start event of an interaction, such as repositioning a sound by manipulating an orb. After this action has been archived Qualitative data from participants' reports can be established by comparing transcribed verbal accounts of the set task and system response. Both data types can be considered in context by reviewing the overhead broadcast documented by the system, archiving actual position of participants and orbs against archived positional data.

3. CONCLUSIONS

The compositional approach is not modeled on a *'fixed or even consistent excitation-sonification relationship'* [Paine G. 2004] many elements of the sound-scape generated are through transformative synthesis methods, in this instance the creation of sound through traditionally unrelated real world variables. Neither are the sounds randomly generated; the capture and transformation of variables such as heat, light, proximity, motion and time create values that could be mapped to conventional parameters for musical control of predefined note sequences, loops and formally structured phrases but in this adaptive approach through a process of observation, listening and sound design these parameters are treated as explicit elements of the real time composition environment, ie the ORBs are designed to be responsive to their physical environment, with adaptive behaviours that motivate human interaction.

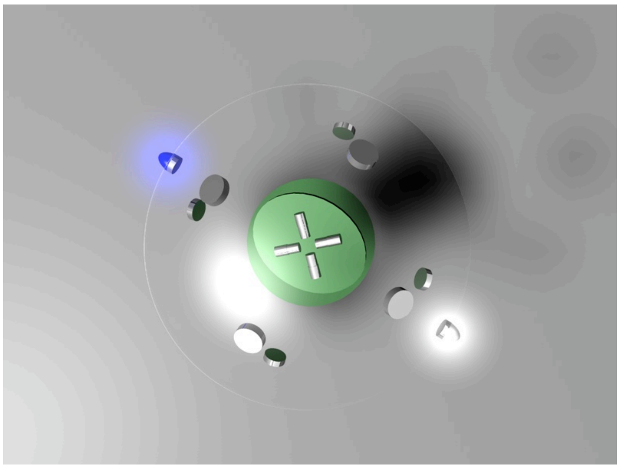


Figure 5. *Orb in adapt mode.*

(Overhead camera view (simulated), Showing a single Orb As an Orb is picked up sensitivity is heightened by re mapping parameters, accelerometer sensitivity maps motion, angle, orientation to sound diffusion while archiving lifting motion as a new behavior or gesture for the current synthesized sound object)

The system design approach is adaptive, one that aims to create synthesis to express physical real world properties in collaboration with participants through social interaction, sound synthesis and diffusion.

The emphasis on compositional content *[1] (Bongers 2002)* rather than purely refining the interface technology has proved to be a significant design methodology, each interface element is fairly simple, basic electronics are used, with this system the combination of participants behavior, adaptive software and 'smart' interfaces creates a new compositional process. Through further observation and refinement of this type of system a deeper understanding of 'play states' and collaborative compositional processes will be described. *"Response to musical stimuli can cause significant changes in both behavior and brain activity"[5] (Machover T. 2004)*

Developing systems that adapt and respond to these essential elements of musical activity is a demanding challenge to this field. Consideration of social interaction through the medium of sound is a core concern of this research; how we perceive and interact with sound environments or interface objects that adapt to our behavior. In this sense the Orb3 interface is 'smart' our social interactions and interplay are part of the 'instrument' but the instrument is not merely a separate controller or extension of an individual performer, it is a socially mediated compositional environment with the potential to adapt to emergent behavior. 'An adaptive systems approach that exhibits process driven collaboration.' [13]

4. REFERENCES

[1] Bongers B. Harris Y. A (2002) Structured Instrument Design Approach: The Video-Organ *Proceedings of the Conference on New Instruments for Musical Expression, May 2002, Media Lab Europe, Dublin Ireland.*

[2] Bowers, J. (2001) TONETABLE: A Multi-User, Mixed Media, Interactive Installation P2, and 4,5.

[3] Franco E. Griffith N. Fernstro E. (2004) " Issues for Designing a flexible expressive audio visual system for real-time performance & composition" *Proceedings of the Conference on New Instruments for Musical Expression, June 3-5, 2004 Shizuoka University of Art and Culture Hamamatsu, Japan*

[4] Lippe, C (2002) "Real time Interaction Among Composers, Performers, and Computer Systems" *Information Processing Society of Japan SIG Notes, Volume 2002, Number 123, pp. 1-6.*

[5] Machover T. Shaping Minds Musically *BT Technology Journal Vol 22 No 4 October 2004*

[6] Newton-Dunn H., Nakanon H., Gibson J. "Block Jam: A Tangible Interface for Interactive Music" *Proceedings of the Conference on New Interfaces for Musical Expression*, Montreal, Canada *2003*

[7] *Paine G. (2004) "Gesture and Musical Interaction: Interactive Engagement Through Dynamic Morphology" Proceedings of the Conference on New Instruments for Musical Expression, June 3-5, 2004 Shizuoka University of Art and Culture Hamamatsu, Japan*

[8] Patten J., Recht B., Ishii H."Audiopad: A Tag-based Interface for Musical Performance" *Proceedings of the Conference on New Interfaces for Musical Expression*, Dublin, Ireland, 2002.

