# DIFFUSION EVOLVED: NEW MUSICAL INTERFACES APPLIED TO DIFFUSION PERFORMANCE

By

## BRIDGET DOUGHERTY JOHNSON

An exegesis

submitted to the New Zealand School of Music

in partial fulfilment of the requirements for the degree of Masters of the Musical Arts

in Composition

Supervisor: Dr. Ajay Kapur

Secondary Supervisor: Dr. Dugal McKinnon

2013

# Acknowledgements

There are many people without whom this work would not have been possible. To all those who have been involved I would like to express my gratitude but in particular:

To Ajay Kapur and Dugal McKinnon. You supported this project from its conception. For always encouraging my ideas no matter how ridiculous or achievable they were.

To Florian Hollerweger. If you hadn't shared your experience and expertise in spatialisation, this project would still be just a good idea, and for teaching me that math is fun.

To Blake Johnston. For your many hours, coding, rehearsing, performing, filming, calibrating and collaborating. It's been an absolute pleasure to work with you, and I look forward to more in the future.

To Owen Vallis and Jordan Hochenbaum. For your encouragement, patience and technical support throughout all my work with the *Bricktable*. Also for building it!

To Jason Wright. For the beats, beer, coffee, chats, installations, furniture and gear movement, photography and funk. You've been such a huge part of my development as an artist and as a person. If there were more people in the world like you, it would be a far better place. Also I ate all your lollies.

To Ane and Jon, George and Wai, and team Dougherty. You have become my family away from home, I love you all.

To my loving family; Mum, Dad, Tim, Cian and Madeline. For a while there we thought this would never happen. For your support through the good times and particularly through the bad, for your strength, encouragement and for showing me I would always have you guys no matter what, and that nothing else is as important.

And finally to James Wassell Murphy.........

if there were words, I would write them.

# Table of Contents

# Contents

# Table of Figures

# List of DVD Portfolio Contents

- High resolution code diagrams

- nebular performance video

- nebular performance 8 channel interleaved audio files, configured as per figure 17.1

- PDF Document of Exegesis:
*Diffusion Evolved: New Musical Interfaces Applied to Diffusion Performance*

- *tactile.space* demonstration video

# Glossary

**Audio Object** – the visual representation of each audio file on the GUI.

**Control Area** – the circular area of the GUI inside the representation of the speakers. Within this area the movement of objects affects the audio in real-time. Outside of this space, objects may move freely without affecting the audio.

**Fiducial** – in computer science image technology the term fiducial refers to an object in the field of view. In this case, it specifically refers to a physical object placed on the touch table surface.

**Gain Factor** – rather than being a direct gain, the gain factor is a calculated value that is then multiplied by the audio signal, to give the final gain of each speaker.

**GUI** – the graphical user interface. The visual aspect of the interface with which performers interacts, and via which they reieve visual feedback.

**Listening Point** – the representation of the very centre of the speaker array. Generally this is centred around the position where the performer is standing.

**Phantom Source** – the phenomenon whereby two speakers' gains are weighted to create the perception of the sound emanating from a specific point source location between the two speakers.

**Regular speaker arrays** – configurations of speakers that are entirely equidistant. That is, each speaker is the same distance from the centre and from its adjacent speakers.

**Spreader** – an Audio Object that has been spread out into an arc shape to create a wider sound source.

**SweetSpot** – the area that a listener can be situated inside and get an accurate depiction of the sound field. The width of the sweet spot is influenced by the number of, and positioning of, the speakers.

# Abstract

This exegesis takes a critical look at the performance paradigm of sound diffusion. In making a shift away from the sixty-year-old practice of performing on a mixing desk or other fader bank console, it proposes and outlines a goal towards intuitive and transparent relationships between performance gesture and spatial trajectories. This is achieved by a coupling of the two previously segmented fields within electroacoustic: spatialisation and interface design. This research explains how connections between the two fields and an embracing of contemporary technological developments, with a goal toward increasing the liveness and gestural input that currently limit sound diffusion practice, could extend the art form into a virtuosic and compelling gestural performance art. The exegesis introduces and describes the author's research and development of *tactile.space*, a new multi-touch tool developed on the *Bricktable* for live sound diffusion. *tactile.space* is intended as a contribution to the growing research area of user interfaces developed specifically for the performance of sound in space. It affords performers a new level of gestural interaction with the space of the concert hall and the audience members and redefines multiple standardised interactions between the performer and the space, the gesture, the audience, and the sound in a diffusion concert.

# Chapter 1 - Introduction

For over half a century, electroacoustic composers have regarded the spatial positioning and movement of sounds in their pieces as an element to be manipulated and performed in a concert setting. This act of sound diffusion has taken place on the largely unchanged performance interface of a mixing desk console. Meanwhile, methods of fixed media spatialisation in composition have fully embraced ongoing technological developments. This begs the question, why have the same technological advances not been incorporated into the gestural performance of sound diffusion? *tactile.space* is an attempt to realise some of the possibilities that contemporary technology could afford on the gestural performance of real-time movement of sound through space.

## 1.1 Motivation

*Changes in motion, pressure, velocity, light, gravity, skin conductivity or muscle tension, almost anything, can now become a 'musical controller.'*

*- Sergi Jordà (2007, 97)*

In an age where almost anything can become a musical controller, we need to think carefully about gestural performance input and how to meaningfully relate it to sonic output.

Now that the limitations of what can produce or affect the musical output in our performances are set only by the imagination, it has become even more important to seriously consider the relationship between performance action and sonic output, as well as the perceivable embodiment of these two factors by the audience member.

The main objective of the *tactile.space* project was to design a user interface that would afford the gestural micro-control of space. Of equal significance was the goal to look for ways in which the art of sound diffusion as a performance practice could be extended to develop an intuitive approach to performance and enhance performative transparency for audience members. It was hoped that any interface that successfully addressed these two problems would also lend itself to new forms or techniques in diffusion practice.

## 1.2 Overview

This first section has given us a brief introduction to the concepts explored throughout the exegesis; an overview of the content of the exegesis now follows. Figure 1 is a structure overview of the exegesis in diagrammatic form.

Chapter Two will first examine related works that served to catalyse the project. Firstly, significant historical developments in the spatialisation of electroacoustic art music are discussed, along with introductions of some of the most significant techniques used today. Secondly, sound diffusion as an art form is explored. Several notable diffusion systems and their performative approaches are introduced, and performance techniques of these systems are discussed.

In suggesting a shift to a more gestural performance interface for sound diffusion, Chapter Three presents the tools used to develop a new interface for diffusion artists. The chapter includes a brief history of multi-touch environments in music and the software and hardware involved in such systems. Chapter Three focuses on the reasons for choosing these specific tools and how they have contributed to the outcome of the project.

Chapter Four presents a case study on *tactile.space*, a multi-touch interface developed as part of the author's research in diffusion performance interface development. Specific features of the interface will be described, with emphasis placed on how these attributes allow a performer to have intuitive and gestural control, increasing the performative aspects of diffusion concerts. This chapter also explains the technical theory that has been implemented with each spatialisation technique.

Chapter Five is an evaluation of *tactile.space*. It includes a discussion of some of the diffusion techniques performers found to be successful and innovative in diffusion practice. Data presented in this chapter were collected from a user study taken by those who performed on the interface. Expanding relationships between performer and audience member are also examined with discussion of concert setups, and how these can be best designed for the gestural performance of diffusion on *tactile.space*.

The final chapter concludes the exegesis by commenting on the perceived success of such a tool for diffusion practice, and the implications it has for diffusion as an art form. Avenues for future research afforded by the project are also discussed, with a focus on diffusion performance research. Finally, uses of *tactile.space* outside the specific diffusion paradigm are suggested.

Diffusion Evolved: New Musical Interfaces Applied to Diffusion Performance

Chapter 2: Spatialisation History and Developments
* Approaches To Spatialisation
* Spatialisation As An Art Form
* Diffusion Systems

Chapter 3: Tools For Multi-Touch Computing
* Open Source Systems For Multi-Touch Computing
* The Bricktable

Chapter 4: Description of New Diffusion Tool: tactile.space
* Configurability
* Modes
* Source Spreading
* The GUI
* Speaker Routing
* Live Input

Chapter 5: Evaluation of tactile.space
* User Study
* Concert Setting

**Figure 1: Exegesis Overview**

# Chapter 2 - Related Work: The Development of Sound Diffusion

*The trouble with space is it's the whole piece. It's the sounds and everything. Space is the whole thing.*

*– Denis Smalley (Austin 2000, 18)*

Electronic art music has always been concerned with the space it encapsulates. Throughout history there have been many approaches to the aesthetic engagement with space and the possibilities for performance of spatial elements. This chapter gives an introduction to the means by which composers can engage with space in multi-channel works. A specific focus is placed on the relevance of diffusion practice.

A brief history of the developments of multi-channel approaches to spatialisation is introduced, along with a discussion of the pros and cons of the most relevant techniques for spatialisation. The chapter will then present a discussion of the artistic qualities and considerations of spatialisation as a performative art; the final section highlights existing systems that have been developed for sound diffusion concerts.

## 2.1 Approaches to Spatialisation

Electronic music has explored multiple techniques for the musical expression of sound in space. This section will introduce some historical developments in approaches to the positioning of sound through loudspeaker arrays. The techniques described are not only relevant to sound art but to commercial audio playback and science.

### 2.1.1 Amplitude Panning

Spatialisation techniques were originally explored in contexts other than those of electroacoustic art music. In a 1940 concert at Carnegie Hall, Frank Jewet, president of major audio company, Bell Laboratories, explained to audiences

> three loudspeakers were hung from the roof, one at either side of the stage and one in the middle. Three large amplifiers in the hall's basement would carry each track to one of the loudspeakers, creating a 'stereophonic' effect and immersing the listeners in a sort of musical 'third-dimension' that would allow them to hear spatial relations of the various musicians at the time the recordings were made. The recorded music would be rendered absolutely life-like.
>
> - Frank Jewet, (Milner 2009, 50).