[9] Weinberg G. Key Note Address: Interconnected Musical Networks –Bringing Expression and Thoughtfulness to Collaborative Group Playing" *In Proceedings of the International Computer Music Conference 1-6 November, 2004 University of Miami, Florida, USA*

[10] *Woolf S. Experiments with Reactive Robotic Sound Sculptures. ALife VIII: Workshop proceedings 2002 P2, 3.*

[11] Livingstone D. and Miranda E. "Composition for Ubiquitous Responsive Sound Environments". *In Proceedings of the International Computer Music Conference 1-6 November, 2004 University of Miami, Florida, USA*

[12] Livingstone D. (1998) "The Space between the Assumed Real and The Digital Virtual". *In Ascott, R. Reframing Consciousness – Art, Mind & Technology Intellect Books pp138-143*

---

[i] *Composer and Listener Objects detailed here form part of a larger integrated system included in proceedings ICMC 2004* [11]

[ii] *Details of this symbol based adaptive tracking system are included in proceedings ICMC 2004* [11]

TACTILE COMPOSITION SYSTEMS FOR COLLABORATIVE FREE SOUND

Dan Livingstone
Computer Music Research

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth PL148AA United Kingdom
dlivingstone@plymouth.ac.uk

Chris O'Shea
Digital Media Artist

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth PL148AA United Kingdom
me@chrisoshea.org

ABSTRACT

Numerous innovative controllers and collaborative tactile interfaces have been developed for social interaction with sound. This evolutionary field of interaction design has led to a wide range of compositional models that increasingly mirror the open source methodologies developed by the creators of such systems. The authors consider the software integration of such systems and propose a potential model for free sound composition. We speculate on how these integrative approaches are leading to new compositional frameworks for distributed composition, providing an overview of how an open source development approach influences the structure, interaction design and compositional output of such systems. The range of related works in this field is considerable, selected examples are considered in terms of interaction models & compositional approaches that offer a free sound or open source model for social collaboration with the potential for distributed composition.

1. INTRODUCTION

Our discussion focuses on the potential of collaborative tactile interfaces to extend the notion of free sound composition. This is not an exhaustive or comparative survey, instead we have chosen to focus on a small range of tactile or tangible interfaces that each offers a different interaction framework for participants to explore. Several of these examples are well documented by the original authors; others are lesser-known systems that offer complimentary approaches. In most cases the designers of these systems had specific audiences or interaction methods as a design objective of each system. We summarise the core features of each and provide a brief analysis as to how each interaction model can contribute to a wider knowledge of interaction design for tactile or tangible collaborative composition systems.

2. EXAMPLE SYSTEMS

Each system has a tangible interaction model; *Soundgarten* [7] offers a toy like collective floor based building process, combining elements of a single object, enabling children to record, modify and arrange samples.

ISS Cube [5] functions as a collaborative table top spatial mixing surface for up to four participants using simple movement of tactile objects, this builds on a more strategic model of play and exchange with intuitive interaction based on relative movement and location of small discs, reminiscent of many board games and intuitive to use with collaborative 'positioning' of predefined sound samples. *Audiopad* [4] is a well-documented work in the audiovisual tactile mixer field offering real-time visual feedback in addition to a tangible control interface with tactile elements. The fourth example, *RGB Player* [1] allows manipulation of sequence and pattern, either collaborative or turn based placing/removal of colored objects. *Block Jam* [3] combines an element of building or assembly to construct a sequence, pattern variation and control of audio flow. A different approach can be seen in the design of *ReacTable* [2] in this system of tangible objects, textural qualities, topological markers and simple gestures are combined to trigger or represent different types of synthesis. Each example is intended for a different type of social interaction, for example; group discovery, individual or turn based interaction, collective play and collaboration.

3. INTERACTION MODELS

These examples have interaction models and functionality that can be simply categorised as Exploratory, Organisational, Sequential and Relational. It is interesting to note that the example aimed at the youngest audience naturally offers the personalisation of the sound-scape through live sampling, whereas the potentially most compositionally experimental work uses visual metaphor to indicate sound synthesis processes.

3.1. Exploratory model

Soundgarten is *"a tangible interface that enables children to record, modify and arrange sound samples in a playful way"*[7]. The project is aimed at 4 to 6 year olds, with the objective of developing early musical education with pre-school children. The interface for Soundgarten resembles a children's toy, where the surface of the garden is the performance stage.

The garden has 19 plug holes that allow sound samples to be triggered by plugging in a *mushroom*. The 3 vertical levels of the garden control the volume of each sample. A microphone, called a *shuffle* in Soundgarten enables a child to record sounds in their environment, and by plugging a mushroom object into the shuffle, the recorded sound can then be plugged into the garden. As the microphone is wireless it allows the children to roam around to find interesting sounds to record, rather then being confined to being sat around the project, which would inevitable lead to the recording of sounds already being produced by the garden. As well as the ability to record sounds, Soundgarten is loaded with a set of pre-defined sample banks. Each sample mushroom contains an icon on the top to indicate the sound produced. The colour of the icon also denotes the type of sound, such as blue for environment sounds, such as wind blowing, alarm clock or dog barking and brown colour for instruments, such as drums or violin. Soundgarten also enables a set of effects on the sounds associated with each mushroom. Filters such as echo, resonance, play backwards, increase & decrease pitch can be applied to a sample via *attribute* objects. These attributes resemble a flower petal or leaf and can be plugged into the top of a mushroom and adding more than one attribute will combine the effects.

The designs of the tactile attribute objects in Soundgarten don't seem to correspond to the effect on the sounds, such as echo or increase pitch, but in this case that isn't necessarily a problem. Given the target audience of the project, a child of this age wouldn't have a grasp on technical working of those effects, but would just need to remember what each attribute object did. Soundgarten aspires to be extendable, "Like Lego, Fisher Technique and other constructive toy systems SOUNDGARTEN provides an open system, which can be expanded indefinitely" [8]. This may be the case with the production of new sound samples, tangible objects and perhaps a larger playing surface, it won't be able to achieve a free open play environment such as Lego, due to the structure of the plug holes and objects themselves. The ability to combine sounds, by plugging them on top of each other would make this more open. This would enable detailed gardens to be built by the children.

3.2. Organisational model

The Interactive Surround Sound (ISS) Cube is a surround sound mixer that allows users to spatially position a sound using tactile objects. The aim of this project is that *"users of the system can easily change their mood by recreating their spatial sound scenery. For example, nature sounds can be positioned within the space to create a calm and natural environment"* [5]

ISS Cube has 4 coloured pucks, called *carriers* that allow the users to select a predefined sound sample by moving it to the edge of the surface, where a selection menu will appear. Once a sample is selected, moving the carrier across the surface will spatially position the sound using a 4 speaker set-up. Each corner of the surface representing one of the speakers, so the sounds pan between each speaker based on the position of the carrier relative to the surface corner.