This was the introduction in the United States to stereophonic sound systems. By 1940, though, Alan Blumlein had already patented his ideas and techniques for stereo sound in the United Kingdom (Milner 2009).

From stereo, spatialisation techniques evolved to quadraphonic systems, subsequently growing to include various forms of octophony and the more commercial Dolby 5.1 and 7.1 surround sound setups. Modern speaker orchestras often consist of well over one hundred individual loudspeakers. As soon as multi-channel systems became popularised, the question arose: how can performers meaningfully use and control all of the speakers in such an array, in an expressive and musical way? It is this question that drives this author's research and the development of *tactile.space*. However, *tactile.space* is just one new approach to achieving a meaningful relationship between sound and space. There have been prior attempts to do this, some of which will be discussed below.

With stereophonic sound, it quickly became apparent that one could not only position sounds at the far ends of the poles created by the distance between the two speakers, but one could also create 'phantom sources' anywhere within the two speakers' field. A phantom source is one where the gain of each of the speakers is weighted such that the sound source appears to be coming from a 'phantom' speaker somewhere between the two physical speakers. With the discovery of phantom sources, a new field of study emerged, focusing on the mathematics behind the accurate positioning of such phantom sources. Research in this area included a cognitive approach to human hearing and will be discussed in greater depth in section 4.2.2.1. Stereo techniques also form the basis of many diffusion systems, which will be discussed shortly.

In the 1990s, a technique known as vector base amplitude panning (VBAP) emerged; the technique served to further the principles of stereophonic panning (Pulkki 1997). VBAP expanded the notion of frontal image 'stereo' into 'pantophonic (two-dimensional) pair-wise panning,' and subsequently into triangulation algorithms for source positioning across three-dimensional speaker arrays. As pair-wise panning is implemented into the *tactile.space* project, it will be discussed in greater detail in section 4.2.3.

## 2.1.2 Holophonic Sound Field Production

Research into meaningful uses of multi-channel speaker arrays followed many directions. VBAP represents the direction of discrete source positioning. Another popular direction was that of re-creating an entire sound field in the most accurate way possible, a technique known as the holophonic approach to spatialisation. The idea was first introduced by Michael Gerzon in the 1970s and is known as Ambisonics (Gerzon 1977). This first instantiation of ambisonic sound field reproductions are known today as first order Ambisonics. Ambisonic techniques were then widely extended in the 1990s by many researchers, who built upon Gerzon's first-order into higher-order Ambisonics.

With each higher order of Ambisonics, a more accurate depiction of the sound field can be obtained. With each higher order, there is need for a greater number of loudspeakers to reproduce the sound. The use of large arrays of loudspeakers led to explorations in 3-dimensional speaker layouts, a paradigm currently gaining popularity in spatialisation research institutions such as SARC in Belfast, Ireland and CREATE in Santa Barbara, USA.

Ambisonic audio files use a unique audio file format that enables them to be transferred throughout different systems and maintain their accurate depiction of a holophonic sound field. There are two separate stages to the format: encoding and decoding. The encoding stage stores the audio files with the spatial positioning information embedded into them. This stage is completely independent of the decoding process, which uses the information of physical speaker locations to transmit the spatial information across the speaker array (Jérôme and Sébastien 2004). One of the benefits of Ambisonics is that it is a very portable format: it does not rely on knowing the configuration of loudspeakers at the time of encoding, as is standard in Dolby Digital or other surround sound formats. This also means ambisonic files can be decoded to varying speaker numbers and configurations. In spite of these advantages, Ambisonics only achieves realistic spatialisation with regular speaker arrangements. Furthermore, there is a minimum speaker number necessary to reproduce each order of Ambisonics, the higher the order the greater the minimum number of speaker required but the more accurate the sound field depiction will be. Each higher order of Ambisonics also achieves a wider sweet spot where the sound field depiction is perceived most accurately.

There are some disadvantages in current ambisonic techniques: Ambisonics does not allow for any sense of distance to be encoded in the audio; instead, all sounds will appear to be emanating from the edge of the speaker array unless digital signal processing (DSP) is in place to counteract this occurrence. Ambisonics is also very reliant on the so-called sweet-spot: the sound field is often only represented accurately to a select few who sit right in the centre of the speaker array.

Another approach to spatialisation, which also takes a holophonic stance and aims to recreate an entire sound field, is wave field synthesis (WFS). WFS is built on the Huygens-Fresnel Principle, which states that any point on a wave front can be understood as the source of a secondary wave front. In the context of audio, this means that it is possible to recreate the radiation pattern of any sound source with secondary sources (i.e. loud-speakers). The Kirchhoff-Helmholtz Integral furthers the competence of WFS to add the mathematics behind the theory that one can capture and also reproduce a sound field by determining the pressure and pressure gradient on that sound field's surface (Jérôme, Rozenn, and Sébastien 2003). In order to do this in practice, a very large number of speakers surrounding a space is required, with a minute distance between the speakers. Hence this approach is not practical for most setups, however it has been explored and is in use in a few institutions, mostly throughout Europe and the United Kingdom[*].

The techniques so far discussed are all highly relevant to the field of spatialisation in composition of electroacoustic works, but they do not address the issue of spatial performance.

**2.2 Spatialisation As An Art Form**

One would expect that with all the developments of spatialisation techniques over the last half-century, this technological advancement would have been mirrored in performance interface development. However, this is not necessarily the case.

From the early 1950s, the French schools of acousmatic music started to think about how sounds could be spread across multi-channel speaker arrays as a performance technique. In 1951, Pierre Schaeffer and Pierre Henry debuted their performance interface *potentiometer d'espace*. This was arguably the first tool for live diffusion, as it allowed the real-time movement of sound across a tetrahedral speaker array through the use of potentiometers (basic vertical faders) that controlled the direct gain of each speaker (Barrett 2007). This performance catalysed a long tradition that is still very much alive today: that of the live sound diffusion concert.

---

[*] For a deeper examination of both Ambisonics and Wave Field Synthesis please refer to (Jérôme, Rozenn, and Sébastien 2003).

Diffusion performance generally consists of the real-time spreading of a stereo audio file composition across a number of loudspeakers within a concert setting. The role of the performer is to dynamically adjust the volumes of each speaker or sets of speakers by controlling a fader bank, on which each fader corresponds directly to the gain of its respective speaker. Where a significantly large number of speakers are being used, they will often be grouped in various ways such that the performer controls each multiple-speaker group by adjusting one fader. The mapping and grouping of these faders informs the possibilities for spatial trajectories within a given concert. In this regard, today's diffusion concerts are performed with the same technology (the potentiometer) that Schaeffer and Henry used in 1951: the performance interaction has not moved beyond the fader bank. The performance on the user interface of a mixing console lacks any evident relationship between the movement of sliding a fader and the resultant change in spatialisation as listener or performer, one can not directly map the space of the performance to the space of the diffusions gestures.

Different composers see diffusion in different ways. Simon Emmerson, author of *Living Electronic Music*, views diffusion as a composer's recognition that the studio environment (where the piece is conceived) will never be the same as the concert hall (where the piece is performed) (Emmerson 2007). Therefore, diffusion is a way to re-interpret the piece to allow it to make sense spatially in its new environment. As such, performances using diffusion represent an active interpretation of the piece, in the same way that any classical music performer will always subtly give a personal interpretation of a piece in their own performance. The diffusion artist too is personally interpreting the piece through his/her spatial performance.

Denis Smalley, one of today's most notable diffusion artists, also views diffusion as an interpretation of a new space. Smalley believes that in a live diffusion the concert hall is brought to life through attributes of the piece itself. Without an active diffusion, each piece will be set to activate the room in which it was composed, the diffusion is what takes

the sound out of the individual speaker and makes it an activation of the concert hall as an entity of space. In discussing performance practice, Smalley suggests that one should never seek to contradict what is composed into spaces within the sounds or within the piece, rather one should seek to enhance them through the spatialisation of the piece as a whole (Austin 2002).

It is these understandings of what a diffusion performance is, or what it has the potential to be, that drove *tactile.space* as a specific tool for gestural spatialisation. As noted above, diffusion is currently still performed with a set of vertical potentiometers (a fader bank), as it was by Schaeffer and Henry in 1951. It is the author's belief that with development of the performance interface, the intricate understandings of spatial interpretation — as articulated by Smalley and others — will become not only more transparent to audiences, but also will allow a more intuitive and expressive performance by the artist. This will initiate a feedback loop with the ultimate result being a higher level of sophistication and virtuosity in the art of diffusion.

## 2.3 Diffusion Systems

While relatively little has changed in their user interfaces, diffusion systems as a whole have seen much technical development in the decades since their invention. As systems became capable of handling multiple speakers at once, institutions began to build loud-speaker orchestras. These orchestras developed over time from around 20 speakers in the 1970s to many contemporary orchestras easily exceeding 100 speakers. The first major diffusion developments (after Schaeffer's 1951 *potentiometer d'espace*) occurred in the 1970s when large-scale diffusion systems were developed. The early diffusion systems were base on the Acousmonium, a loudspeaker orchestra developed by the French GRM research group (the same group that Schaeffer and Henry had been working with) (Desantos, Roads, and Bayle 1997). It was through working with the Acousmonium that Denis Smalley was able to introduce the loudspeaker orchestra to the United Kingdom (Austin 2002). The subsequent subsections will describe firstly two pioneering systems that have evolved

over time, and then two more recent diffusion systems that attempt to expand the possible sonic trajectories in sound diffusion.

### 2.3.1 The Gmebaphone
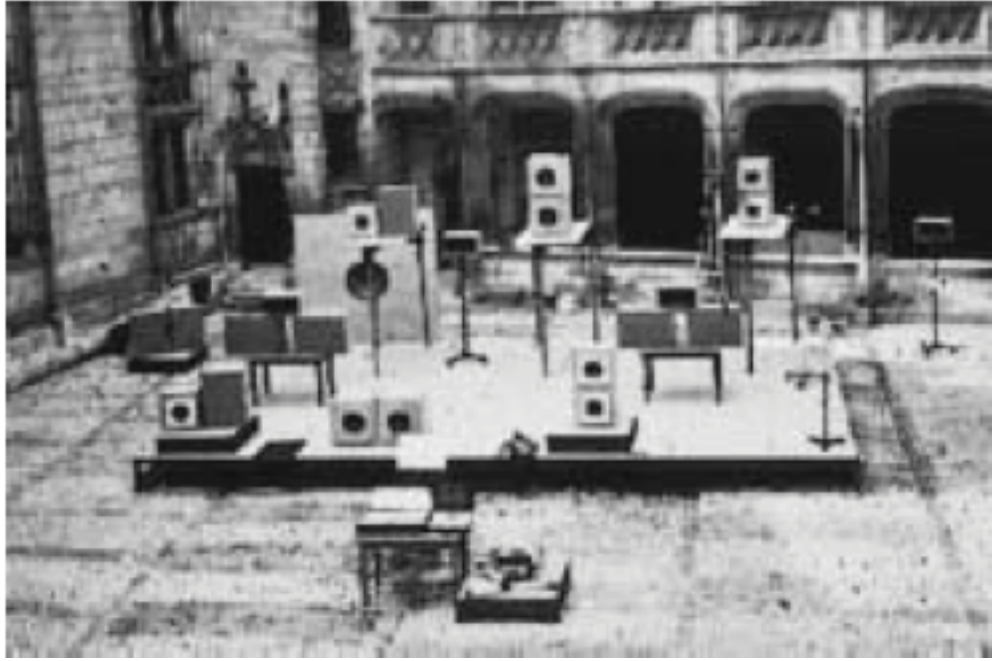


**Figure 2: The 1973 Gmebaphone (Clozier 2001, 82)**

One example of a pioneering diffusion system is the Gmebaphone. Developed at the Institut International de Musique Electroacoustique de Bourges (IMEB), it was first showcased in 1973 at the Third International Festival of Experimental Music (Clozier 2001). The system is still in use today, and has undergone numerous developments as the technology and the techniques of diffusion artists and researchers have evolved.

The redesign most relevant to this exegesis was that of the Gmebaphone 6, realised in 1997. While the system had already incorporated some digital technology in its 1992-3 instrument, the 1997 version featured a completely new digital console. This version was renamed the Cybernèphone because of its integration of computer systems: it was capable of being controlled remotely over a network and featured computer-assisted memory, which allowed a composer to pre-program the diffusion to enact specified sonic trajectories. As

a digital system, the Cybernèphone includes two control screens; however, these are used only for graphical feedback, the performer's physical interaction taking place using traditional computing methods such as mouse and keyboard, or through the traditional diffusion physical interaction of controlling faders. While the Cybernèphone utilises advanced computing techniques, the performance interaction remains largely unchanged from the earliest methods of diffusion performance.

### 2.3.2 BEAST

The Birmingham Electro-Acoustic Sound Theatre (BEAST) is another renowned diffusion system that has undergone numerous developments since its inception. BEAST was founded in 1982 at the University of Birmingham by Jonty Harrison (Harrison and Wilson 2010). The system was originally based upon the French GRM Acousmonium (Gayou 2007), with the system's main purpose being the diffusion of stereo and multi-channel electroacoustic works. BEAST has always sought to be a highly modular system, giving any performer or composer the freedom to customise the system to the specific needs of their piece and the particular performance environment. This is a feature that has inspired the modularity of *tactile.space*, as discussed in Section 4.1. The modularity is integral to the system's success. As discussed by Harrison (1999), the configurability encourages the concept that a diffusion should be an active interpretation of a performance space.

The standard mixing desk console was designed for multiple inputs to be actively merged to fewer outputs for audio recording and production. The opposite of this is desired in a diffusion system. Generally a diffusion will require minimal inputs (most commonly two but possibly up to eight) to be spread across a large number of outputs. In light of this problematic inverted use of the standard mixing desk, engineers at BEAST have designed a custom mixing desk console that balances the input and output controls in a manner more suitable to sound diffusion demands (Mooney 2005). The advantage of the customised console is that the user exhibits greater control and configurability of the output mappings to a greater number of loudspeakers.

The BEAST's system blurs the lines between composition and performance. The spatial aspects are somewhat composed and embedded into the piece but are also re-assessed in the live setting and worked specifically for a particular piece to be brought to its full potential through the incorporation of the characteristics of the individual concert setting. While the BEAST system, like the Cybernèphone, incorporates graphical feedback screens and computer-assisted diffusion techniques, the gestural interactions by the performer are again limited to mouse and keyboard computer input and motorised mixing desk faders. Because of this mode of interaction, the spatialisation gestures that can be potentially performed live are limited by the configuration and mapping of the fader banks. In James Mooney's thesis on sound diffusion systems, in evaluating BEAST he states,

> the biggest criticism is [the] relatively poor interface ergonomics, stemming from the fact that diffusion must be executed on a one-fader-to-one-loudspeaker [or group of loudspeakers] basis

> - James Mooney (2005, 206).

A common factor among both these systems is that while they focus heavily on the technical dispersion of sound, none of them have conducted any significant research into the physical gestures controlling the output of this spatial movement. Hence we see that in spite of the technical progress of the sound systems, the restriction that the performance interface — consisting of fader banks and standard keyboard/mouse — places on the possibilities for performance of sonic trajectories.

### 2.3.3 GSMAX

This is not the first research to note fader bank ergonomic limitations as problematic and detrimental to the development of diffusion art. Traux (1999) addresses the issue of limited trajectories in diffusion techniques and offers the GSMAX software as an approach to minimising some of these limitations. Traux summarises a few of his concerns with diffusion practice as follows:

1. A two-channel bottleneck where spatialisation is limited by the use of stereo audio files.

2. The limitations of manual control; as mentioned, there exist many limitations on spatial movement by the mapping and configuration of a mixing desk.

3. A lack of available rehearsal time not only in performance space but with a performance system as well.

While Traux's points are valid, it cannot be forgotten that some of the aforementioned concerns are addressed in the computer-assisted diffusion software platforms that exist within the Cybernèphone and BEAST systems, both of which allow up to 8-channel input and are available to composers before the time of rehearsal.

The GSMAX software uses an 8x8 matrix-style graphical user interface built in Cycling 74's visual programming environment Max/MSP, to map up to eight audio tracks to a range of predetermined sonic trajectories that would be unachievable with a mixing desk and fader bank. Users retain real-time control with the ability at any point to overwrite the movements and program new ones on the fly. Traux points out that

> to extend diffusion techniques to multiple channels, computer control is required since manual control with two hands is limited in the two-channel model. Smooth trajectories are also difficult to control with stereo, and those in contrary motion are physically impossible to control with a conventional mixer
>
> - Barry Traux (1999, 145).

While the development of this software, and others like it, certainly recognise and seek to overcome some of the limitations in traditional diffusion setups, it is pertinent to note that even such sophisticated software advances only computer-assisted spatialisation. Performance interaction is once again limited to mixing desk fader banks or mouse input.

### 2.3.4 M2

The M2 Diffusion system was developed by Adrian Moore, Dave Moore, and James

Mooney at the University of Sheffield. The system was introduced to the community at the International Computer Music Conference (ICMC) in 2004 (Moore, Moore, and Mooney 2004). The M2 system was developed as a specific response to the limitations of what were the current diffusion systems. M2 addresses specifically the issues of portability and modularity of a system. A common critique of diffusion systems is the relationship of one fader to specifically control the gain of one, or one group of, loudspeakers. In response to this, the M2 system features use of a MIDI fader bank, rather than a traditional mixing console, as the user interface. This interface, whilst affording only the same physical interaction from the performer, does allow for a dynamic mapping of the gestures to sonic outputs. As with BEAST, Cybernèphone, and GSMAX, specific software has been produced to run the system, in this case the SuperDiffuse Client and Server (Mooney 2005). With this software a performer can configure the 32 MIDI faders to control whichever parameters they wish. As the physical fader bank is an abstraction of the system there is nothing to suggest that this needs to be a conventional link between fader and speaker gain (this however, is entirely possible and in most cases highly probable). Further to the customisable fader mapping, a performer may map more than one parameter to each fader, allowing for a wider array of possibilities for real-time sonic trajectories.

The M2 system does not restrict the live performance to that of spatial control based on speaker gains. The user interface can be mapped to a number of effects parameters, therefore blurring the line between composition and performance and expanding the notions of a traditional diffusion past a real-time engagement of only space. However, it is important to note that the DSP effects are generally used for their spatial attributes: phasing, frequency shifting for Doppler effect etc. Although the system is potentially capable of more than artistic performance of space, it is this type of performance that the system was specifically conceived for, and is in use as.

One step towards potential future research in the diffusion paradigm that the M2 system represents is the notion that a diffusion user interface should be specifically designed as a diffusion user interface. In discussing the mixing desk console as a diffusion device,

Mooney reminds us that,

> ultimately, the technology [the mixing desk] has not been designed specifically for sound diffusion, and is therefore not as well suited for this purpose as might be desirable
>
> - James Mooney (2005, 199)

Arguably, the use of a bank of MIDI faders is not necessarily a significant progression towards the design of new performance tools specifically for sound diffusion. However, it is certainly a step in this direction. Research into the evaluation of diffusion systems and development of new systems that address the issues of configurability and performance gestural limitations as described by both Mooney and Traux, have laid the groundwork for a new field of diffusion research to evolve exploring relationships between gesture and space.