A second type of tactile object, a smaller white puck, controls the volume of each sample. The closer the sample carrier to the volume, the louder it becomes within the space.

The focus of ISS Cube is to allow collaborative mixing of sounds in a space; "due to the multiple input devices, the square tabletop display, which enables equal access from all sides, invites collaborative interaction" [5]. As there are only 4 carrier objects to control samples, this only allows 4 people to collaborate of the positioning of sounds at a time, with enough space around the table for spectators.

3.2.1 Audiopad

Audiopad is a tactile interface for musical performance. The initial aim was to increase the stage presence of laptop style performers. Audiopad is essentially a mixer, allowing performers to trigger sound samples, control volume and various effects on those samples.

Interaction with Audiopad is via a series of pucks, each with a different action. Sample pucks are used to carry sound sample banks. By moving a sample puck over an area of the interface a performer can then select a group of samples from the graphical menu. A selector puck placed near a sample put brings up a graphical tree menu for choosing a sample. A microphone puck controls the volume of a sample based on the distance between.

A projected graphical interface provides instant feedback to the performers using Audiopad. Graphics are placed over the position of each puck, providing local details about the selected sample, volume, on / off state and applied effects. *"Our exploration suggests that this seamless coupling of physical input and graphical output can yield a musical interface that has great flexibility and expressive control"* [4]. The level of sound control in Audiopad is based around selecting predefined samples, altering their volume and applying effects. Samples are held in Ableton Live, with control parameters being passed to it via MIDI by the tracking interface. Effect filters, such as delay or low pass are assigned to different groups of samples, so a performer is not free to add every effect to each sample, however this focused approach leads to a level of intuitive interaction which is highly accessible.

3.3. Sequential model

RGB Player began as a *"dynamic physical interface that would allow any everyday object to become a device of interaction"* [1]. Through the artist's own interest in creating sound from visuals, RGB Player was an exploration into the reverse of this process. The main compositional feature of RGB Player is the ability to create a patterned sequence by placing objects in a line around the disc. Drum sequences that increase and decreases can be built up, mixed in with fast repeating guitar samples and piano keys for examples.

The interface for RGB Player consists of a rotating glass disc and a slit in the surface that scans any objects that pass over it. Beside it stands a variety of small colourful children's toys, that when placed on the rotating disc, trigger sound samples as they pass over the scanner.

Toys and objects placed onto RGB Player are scanned by an internal webcam, which then translates the RGB values into one of 6 sound samples, depending on it's nearest colours, from bass to drums and piano. The distance of the object to the centre of the instrument determines the pitch of the sample played, with less distance emitting a higher pitch.

The rotating disc in RGB Player serves as a good metaphor for the loop of a sound sample. With each full cycle the composition goes back to the starting point to begin the sequence again. The only downside to this is the inability to stop the disc from rotating, so one eventually becomes dizzy following objects around and trying to generate a pattern.

3.3.1 Block Jam

Block Jam is a musical sequencer that allows players to control the order of sound samples using a series of connected tangible blocks. "Block Jam is not a musical instrument; it is an alternative means of controlling a sequencer. It has no means of continuous control or gesture" [3]. The aim of Block Jam is to create an accessible collaborative musical interface.

The player interacts with Block Jam via 26 physical blocks. Each block contains visual feedback via an LED matrix, a push button and rotating dial style input. Initially players start with a *play* block, to which sample blocks can be connected, by putting them side by side they lock into place.

The visual feedback on Block Jam displays the state of each block, which indicates the direction of play in the sequence, such as straight, corner (change direction) or a gate (rotate direction). The player can select from one of three sound sample banks for each block by rotating their finger on the dial interface, with each sample bank containing 5 sounds. The colour displayed on the block indicates which sample bank is currently active for that block (red, orange, green). It is unclear as to why these colours were chosen, as they do match that of traffic signals, which would suggest a stop or go action, but this is not the case.

The speed of the musical sequence in Block Jam is determined by the length of time the player pushes the button on the play block before releasing. Each sample contains 3 variations to match the 3 possible speeds of play, as apposed to simply speeding up or slowing down the rate of play of one sample.

3.4. Relational model

ReacTable is an instrument for collaborative performance. At the time of writing the system leads the field toward the design of tangible objects in relation to the sounds generated. Haptic encoding such as object shapes, surface texture and colour have been explored. Surface texture gives users indication as to the timbral properties of a sound, *"Noise generators have a completely irregular texture and different types of sanding paper can represent a granular synthesizer"* [2]. In earlier versions it is unclear as to whether surface texture communicates effectively as a method of identification, as a performer would have to at least understand the terminology and process of each, like saw-tooth generator for example. The ReacTable development team has also experimented with surface materials, such as using plastic (for synthetic sounds) and wood for organic sounds.

The Reactable team has recently been exploring the use of topological markers to indicate the relationship between object and interaction. In addition they are refining their camera tracking methods by using these markers to refine object identification, orientation and relative position.

4. A FREE SOUND MODEL

Each model discussed has core elements that help to define a Free Sound Integrative model. From the exploratory model, the process of building, reconfiguring and live sampling participants provides an open inclusive form of interaction. The Organisational model shows that conventional control mechanisms can be far more intuitive using tactile objects supported by visuals that reinforce interaction and functionality in a combined perceptual interface. The Sequential model offers a tactile method of assembly of scored elements with pattern variation, a reconfigurable linear process. The Relational model establishes a potentially more direct kinaesthetic linkage between objects, textures and sound properties, a form of haptic encoding. The distributed model allows for virtual interaction within a shared compositional online space, where participants create spatial and visual relationships while exploring a range of sound juxtapositions that can be added to through file upload and exchange.