## 2.4 Summary

This chapter has shown that although the field of electronic art music has encouraged a wide range of development in the techniques employed by developers of diffusion systems, performers have been content to ignore the development of gestural interaction. Furthermore this interaction has continually hindered the creative possibilities for the manipulation and activation of spatial trajectories. Systems such as BEAST, the Cybernèphone, the GSMAX, and the M2 have all developed software to expand the range of spatial possibilities in sound diffusion, but in doing so have left a gap in the field for research into the gestural input of the diffusion artist. This gap has become increasingly problematic as other forms of sound art embrace new technologies and redefine the relationship between performers' physical action and the sonic output of their work (Croft 2007): diffusion art is being left behind. This thesis shows that the design of performance interfaces specifically for the paradigm of sound diffusion can enhance the complexity of sonic trajectories achieved through performance gestures, thus increasing relationships between the performer and audience member.

# Chapter 3 - Tools

In an attempt to address some of the limitations of the diffusion paradigm outlined in Chapter Two, tools that have potential to encourage a holistic evolution of a gestural practice such as diffusion will now be introduced. Multi-touch computing is a relatively new technological advancement but has already proved its worth as an expressive platform upon which musical interfaces can be built. The first section of this chapter will explore how musical applications have been developed for this technology; a focus will be placed on how this exhibits strong possibilities for its development as a tool for the gestural performance of space. This is followed by an explanation of how a tangible tabletop surface operates: the hardware and software of such a system are discussed. In doing so, *Bricktable* (Hochenbaum and Vallis 2009), the specific tabletop platform that has been used for the development of this research, is introduced.

## 3.1 Open Source Systems For Multi-Touch Computing

Throughout the late 1990s and early 2000s, many projects developed by designers, engineers, and computer scientists explored the possibilities of multi-touch surfaces. The

commercial success of the technology is evident in the popularity of the iPhone[*], iPad, and Android[**] devices. It did not take long for composers, performers, and live electronic musicians to see a potential to exploit multi-touch technology, adopting it specifically for sound installations and live electronic performance.

Some early examples of successful tabletop interfaces for musical performance and installation were the AudioPad (Patten, Recht, and Ishii 2002), and Block Jam (Newton-Dunn, Nakano, and Gibson 2003). Both of these interfaces were presented at the New Interfaces for Musical Expression (NIME) conference, an international community of electronic luthiers, and paved the way for tangible user interaction in music. Both devices use physical objects placed on the tabletop surface to trigger audio samples and manipulate their musical parameters.

The *reacTable* (Jordà et al. 2005) has proven to be one of the most significant projects in multi-touch applications for live computer music performance. The first instance of the *reacTable* used fiducial objects placed on the table to transmit data to a modular synthesis unit for performance. Additionally, it was used for multi-user collaborations as an installation piece. The *reacTable* showed that multi-touch gesture-based interfaces possessed both the intuitive nature needed to increase the learnability of a new instrument and also the sophistication required by experienced live electronic musicians to create a meaningful expressive performance. It explored some of the complex mapping systems that such interaction affords, with a range of different fiducials transmitting data to synthesis parameters: some fiducials acted as oscillators; others served as modulation sources. The *reacTable* provided graphical feedback on the interface, explaining the relationships between fiducials and helping inexperienced users effectively map their gestures to craft their desired sonic results (Jordà 2008). Later instances of the *reacTable* explored the possibilities for the creation of Western tonal music (Jordà and Alonso 2006), and more recently its application as a tool for studio mixing (Carrascal and Jordà 2011).

---

[*]    http://www.apple.com/
[**]    http://www.android.com/

A major turning point in tabletop development was the addition of finger tracking. Upon its introduction, tabletop tracking interfaces were no longer limited to the use of fiducial objects: the reacTIVision vision tracking software was capable of recognising a finger's placement on the surface of the table (Kaltenbrunner 2009). The first step towards this extension of tracking capabilities involved the placing of small fiducial-like objects on ones fingertips. Shortly thereafter, though, fully functional finger tracking requiring no modification to the finger itself was introduced. One of the major advantages of finger tracking is that it allows interaction with virtual objects within a custom-built GUI. With a custom GUI, the object a user interacts with can be application-specific rather than merely system-specific. It also removes any limitations that access to said fiducials may cause; as such, there is no limit on the number of fingers that a table can recognise, providing users with a wider range of possibilities for multi-user interactions. The size of a single finger in comparison to a fiducial object also affords a higher degree of resolution and accuracy in touchable parameters, giving users finer control of the interface, and increasing possibilities for virtuosic performance.

The advent of finger tracking, combined with the release of multi-touch devies and the inclusion of multi-touch functionality in the trackpad of many personal computers, catalysed a new wave of gestural interaction. Research in the field now extends touch-based relationships to include intuitive human gestural movements and incorporates recognition of the constrictions of physical human form. Projects such as MStretchSynth and MDrumSynth (Scheil 2010) explore complex relationships between the points of touch data to provide a higher level of expressivity within musical controllers. Studies have also examined the effectiveness of these relationship-based mappings in comparison to traditional Cartesian-style x-y mappings (Johnston, Vallis, and Kapur 2012). These studies took into account software considerations of human gestural capabilities such as average handspan. Research of this nature has served to evolve the development of multi-touch surfaces from an experimental interface into a gestural, intuitive, and highly expressive musical instrument. Thus, an ideal platform for the development of a new performance interface for sound diffusion.

Sergi Jordà, member of the *reacTable* research group, draws parallels between musical multi-touch devices and the original non-musical intentions of multi-touch surfaces. These surfaces were seen as a way to effectively share data between users: common demonstrations for new surfaces are applications allowing the sharing of personal photos or visualisations of graphs. Such applications are prime examples of situations where users benefit from an easy way to share access to the data presented with others. The *reacTable* showed that this concept could be extended to the sharing of data between users or performers, thus making the tangible table-top surface an ideal platform for multi-user collaborations (Jordà 2008). The concept also exhibits potential for further extension, allowing data to be creatively shared between the performer and the audience. As discussed earlier, one of the problematic performance characteristics the paradigm of diffusion art has caused is a disconnect between performative gesture and sonic output. Hence, the performance interface of the touch table has great potential to resolve some of this dissonance as a performance tool for live diffusion, exhibiting transparent input/output relationships.

## 3.2 The *Bricktable*

The decision was made to pursue diffusion interface development using the *Bricktable*, a tangible table-top surface developed by Owen Vallis and Jordan Hochenbaum. Table-top surfaces have proved their place in electronic music through examples like the AudioPad and the *reacTable*. Their gestural capabilities and intuitive nature make this an ideal platform for development of gestural performance interactions. The hardware and software involved in running and developing for the *Bricktable* will be discussed in subsequent sections.

### 3.2.1 Hardware

The *Bricktable* uses the diffused illumination technique for vision tracking, allowing for detection of both fingers and objects. In a diffused illumination system, infrared light is shone upon an acrylic screen from inside the table structure. The light is reflected off the screen and diffuses throughout the structure; the opposite side of the screen then acts as the touch surface. When a finger touches the screen, the finger reflects significantly more light than is diffused amongst the rest of the screen. This enables the camera to track the presence and positioning of the finger (Kaltenbrunner and Bencina 2005). Unlike many commercial touch screen products, there is no limit to the amount of fingers that may be tracked at any one time. It is important that the lighting be infrared, as the diffused lighting needs to be of a different wavelength to that of the projector. In order for users to view the GUI (Graphical User Interface), it must be projected in the visible spectrum of light, therefore the illumination lighting must be infrared (Kaltenbrunner and Bencina 2005). The tracking software is then able to detect significant differences in the lighting when a finger is placed on the table's surface.
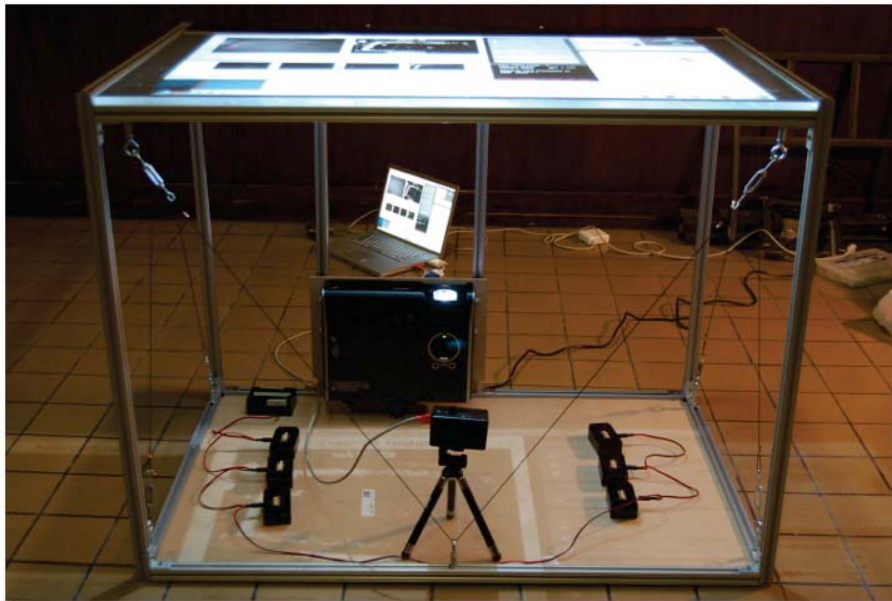


**Figure 3: The** *Bricktable*

The GUI is projected onto the acrylic tabletop surface from inside the structure. This al-

lows the user to feel as if he/she is touching virtual objects within the GUI. Fast accurate tracking gives the feeling of dragging an object across the screen; this mode of interaction has become popular with smart phone technology, and thus is becoming increasingly familiar to the general public.
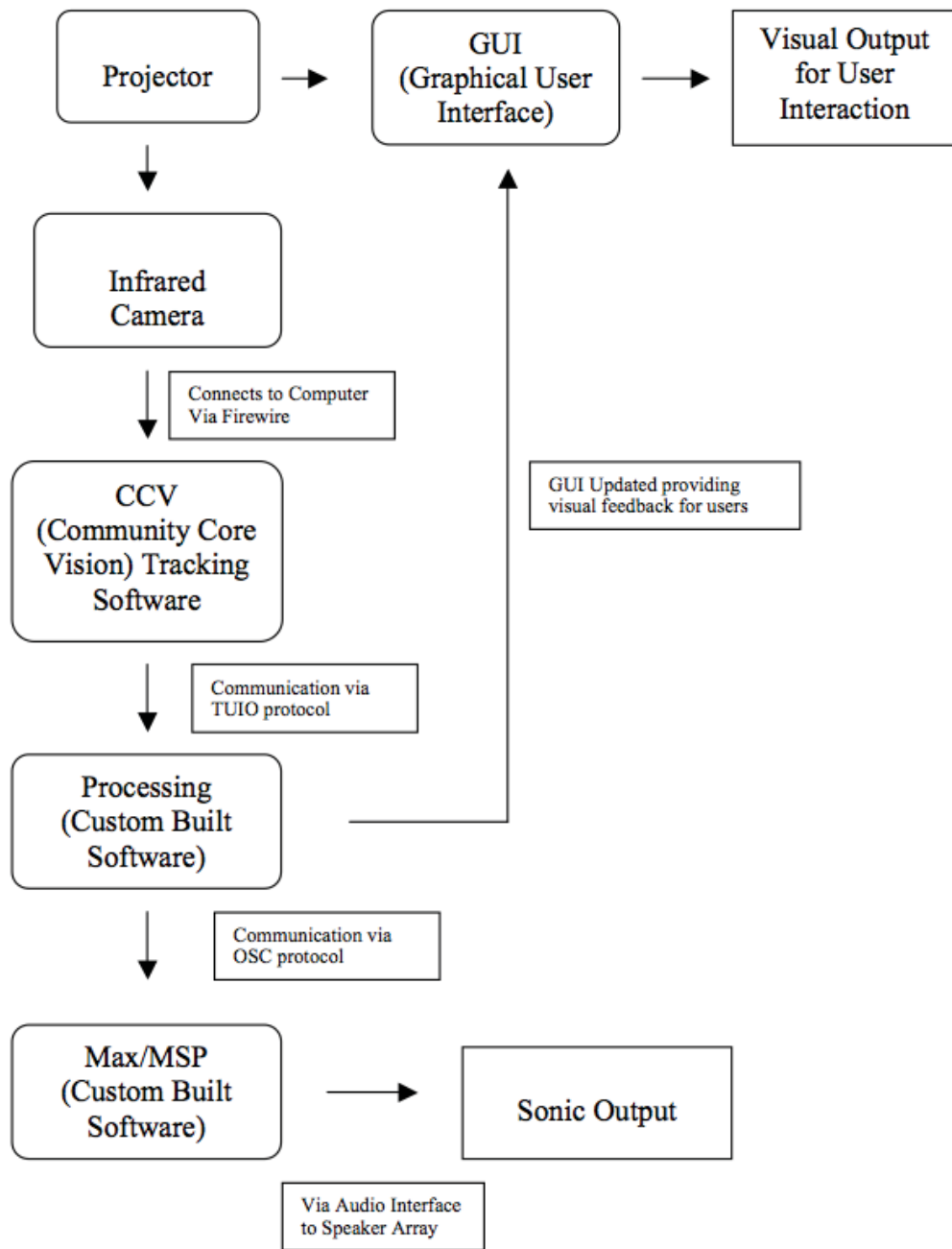
### 3.2.2 Software

The *Bricktable* makes use of the *reacTIVision* (Kaltenbrunner and Bencina 2005) vision tracking framework. The data chain begins with the open-source software Community Core Vision (CCV)[*]. CCV is the tracking software that recognises the touch data from the view of the camera and sends it via the Tangible User Interface Objects (TUIO) (Kaltenbrunner et al. 2005) protocol which is received in Processing[**] using the TUIO client library, see Figure 4.

The bulk of the custom-built software for *tactile.space* is written in Processing. Processing is a Java-based language that was developed by Ben Fry and Casey Reas specifically for artists as an introductory language into computer programming (Reas and Fry 2007). The program written in Processing is responsible for drawing and updating the GUI, for the processing of TUIO data, and updating the GUI based on this. The Processing-built application is also responsible for calculating values and sending them as Open Sound Control (OSC) (Wright, et al 2003) messages to Cycling74's Max/MSP[***] . Max/MSP deals with the audio, playing the audio files and receiving the appropriate gain factors to be sent to each speaker. Max/MSP is also responsible for a limited amount of DSP, specifically when *tactile.space* is being used in VBAP mode. The DSP is in place only in order to embed a sense of distance within the VBAP algorithm and a closer simulation of the realistic movement of sound through space. Both distance encoding and VBAP mode will be discussed in more depth in sections 4.5.2 and 4.2.3 respectively. Whilst very early installments of the program used the Minim[****] audio library inside of Processing to handle the audio playing,

---

[*]      NUI Group Community. "Community Core Vision (CCV)". http://ccv.nuigroup.com/
[**]     Fry, B. and Reas, C. Processing 1.5.1. http://processing.org/ 2010
[***]    Cycling 74. Max/MSP. http://cycling74.com/products/maxmspjitter/
[****]   Minim Audio Library. http://code.compartmental.net/tools/minim/

it very quickly became apparent that in order to create a meaningful multi-channel diffusion, a higher level of sophistication would be required from the audio unit, therefore prompting the development of a custom-built Max/MSP patch (see section 4.5).

```
┌──────────┐      ┌──────────────┐      ┌──────────────┐
│          │      │     GUI      │      │ Visual Output│
│ Projector│ ───► │(Graphical User│ ──► │  for User    │
│          │      │  Interface)  │      │ Interaction  │
└──────────┘      └──────────────┘      └──────────────┘
      │
      ▼
┌──────────┐
│ Infrared │
│  Camera  │
└──────────┘
      │   Connects to Computer
      │   Via Firewire
      ▼
┌──────────┐
│   CCV    │           GUI Updated providing
│(Community Core│       visual feedback for users
│Vision) Tracking│
│ Software │
└──────────┘
      │   Communication via
      │   TUIO protocol
      ▼
┌──────────┐
│Processing│
│(Custom Built│
│ Software)│
└──────────┘
      │   Communication via
      │   OSC protocol
      ▼
┌──────────┐      ┌──────────────┐
│ Max/MSP  │      │              │
│(Custom Built│ ──►│ Sonic Output │
│ Software)│      │              │
└──────────┘      └──────────────┘
      Via Audio Interface
      to Speaker Array
```

**Figure 4: *tactile.space* data flow**

The framework developed and supported by open source communities has ensured the ongoing development of touch table musical applications. Programming environments such as Processing and Max/MSP have made the development of these applications possible for those lacking computer science and software engineering backgrounds, allowing

artists direct input into the direction their technology takes. This has catalysed a rapid increase in the development of musical interfaces. The *Bricktable* is a prime example of the performer-turned-luthier taking control of the advancement of his or her performance tools.

The expressivity and gestural control that the *Bricktable* affords warrants an exploration into its capabilities as a performance interface, with the ultimate outcome of such explorations being the introduction of a new element to the diffusion performance paradigm, and a marriage of the fields of interface design and spatialisation performance.

# Chapter 4 - Implementation

*The control interface must ideally be fairly simple and intuitive, but nonetheless able to provide the performer with all the necessary means to achieving the musical communication 'in real time'.*

*- James Mooney (2005, 179)*

This chapter describes the features of the diffusion interface *tactile.space* and the way they have been developed to encourage transparent relationships between performance gesture and sonic output as well as to increase the possibilities for performance of spatialised sound. As Mooney suggests, the interface was designed to be as simple and gesturally intuitive as possible, whilst still allowing an expressive and highly configurable platform for performance. The possibilities for user configuration are explored, and suggestions for some reasons why a performer may make these decisions are provided. In discussing the modes of spatialisation that *tactile.space* is capable of, the theory behind these techniques is discussed, as are the methods by which such modes of spatialisation are adjusted to work most effectively in the *tactile.space* environment.