A Free sound approach for the creation of collaborative tactile composition can be described as one which integrates key features of all of the above elements. Common features in this type of system would be not only interaction to influence spatialisation of predefined samples, but the ability to add new source material through file upload, live sampling or real-time synthesis. Sequencing and flow of sound elements over a distributed composition also allows for different compositional elements to be modified either simultaneously or in direct response to the interaction of other users. This also suggests that an evolutionary or algorithmic approach to generate new composition from shared elements would extend the open nature of such works.

4.1. A Free Sound Approach

To establish this model, a framework needs to be in place for integration between programming software, sound applications [6], visual output and tracking systems.

We have implemented one such approach, to illustrate an integrative methodology, more recent development of the Sonicforms platform has been motivated by further use of Open source libraries, currently in development.

4.3 Sonicforms

We introduce Sonicforms [8], an online open source research platform that promotes a free sound approach for the development of new audio or visual works, mediated through a tangible tactile interfaces. The structure of the project allows a range of open source software to be used in the creation of the visual and audio output, whilst the tracking system sends information about each tactile object to the software. Using TEMP, a communication gateway server, any messages can be communicated between programs, such as UDP and TCP. The project is open to contribution, enabling other audiovisual artists and developers to implement audiovisual works for tactile collaborative interaction.

5. CONCLUSIONS

In order to implement a successful tactile interface for collaborative composition, it is extremely useful to consider the interaction models evident in related systems. Each system discussed has received highly favourable responses from audiences & participants.

These systems function effectively and have been professionally implemented. Each example discussed has gone through a process of refinement in terms of interface, interaction mode and sound control, in several cases these are long term projects supported by an integrated team of researchers and practitioners. When considering a 'Free Sound' or an adaptive compositional approach to tactile interaction the underpinning technologies can determine the methods available. In some cases the limitations of a specific software based approach may define the resulting compositional parameters of such systems. By considering the interaction models embedded within each system we have been able to draw on this 'best practice' to identify the core elements of an adaptive or 'Free Sound' approach, we have also considered the potential software limitations and provide an example system that utilises an integrative software approach, including a summary of software integration to extend the interaction, composition and broadcast potential of tactile compositional environments for collaborative composition. In conclusion, the open source community continues to provide versatile tools, extensions, externals and libraries that support and extend a broad range of approaches, identifying interaction models within tangible collaborative music systems is a very useful methodology for identifying the most effective development route.

6. REFERENCES

[1] Barter T. RGB Player, exhibited next2004 (Denmark), Royal College of Art Degree Show 2004 (extract from interview by author.)

[2] Kaltenbrunner, M. & O'Modhrain, S. & Costanza, E "Object Design Considerations for Tangible Musical Interfaces" *Proceedings of the COST287-ConGAS Symposium on Gesture Interfaces for Multimedia Systems, Leeds (UK) 2003*

[3] Newton-Dunn H., Nakanon H., Gibson J. "Block Jam: A Tangible Interface for Interactive Music" *Proceedings of the Conference on New Interfaces for Musical Expression*, Montreal, Canada *2003*

[4] Patten J., Recht B., Ishii H."Audiopad: A Tag-based Interface for Musical Performance" *Proceedings of the Conference on New Interfaces for Musical Expression*, *Dublin, Ireland, 2002.*

[5] Quarta, M. ISS Cube, *exhibited at Ars Electronica, Cybersonica,, Bafta Interactive Festival 2003* http://www.somethingonline.org/txt/d_iss.html

[6] Wright M., Freed A., Momeni A. "OpenSoundControl: State of theArt 2003" *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03), Montreal, Canada. 2003*

[7] Woolf M. Documentation of exhibited work http://www.soundgarten.com/SG_full_documentation.pdf (extract from interview by author.) _

[8] Sonicforms - http://www.sonicforms.org/ (open source research platform)

ORB3 – MUSICAL ROBOTS WITHIN AN ADAPTIVE SOCIAL COMPOSITION SYSTEM

Dan Livingstone
Computer Music Research

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth PL148AA United Kingdom
dlivingstone@plymouth.ac.uk

Eduardo Miranda
Computer Music Research

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth PL148AA United Kingdom
eduardo.miranda@plymouth.ac.uk

ABSTRACT

Gesture capture, motion tracking and 3D visualisation technologies have generated many new musical forms, often extending the mannerisms or behaviours of a given performer or discipline, providing new compositional frameworks for real time synthesis in response to action.
In many cases these approaches are presented within a single domain, a live stage performance, a site specific installation, a shared networked visualisation of collaborative composition. The reality is that these 'interactivating spaces' [1] whether haptic, [5] tactile [9] or ubiquitous [11] is that they manifest new forms of interaction, between people, systems and the medium of sound.
Free Sound can be understood to be an extension of the 'open work' where the base materials for a compositional process are created through a model of exchange, interaction and resynthesis. The resulting output of these activities can be broadcast and disseminated through a range of technologies to both social and private spaces. This research suggests that there are new interaction models and social compositional frameworks to be found in these cybrid spaces, a previously intangible location often dominated by the broadcast and publishing industry. A marketing model defined by revenue streams and a value chain. In the case of socially mediated composition or 'free sound' there is still a value chain, it's investors and beneficiaries are the open source community, the collaborators and participants within such mediated systems and the resulting free sound.

Keywords

Adaptive System, Sound Installation, Smart Interfaces, Music Robots, Spatial Music, Conscious Subconscious Interaction, Interaction models.

1. INTRODUCTION

In the design of new interfacing methods [1] for sound manipulation and control it is often the case that the primary focus is the point of tactile interaction, the exploration of new gestural controllers or methods for mapping and transforming data to create sound material [2]. This approach has led to the development of numerous novel and individual interfaces [3], in many cases the interaction mode is learnt by the user, in order to complete the feedback loop, thereby achieving dynamic results through an exploratory model of interaction.
With a modular adaptive systems approach the emphasis is on providing an interface framework for different types of interaction that can be initiated by both users and 'smart' interfaces, ie new interaction behaviors can be identified by the system independently, in response to users actions, whether direct tactile control or simple movement, location, gesture or position. Figure 1 illustrates the ORB3 system, an example of an adaptive systems approach.