# 4.1 Configurability

The system architecture of *tactile.space* is such that the interface is highly configurable for individual performers' wants and needs. This will eventually allow it to be run on any multi-touch surface rather then as an institution- and system-specific application. It also allows the interface to quickly and easily cater to the specific technical and aesthetic needs of each composition, making a smooth transition possible between pieces in a concert setting without placing restrictions on possible configurations for the individual performer.
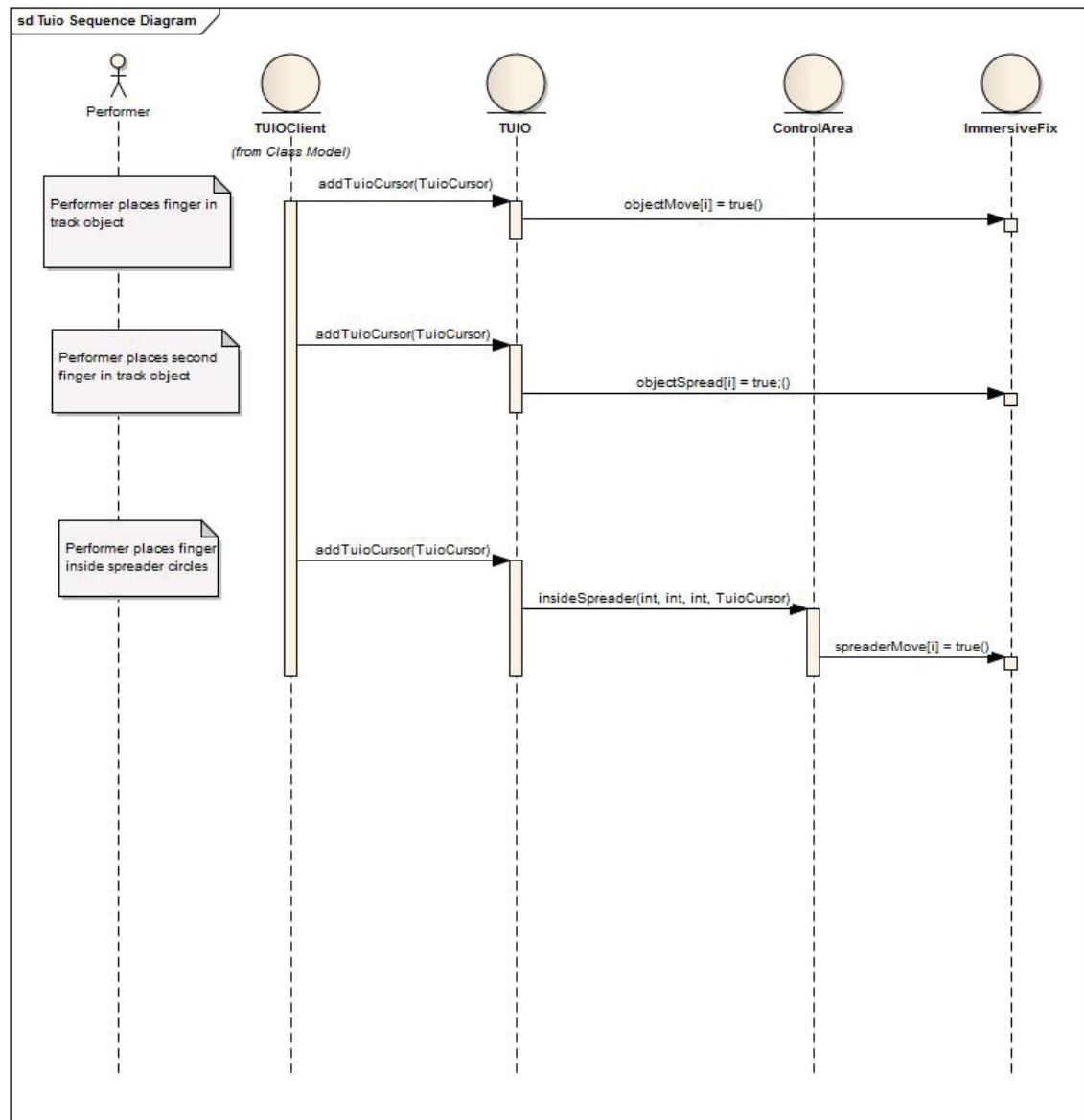
**Figure 5: Performance Interaction Sequence Diagram**

In its simplest form the user drags a visual representation of an audio track and positions it amongst a representation of an array of loudspeakers. Calculations are made to position that audio in the desired location. Users may also place two fingers inside an object to cause the object to spread and its location to come from a wider spatialised area rather than a direct source position. Once an object is a spreader the user may update size and position of the spread object (source spreading is explained further in section 4.3).

Figure 5 is a sequence diagram explaining these interactions and the effects they cause with in the Processing sketch that controls the interface. The first column shows each of the main interactions with audio objects a user can make, each of these interactions is then followed through showing the classes within the code that are updated and called based on this interaction. Each of the other columns represents classes within that code that are active. Immersive Fix is the main body of the code. See appendix two for further explanation of class relationships.
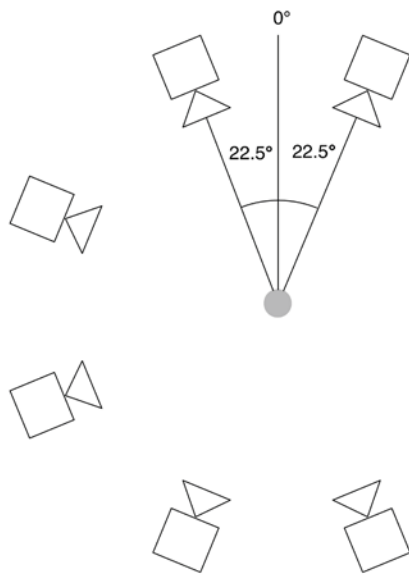


**Figure 6.1: European Octophony Standard**
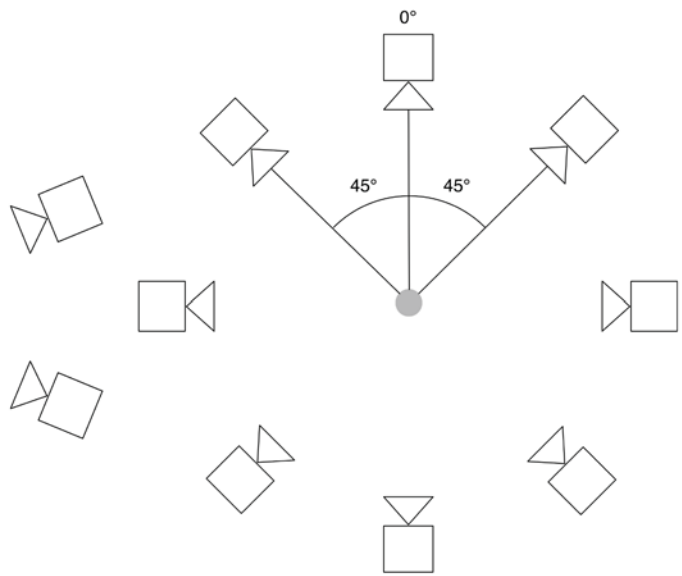


**Figure 6.2 US Octophony Standard**

To realise an adequate level of modularity within the system, the program is set up to work with any possible number of speakers, provided the amount of outputs are made available.

The 'numOfSpeakers' parameter in the Processing sketch can be adjusted to match the number of physical speakers in use. The user may then also choose between the European and US standards for speaker positioning. These standards for octophony are as follows: front speakers at 22.5° and 337.5° from listener (see Figure 6.1) in European configuration and a single speaker 0° to the listener in the U.S. standard positioning scheme (see Figure 6.2).

Users may also choose from different 'modes,' discussed below, which determine the type of spatialisation algorithms to be used; users also have the ability to turn on or off various features to customize the level of graphical user feedback.

## 4.2 Modes

For any new interface to be considered as a serious tool for live diffusion it first needs to be able to achieve all that is possible from a more traditional diffusion system. Without this, one could never expect artists of the established performance paradigm to consider a change to a new interface. Once this has been considered and achieved, new ideas for techniques that a new interaction device may afford on the performance practice can be explored. As such, *tactile.space* can be run in varying modes, allowing performers to choose their spatialisation techniques and algorithms. Performers input the mode they wish to use before the Processing program is executed. The sophistication and developmental stages of the program are evident in the development of each mode. As such, while the third mode is the most comprehensive, there could be aesthetic reasons that warrant the use of the quadraphonic panning mode and/or the stereo pairing mode. It is therefore left to the performer to choose their method of spatialisation depending both on the configuration of speakers available to them and on the aesthetic needs for their piece. A discussion of the individual modes follows in the subsequent subsections.

### 4.2.1 Quadraphonic Panning Mode

The quadraphonic panning mode was the first attempt at moving beyond two channels. This mode works specifically with four speakers and assumes an equidistant regular quadraphonic speaker array as shown in Figure 7. This mode assumes a speaker layout as described by Figure 17.1. This is a common speaker layout and is consistent amongst all modes of *tactile.space* (see section 4.6).
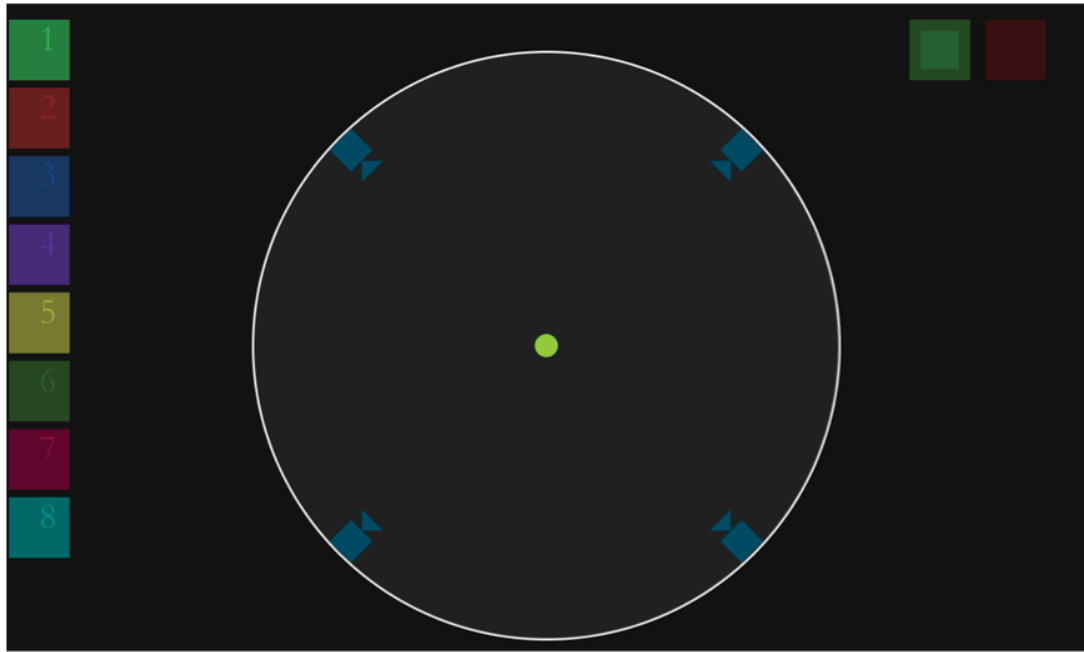


**Figure 7: Quadraphonic Panning Mode**

The quadraphonic panning mode was inspired by Ambisonics research; it involves all speakers playing all the time. Spreading the overall gain across the entire array and dynamically weighting the gain factor of each speaker based on the position of the audio object creates the spatial positioning and movement. The weighting of the gain of each speaker contributes to the perceived source position. The spatialisation algorithm makes use of a simple linear stereo technique to incorporate four speakers. The Cartesian positional data from within the control area is first scaled between 0 and 0.9 being the range of appropriate gain factors; next, a calculation of the inverse of this number is made. Provided four floating point numbers:

These numbers are then used to provide a positional spread between the four speakers and

$$scaledX = posX \text{ (scaled between 0 and 0.9)}$$
$$scaledY = posY \text{ (scaled between 0 and 0.9)}$$
$$inverseX = 0.9 - scaledX$$
$$inverseY = 0.9 - scaledY$$

send the data as a gain factor to each of the speakers.

This mode was enough to prove that the project had potential and that multi-touch was

$$speakerOne = |inverseX * inverseY|$$
$$speakerTwo = |scaledX * inverseY|$$
$$speakerThree = |inverseX * scaledY|$$
$$speakerFour = |scaledX * scaledY|$$

indeed an intuitive gestural way to move sound throughout a speaker array. However, it very quickly became apparent that a quadraphonic setup of this nature lacked the sophistication needed for the application to become successful, intriguing, and capable of affording the possibility of virtuosic performances. Additionally, the limitations for spatialisation that come with the use of only four speakers motivated the development of the modes discussed in following subsections.

**4.2.2 Stereo Pairing Mode**

As mentioned previously, it is important that the program can emulate current diffusion practice as well as encourage an extension of live spatialisation techniques. The traditional method of spatialising a split stereo file across numerous left/right pairs of speakers has already been discussed. It was this practice that informed the method of the stereo pairing mode in *tactile.space*.

*4.2.2.1 Stereo Pairing Mode Theory*

As the field of psychoacoustics and stereophony developed, it quickly became apparent that early ideas about stereophonic spread and source positioning were not as linear as seemed obvious. Whilst at first thought one might assume that if the left channel is turned up at the same rate as the right channel is turned down, the audio will be perceived as a smooth transition between the two speakers. Whilst this method, known as linear pan-

ning, will preserve the amplitude of the source signal, human ears are far more concerned with the intensity or power of a sound source rather than its amplitude (Rumsey and Mc-Cormick 2006) (Pierce 1999). As such, when using this method of linear panning we are left with the perception of a hole in the centre of the two speakers where our brains will not accurately localise the sound source. This is also affected by the conflict between the direction of the sound event (the physical sound source) and the direction of the auditory event (the perceived source of the sound) (Blauert 1997).

The aforementioned phenomenon occurs because of the way our brains localise sound sources. In order to determine the position of a sound source the human brain evaluates the difference between the incoming sources at each ear, these are known as binaural cues. Possibly the most influential cue our brain uses to decipher where a sound source is located is the Inter-aural Time Difference (ITD)— the difference in time between a source reaching the left and right ears. The ITD is closely followed in significance by the Inter-aural Level Difference[*] (ILD) (Matthews 1999a) (Blauert 1997). It is the ILD that *tactile. space*, in line with most other panning algorithms, exploits in order to have the listener perceive the specific spatial location of a sound source. In electronic music, this technique of exploitation is known as amplitude panning. While it comes in many forms, the most common use is in a frontal stereo spread.

As mentioned above, the assumed lineal conditioning of the left and right channels does not result in accurate source positioning because of human localisation cues and because of the deciphering of sound intensity rather than amplitude. Research in this area has proposed many ideas to counteract this inaccuracy in spatial perception; by using certain algorithms in order to attribute the appropriate gain factors for each speaker, sounds can more effectively be positioned such that they will be perceived accurately by listeners in the desired location between two speakers. These algorithms have become known within the field as pan pot laws and are now embedded into most commercial panning software

---

[*]     Sometimes known as the Inter-aural Intensity Difference or less accurately the Inter-aural Amplitude Difference.

or mixing desks (Griesinger 2002).

The two most common forms of pan pot laws are the square-root panning law and trigonometric panning laws. The square root panning law preserves the intensity (which our ear is most concerned with) of a sound as it is panned between two speakers. However, this algorithm tends to result in a breakdown towards the far left or far right edges of the spectrum. Our ears tend to perceive a smooth transition through the centre but then a sudden collapse into one speaker at the stereo fields extremities. Therefore this algorithm was not suitable for use in *tactile.space* and other options were considered.

*tactile.space* makes use of a version of the trigonometric panning principals known as the 'sine-cosine' pan law. Whilst research (Griesinger 2002) has proven that this may not necessarily be the most accurate positioning method, *tactile.space* needs to employ an algorithm that will be equally effective at all points in a circular array as well as with varying configurations of speakers. *tactile.space* can also not afford to make assumptions about the head position of the listener. The two more accurate laws, the sine law and the tangent law, both rely heavily on the listener facing a certain direction, be it at the sound source itself or at the 0 degree position (Pulkki 2001). Thus the decision was made for the more generic sine-cosine law, as shown below, to be implemented.

For the above equation $L$ and $R$ equal the left and right gain factors respectively, $\theta$ equals

$$L = cos\theta * i$$
$$R = sin\theta * i$$

the source angle, and $i$ equals input (the audio signal).

Another significant factor in the decision to choose the sine-cosine law was research that outlines the significant difference that frequency plays in the way humans localise sound sources. Certain pan pot laws are more accurate throughout different frequency bands (Griesinger 2002). However, the Processing program of *tactile.space* that implements the algorithms operates independently of any audio content. There is no way of knowing

which type of spectral content will be used by the performer at any time, furthermore *tactile.space* wishes to allow spectral freedom to composers. In light of this, some concessions had to be made when choosing and implementing the pan pot laws. The pan pot laws also rely on knowledge of the original source amplitude, another factor that would necessitate knowledge of the audio signal to express. Hence the concession is made to assume all audio signals are brought in as 0.8. The calculations are then made in Processing as weighted gain factors. The data is sent to Max/MSP, which multiplies the actual audio signal by the gain factor it receives. Thus, whilst not explicitly employing these laws with complete accuracy, the gains are calculated relative to each other and thus the source localisation is remarkably accurate with diffusions on *tactile.space*. The *tactile.space* implementation of the sine-cosine law is stated below:

```
gainLeft = (cos(radians(position)) * 0.8)
gainRight = (sin(radians(position)) * 0.8)
```

It cannot be forgotten in an analysis of accuracy in source positioning diffused by *tactile.space* that visual cues are far stronger then aural cues in the brain's recognition of source positioning (Matthews 1999b) (Blauert 1997). In light of this, it is important to attribute some of the success to the listeners' (and performers') view of the GUI. The perception of source location in *tactile.space* is so influenced by the visual cues of placing an audio object in a position relative to a representation of a speaker that in early versions of the program where less accurate source positioning was used the listener was generally unaware of the extent of the inaccuracies. This phenomenon was intensified by the gestural movements of the performer, who was at most risk of expressing an inability to perceive inaccuracies due to his or her physical gesture and visual cues overwriting the aural cues.

### 4.2.2.2 Stereo Pairing Mode Implementation

The stereo pairing mode works by dividing the speaker array vertically into a series of horizontal stereo pairs, as shown in Figure 8. When an audio object is present within the control area, its Cartesian positional data is first used to distinguish which set of vertical pairs the object is present within and then to determine its left/right position between said pair of speakers.
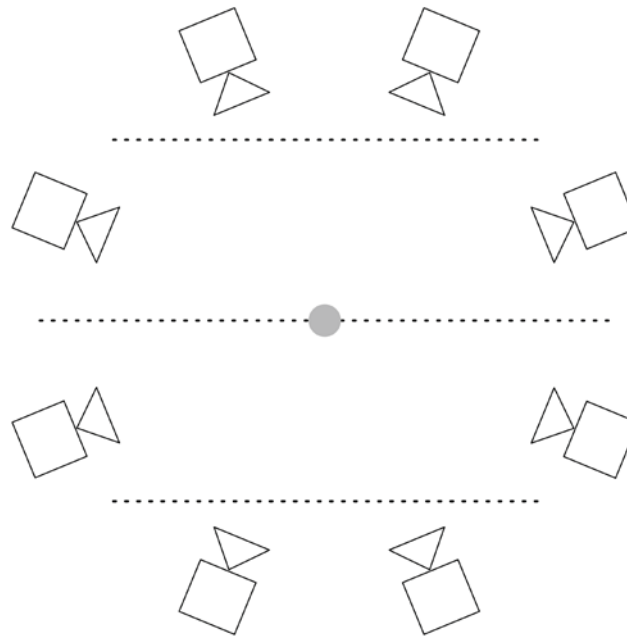
**Figure 8: Division of Stereo Pairs**

The position within the stereo spread is then sent to the trigonometric pan pot law algorithm to decipher appropriate gain factors to be sent to each speaker in the pair. As an audio object is moved from one speaker pair to the next, it is faded out of the original pair. The default fade rate is 500 milliseconds; this may be pre-determined, however, to any amount by the performer before executing the program.