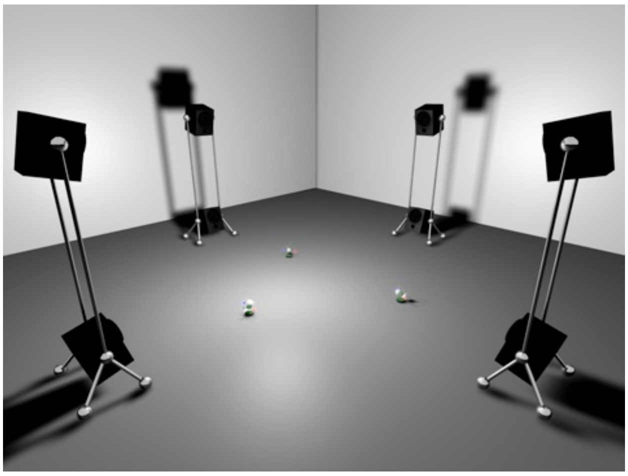


Figure 1. *Auditory Sphere. 8 active speakers angled to provide versatile software controlled diffusion. Diffusion and synthesis generated from environment/interaction data collected by each Orb. Software developed in MAX/MSP running on G4 Apple laptop with M-Audio 410 Firewire mobile multi - channel interface, custom built 'composer – listener' objects (wireless Orb3 interfaces)*

Orb3 is a compositional space that can operate autonomously through each interfaces interaction with its' environment, each Orb collects the base materials for sound composition, participants moving between Orbs influence their motion and data collection behaviours, Orbs can also be picked up and held as a controller for the sound material they have generated, each Orb has the potential, through software, to adapt to its current user, a number of models have been developed to motivate this interaction, based on observation of interaction with a number of novel interfaces and custom controllers:
"…to have a musical response accentuated by the player who sent the original call, to plant a musical "seed" that

would be picked up by the group in various manners, etc." Weinburg [9] A significant observation during performances by children using the Beatbug system developed at MIT was described at the International Computer Music Conference in Miami 2004; the children made exaggerated swooping motions with the Beatbugs as they 'passed' sounds while interacting with the controller. At the time the Beatbugs were not equipped to react to this *emergent behavior*, although neither the audience or children were aware of this at the time. This observation led to the next refinement of this approach to a network of interfaces, potentially using blue tooth technologies and motion detection to refine and utilize this interaction, this anecdote reinforces the value of the adaptive systems approach we present which is a continually evolving field of applied research for novel interfaces and interactive music systems.

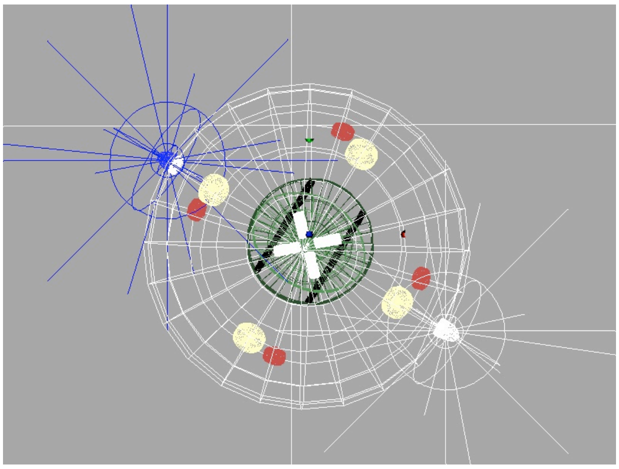


Figure 2. *Orb3 Design Constraints. Design aesthetic can be achieved by designing custom PCB's for sensor placement, careful selection of plastics for manufacture and careful consideration of tactile properties for intuitive interaction.*

2. ORB3 DEVELOPMENT PROCESS

The Orb3 interface design was developed through observation of interaction with wired 'composer and listener' objects[i]. These original objects were static spheres housing a cluster of analogue sensors (Light dependant resistors, bead thermistors, vibration and tilt switches etc) for measuring ambient light, ambient heat, general motion and orientation. The original system comprised four such spheres which could be placed and relocated to generate and vary data used to synthesize sound material for 7.1 sound diffusion controlled by a gesture and motion based video tracking system[ii]

Each sphere contained a total of 8 sensors, wired to a control voltage to midi converter (Infusion Systems Icube) this method worked effectively for developing software and refining synthesis and sound design for prototyping a large-scale adaptive system. Local interaction was less successful due to restricted movement of wired objects and unexpected behaviors and reactions of participants. For example; using the prototype system the shadows cast between spheres as participants moved around the room were recorded by a drop in light values sensed by the sphere affected, causing subtle changes in base sound materials generated for the sound-scape, this was an intended compositional element of the system but on realising this process, many participants could not resist the temptation to explore further, initially cupping or shielding areas of the spheres and inevitably moving and repositioning them, anticipating a direct response. It was immediately apparent that the simplicity of the sphere encouraged a series of interactions that could further inform sound design for socially mediated sound spaces. It also led to the realization that the software techniques applied to the vision system for adding new data relationships based on symbol recognition could be migrated to the interface design for each sphere developing more expressive tactile control, and more significantly, using the relative position and orientation of each sphere as a compositional parameter that could be heard in the diffused sound-scape, that was also registered by visual or tactile feedback on the interface itself. Other observations were that often participants chose to work collaboratively, taking a sphere each, influencing a parameter passing it on, this worked particularly effectively in groups of three, where patterns of motion and exchange had the potential to create rhythm and flow, some general experiments were done with different numbers of spheres to see if this affected interaction modes, it is speculated that providing an odd number of interfaces provides more movement through transfer and exchange and encourages turn taking. It was also noted that during periods of inaction or when participants were more passive different listening modes were reported, this in turn has influenced the sound design of the refined system, incorporating different 'play states' or modes - some further controlled experiments are in data to support these assertions. The logical development of these passive and active modes mediated by participants is to add simple robotics to each sphere to allow each one to move and interact with other spheres independently.

2.1. Design for Collaboration

Having established some significant refinements from the initial prototypes a specification for a more robust adaptive interface was resolved. Primarily a wireless approach was required, high performance with reliable transfer of digital and analogue data from sensors, in addition a wireless microphone embedded in each unit for live sampling. Internal lighting was added to indicate interaction modes and force feedback in response to interaction. These features introduced new design challenges, as the revised design needed onboard power for wireless operation and ideally solar charging to extend session times. A final addition was the inclusion of lasers and proximity sensing to enable quick alignment and event triggering between spheres. A simple method for overhead positional video tracking (max/msp Jitter) using a single fixed camera provides an effective method for documenting movement and behavior of each orb during a live session through time lapse imagining.

A mobile robotic element has been prototyped for each sphere, allowing them to move and reposition themselves autonomously or in 'collaboration' by integrating positional tracking (digital compass/distance and trajectory) and proximity triggers (ultrasound sensors). This dynamic motion provides a visual element that reveals the compositional potential of the system, while demonstrating some of the synthesis and diffusion properties that are influenced by the interaction between or with each sphere. When each Orb is collecting data to influence sound synthesis and diffusion, or being followed or manipulated by participants this collaborative process can be displayed from a top down perspective, using either projection or plasma screen display.

"Most of the systems that allow the creation of sound and image in real-time don't have the capability for organizing events at a global level. This is however, required if the aim is to allow the composition of a piece that involves feedback from events sonic and visual, in the construction of interactive audiovisual compositions." [3] (Franco et al 2004)

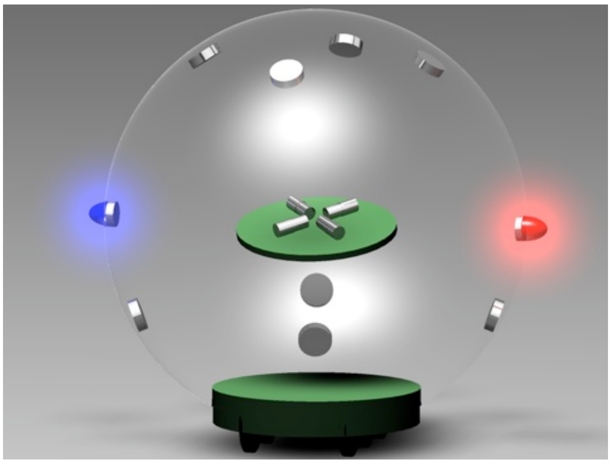


Figure 3. *Wireless Mobile Orb v2.0 in Absorb mode.*

(OrbV2.5 features data collection through light, temperature, orientation, motion sensing, laser alignment, microphone, mobility, rotation, Exploratory interaction transforming data into sound material)

For the purpose of this paper emphasis has been placed on the Orb3 interfaces, the key features are interaction modes and social composition, simply expressed as 'play states'. Sound is the primary medium but in order to make visible the transformative processes underpinning the compositional output ways are being sought to create a visual aesthetic from both the data and interaction of people, making visible behavior and interactions, effectively creating graphical transcription as real time feedback to participants.

Developers of collaborative musical interfaces with tactile, graphical and sensory feed back are developing new terminologies to describe the design process for these systems in terms that begin to articulate their compositional and social modalities. Collaborative interactive music systems, such as 'Block Jam' (Newton-Dunn, Nakano, Gibson 2002) where interconnecting blocks are collaboratively assembled to organise musical phrases and sequences begin to identify new musical forms enabling participants to create *'meaningful musical structures' through 'collaboration and exploration'* [6] (Newton Dunn et. al. 2002)

Other collaborative works such as ToneTable [2] (Bowers J. 2001) use interactive visual elements as an integral interface element, in this case participants manipulate 4 trackballs, 'disturbing' a projected fluid surface with associated textures and diffused sounds, again it is the observation of improvisation and collaboration with a real-time composition system that distinguishes this emerging musical form. The author discusses emergent behaviors and extended engagement as a development of the system design; *'structures of motivation' 'variable participation'* [2] (Bowers, J. 2001) A highly refined table top tactile control surface for two or more participants has been developed by Patten and Brecht, 'Audiopad' [8], which has been extensively exhibited. The system provides a graphically dynamic projected overlay oriented around electronically tagged tracked physical objects or 'pucks' for real-time control of preprogrammed electronic music, moved by hand with fingertip control. Key elements in terms of a compositional model are *'spontaneous reinterpretation'* and a combination of *'visual and tactile dialogue'[8](Patten J. Brecht B. 2003)*

The design and installation for the Orb3 system forms an auditory sphere (fig. 1.) using an 8 Channel sound diffusion through which participants move, view, listen and reconstruct the compositional process through social interaction within it. The audiovisual feedback in response to these varied interaction modalities is an active process, one of content driven collaboration.[iii]

2.2. Communications & Parameters

Each Orb sends data via a 2.4ghz wireless RF interface to a G4 laptop running Max/MSP, a combination of analogue and digital data can be sent and processed by the control software created in max. The software itself is not simply a parameter mapping utillity, it is designed to correlate different data against previous interactions, a form of compositional memory where environmental parameters of previous sessions are compared with current ones to identify repeated behaviors of the system and actions of participants. The software is designed to be adaptive, previously un-recorded or new data configurations are identified and used to compose new sound events or objects. The software sends data to each Orb to indicate it's state and trigger visual or tactile feedback, ie; activate laser/proximity sensing for positioning, activate status leds, activate force feedback.

Each Orb has two compositional states – *Absorb* and *Adapt*. In *Absorb* mode an Orb is autonomous and located on the floor, it's sensors are calibrated to collect environmental data, ambient light, ambient temperature, relative position and orientation, it can also live sample sound for processing - the software controls this calibration which is activated through Orb alignment – each Orb is fitted with proximity sensor, a laser and LDR - placing the three Orbs in a triangle and directing each Laser to the next Orbs locating LDR activates this mode, which is part of the initial setup process.

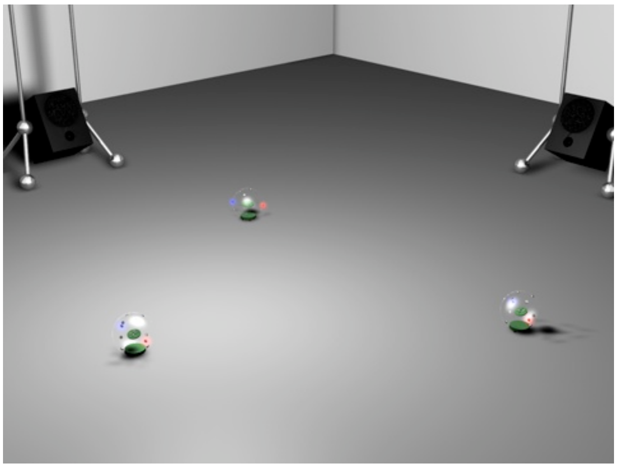


Figure 4. *Laser alignment - Triangulation.*

(Alignment –view of each Orb, lasers are activated, two Orbs are shown in listening 'Absorb' mode after calibration, one (lower right) is about to move out of alignment in response to parameter changes, autonomous – may move or rotate to attract participants, system responds by panning a sound object in relation to Orb location).

Environmental data changes are usually slow in interior environments so these elements are mapped to the timbre and color of sounds created with larger fluctuations affecting diffusion, thus providing an overall structure for the real time composition that is responsive to ambient light, temperature and general movement. *Adapt* mode is activated when the alignment of Orbs is disturbed, either by walking between them, interrupting the laser tracking or by picking them up which also activates vibration sensors and initiates orientation mapping - angle and orientation of each Orb in this state directly influences panning and diffusion rates of synthesized sounds. During *Adapt* mode the laser is deactivated and the ambient temperature measurement is recalibrated to respond to body heat through hand contacts on the Orbs lower surface. Bead thermistors with fast response times are used so as an Orb is passed from one hand to another, or between users, it registers and marks these changes. This data combined with orientation data allows for a range of subtle and dynamic sound events to be initiated by each participant in collaboration with both the system and with other people.

2.3. Emergent Behavior

As an adaptive portable system, the Orb3 environment creates an opportunity for observing and recording forms of emergent behavior in relation to spatial sound interaction, this provides researchers in this field with a structured framework to inform the design of mobile and autonomous interfaces, such as musical robots or adaptive social composition systems.
" we should not forget that humble reactive robotic systems capable of sensing and reflecting the complexity of their environments have the capacity for unpredictable and life like behavior that encourages playful somatic interaction." [10] (Woolf & Beck 2002)

The inclusion of play through collaboration is not a by product of this system, it has been developed explicitly to motivate different responses through consideration of ergonomics and human factors, developing from the considered observations of researchers and practitioners in related fields. The 'play states' or modes titled *Absorb* and *Adapt* have been designed with consideration of both composer/listener object interaction and the listening process or perceptual triggers to motivate participants.
In the 'play state' *Absorb* the Orbs are programmed to activate when certain parameters or sound events are captured, or when conditions match previously encountered sequences, the 'intention to listen' is shown through both the status LEDS and motion/rotation in response to stimulus. This modality can also be attributed to the behavior of participants, who move towards the 'Auditory sphere' of course initially their interest is more likely to be the spectacle of the technology or other participants behavior, however moving into the 'Auditory sphere' shows an intention to participate, *to listen.* Participants interaction at this stage can be described as Subconscious, they are not necessarily aware that their presence and orientation is influencing the system.

The 'play state' *Adapt* is active when the triangular alignment of the three Orbs is disturbed, the software reconfigures itself to a more sensitive state, ready to be interacted with, held, passed, moved in relation to sound synthesis & diffusion as perceived and manipulated by a participant. Applying a different interaction model to the parameter mapping of sensors is an effective way to initiate ORB states. In software terms this is achieved by switching the algorithms mediating data analysis, through patterns stored in short and long term memory (Max objects *capture, decode, funbuff, histo* and *spray* are integrated with *mtr* to record, store and replay streams of data, which are compared against previously collected and live data [short term memory], a form of score following). By picking up an Orb a participant is moving from the *Absorb* state, instead choosing to interact, to explore and through this action perceiving and identifying the source of broadcast sounds, through their manipulation of an Orb. This modality is further reinforced when direct control of sounds are influenced by the participant. Their behavior changes as they *Adapt* to the parameters they have influence over. This can be described as Conscious interaction, a heightened state of attention and engagement, [6] [Newton Dunn et. al. 2002] the intention to collaborate with the system and others using it, improvisation, not simply 'call and response' [4] [Lippe C. 2002] as there are no familiar, formal or structured elements in the form of musical patterns, note sequences or beats inherent in the open nature of this spatial sound environment. A key development with this system is that it continues to adapt while capturing, archiving and broadcasting new behaviors. A range of technologies have been explored a custom PCB with optocouplers was implemented to take multiple outputs from a midi to control voltage converter to send motion and trajectory motor control to each Orb using modified consumer radio control vehicle parts, modification to gearing, a new steerage mechanism and a form of pulse width modulation to control Orb trajectories provides more subtle movement. Multiple overhead color tracking has been tested, a fine balance between ORB speeds and processing of multiple tracked colours (LED combinations as figure 5.) in variable conditions for effective mediation was required. Several task-based experiments using perceptual constructs [12] to establish and refine interaction models have been used to refine compositional processes within the system. Interaction models have been tested by assigning participants simple compositional tasks based on establishing mental models for the relationship between sound objects and their perceived location in the auditory sphere. Quantitative data can be extracted from the short term memory of the system, noting the start event of an interaction, such as repositioning a sound by manipulating an orb. After this action has been archived Qualitative data from participants' reports can be established by comparing transcribed verbal accounts of the set task and system response. Both data types can be considered in context by reviewing the overhead broadcast documented by the system, archiving actual position of participants and orbs against archived positional data.

3. CONCLUSIONS

The compositional approach is not modeled on a *'fixed or even consistent excitation-sonification relationship'* [Paine G. 2004] many elements of the sound-scape generated are through transformative synthesis methods, in this instance the creation of sound through traditionally unrelated real world variables. Neither are the sounds randomly generated; the capture and transformation of variables such as heat, light, proximity, motion and time create values that could be mapped to conventional parameters for musical control of predefined note sequences, loops and formally structured phrases but in this adaptive approach through a process of observation, listening and sound design these parameters are treated as explicit elements of the real time composition environment, ie the ORBs are designed to be responsive to their physical environment, with adaptive behaviours that motivate human interaction.

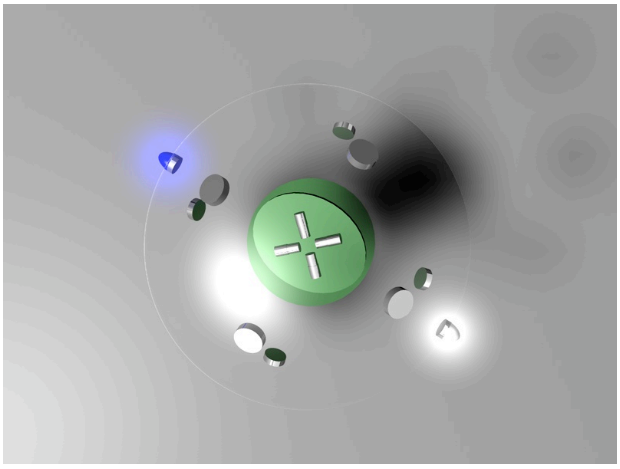


Figure 5. *Orb in adapt mode.*

(Overhead camera view (simulated), Showing a single Orb As an Orb is picked up sensitivity is heightened by re mapping parameters, accelerometer sensitivity maps motion, angle, orientation to sound diffusion while archiving lifting motion as a new behavior or gesture for the current synthesized sound object)

The system design approach is adaptive, one that aims to create synthesis to express physical real world properties in collaboration with participants through social interaction, sound synthesis and diffusion.

The emphasis on compositional content *[1] (Bongers 2002)* rather than purely refining the interface technology has proved to be a significant design methodology, each interface element is fairly simple, basic electronics are used, with this system the combination of participants behavior, adaptive software and 'smart' interfaces creates a new compositional process. Through further observation and refinement of this type of system a deeper understanding of 'play states' and collaborative compositional processes will be described. *"Response to musical stimuli can cause significant changes in both behavior and brain activity"[5] (Machover T. 2004)*

Developing systems that adapt and respond to these essential elements of musical activity is a demanding challenge to this field. Consideration of social interaction through the medium of sound is a core concern of this research; how we perceive and interact with sound environments or interface objects that adapt to our behavior. In this sense the Orb3 interface is 'smart' our social interactions and interplay are part of the 'instrument' but the instrument is not merely a separate controller or extension of an individual performer, it is a socially mediated compositional environment with the potential to adapt to emergent behavior. 'An adaptive systems approach that exhibits process driven collaboration.' [13]

4. REFERENCES

[1] Bongers B. Harris Y. A (2002) Structured Instrument Design Approach: The Video-Organ *Proceedings of the Conference on New Instruments for Musical Expression, May 2002, Media Lab Europe, Dublin Ireland.*

[2] Bowers, J. (2001) TONETABLE: A Multi-User, Mixed Media, Interactive Installation P2, and 4,5.

[3] Franco E. Griffith N. Fernstro E. (2004) " Issues for Designing a flexible expressive audio visual system for real-time performance & composition" *Proceedings of the Conference on New Instruments for Musical Expression, June 3-5, 2004 Shizuoka University of Art and Culture Hamamatsu, Japan*

[4] Lippe, C (2002) "Real time Interaction Among Composers, Performers, and Computer Systems" *Information Processing Society of Japan SIG Notes, Volume 2002, Number 123, pp. 1-6.*

[5] Machover T. Shaping Minds Musically *BT Technology Journal Vol 22 No 4 October 2004*

[6] Newton-Dunn H., Nakanon H., Gibson J. **"**Block Jam: A Tangible Interface for Interactive Music" *Proceedings of the Conference on New Interfaces for Musical Expression*, Montreal, Canada *2003*

[7] *Paine G. (2004) "Gesture and Musical Interaction: Interactive Engagement Through Dynamic Morphology" Proceedings of the Conference on New Instruments for Musical Expression, June 3-5, 2004 Shizuoka University of Art and Culture Hamamatsu, Japan*

[8] Patten J., Recht B., Ishii H."Audiopad: A Tag-based Interface for Musical Performance" *Proceedings of the Conference on New Interfaces for Musical Expression*, Dublin, Ireland, 2002.

[9] Weinberg G. Key Note Address: Interconnected Musical Networks –Bringing Expression and Thoughtfulness to Collaborative Group Playing" *In Proceedings of the International Computer Music Conference 1-6 November, 2004 University of Miami, Florida, USA*

[10] *Woolf S. Experiments with Reactive Robotic Sound Sculptures. ALife VIII: Workshop proceedings 2002 P2, 3.*

[11] Livingstone D. and Miranda E. "Composition for Ubiquitous Responsive Sound Environments". *In Proceedings of the International Computer Music Conference 1-6 November, 2004 University of Miami, Florida, USA*

[12] Livingstone D. (1998) "The Space between the Assumed Real and The Digital Virtual". *In Ascott, R. Reframing Consciousness – Art, Mind & Technology Intellect Books pp138-143*

---

[i] *Composer and Listener Objects detailed here form part of a larger integrated system included in proceedings ICMC 2004* [11]

[ii] *Details of this symbol based adaptive tracking system are included in proceedings ICMC 2004* [11]