Stereo Pairing mode will work with any even number of speakers. Whilst the GUI will always draw the speakers assuming an equidistant pantophonic array, the effectiveness of this mode is not limited to regular speaker layouts. The vertical division of space into each speaker pair will always occur equally amongst pairs. As the calculation of speaker gains in this case is not a direct result of the positioning of speakers, if a performer is able to conceptualise the movement of sound through their desired speaker configuration, the movement of sound will not be affected by the conflicting visualisation of speaker placement on the GUI. It is important to note in relation to irregular speaker configurations, the trigonometric pan pot laws are most effective for speaker pairs no wider then 90 degrees apart.

**4.2.3 VBAP Mode**

Control of stereo positioning provided interesting creative possibilities for the user in Stereo Pairing Mode and further solidified the use of multi-touch tools for performative live diffusion. However with the goal of encouraging intuitive gestural interaction, the physical positioning of audio objects by the performer in this way has a much stronger connection to phantom source positioning than it does to control of stereo spread. VBAP mode was designed as an exploration of the possibilities for dynamic phantom source positioning throughout a pantophonic speaker array.

*4.2.3.1 Vector Base Amplitude Panning Theory*

Vector Base Amplitude Panning (VBAP) is a spatialisation technique conceived by Ville Pulkki in the 1990s (Pulkki 1997). The technique extends what had already been achieved with frontal stereo panning into a two-dimensional ring of pair-wise panning. Speakers are divided into pairs in order to create phantom source positioning of a sound event. Unlike in a stereo pairing system, with VBAP each speaker is part of two pairs, one a pair with the speaker on its right and another with the speaker on its left, as shown in figure 9.
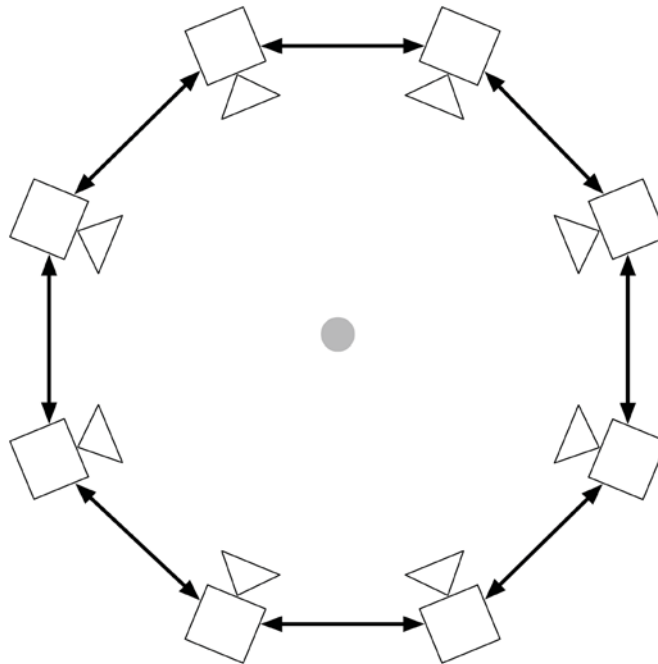
**Figure 9: Pairwise Panning**

When deciphering speaker gains, one must first look at the angle of the source position and find the speakers adjacent to the source. Once the appropriate pair of speakers has been determined, pan pot laws can be used to calculate the appropriate gain factors for each speaker in the pair in order to create a phantom source between them. VBAP relies on equidistant regular speaker arrays but works with any number of speakers, giving it the modularity required to work with the *tactile.space* aesthetic. One of the flaws with the VBAP spatialisation method is there is no way to encode a sense of distance within the source positioning. Thus, all sounds will be perceived at the edge of the speaker array. Resultantly, when the VBAP mode is selected, *tactile.space* employs its own distance encoding to condition audio files independently (see section 4.5.2).

VBAP is now operational within many spatialisation software tools, including the BEAST's and Cybernèphone's software packages and externals available for Max/MSP, Pure Data and SuperCollider. Pulkki furthered his research, extending this notion of pairwise panning to work with 3-D speaker arrays where we see a triangulation of speakers to spread sources throughout a hemisphere of loud speakers. The same research also included techniques as methods to implement varying spread or width of a source positions (Pulkki 1999), which is also implemented in *tactile.space* and discussed in section 4.3.

### 4.2.3.2 Vector Base Amplitude Panning Implementation

The third mode offered by *tactile.space* employs accurate and dynamic source positioning using vector-base amplitude panning methods. It is this mode that shows the strengths of *tactile.space*. Previously VBAP was a rendering algorithm used to embed spatial positions in composition. However, the principles of VBAP lend themselves especially well to live source positioning, which is a far less common form of diffusion. *tactile.space* uses the VBAP method to position by creating phantom sources at any point within the control area.

In VBAP mode, the Cartesian positional data is first reoriented to align the origin of the

system at the listening point. In Processing, the coordinate system is aligned so that the origin of the system is the top left corner of the screen as is the standard for computer systems (see figure 10.1). In order to create phantom source positions through VBAP spatialisation technique we must re-orient the system so the origin is in the middle of our control area; the listening position (see figure 10.2). For example, the Cartesian coordinates for the point shown in figure 10.1 would be (2,3). However the point in figure 10.2 which is in a comparative position spatially, would have the Cartesian coordinates of (-2,-2), hence the need for the re-orientation.
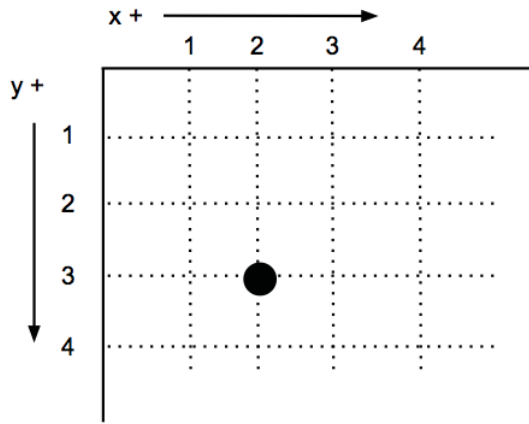


**Figure 10.1: Processing Coordinate System**
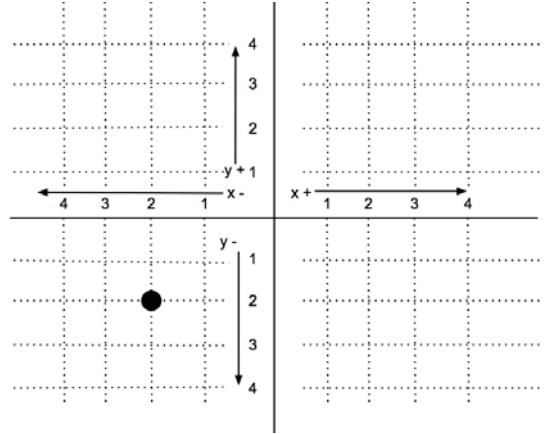


**Figure 10.2: Cartesian Coordinate System**

From here, the Cartesian data is converted to its equivalent polar coordinates (the angle and radius from the origin) using the following algorithms:

$$r = \sqrt{x^2 + y^2}$$

$$\Theta = 360 - ((atan2(x, y) - 360)\%360)$$

For the above equations $r$ equals radius, $\theta$ equals angle, and $x$ and $y$ equal their respective Cartesian values.

The equation for deciphering radius is in its common form. The angle calculations vary slightly from the most common forms of this equation in order to incorporate the correct functionality across all four quadrants of the coordinate system, also allowing the drawing of the arc for the spreader object (see section 4.3) in a clockwise rotation as is required by

Processing, opposite of a conventional anti-clockwise circular rotation in mathematics.

The next step is to analyse this position in relation to the position of the speakers. The program searches to find a match that is the speaker pair within which the audio object sits (as shown in bold in figure 11.2). Once the correct pair of speakers has been found, the angle is scaled to between 0 and 90 degrees, which allows it to be placed into our trigonometric pan pot law algorithm (section 4.2.2.1) and positioned within the two chosen speakers.
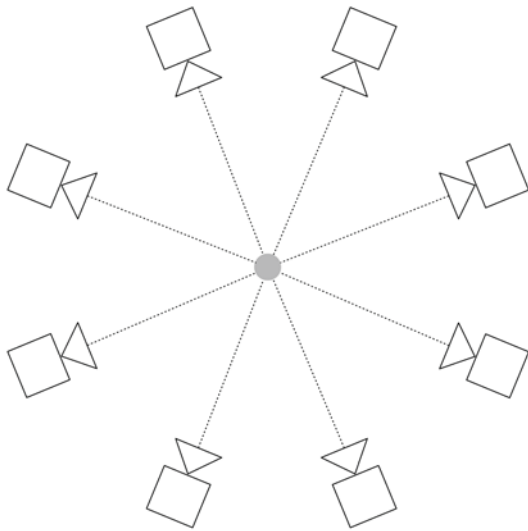
**Figure 11.1: VBAP Division of Adjacent Speakers**
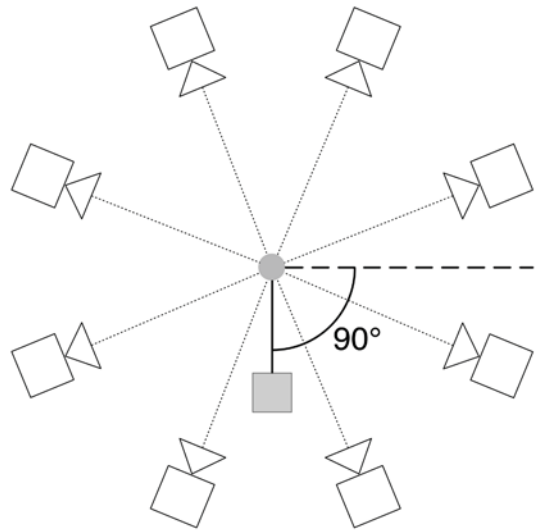
**Figure 11.2: Calculation of Radius and Angle**

Whilst this law was built on an understanding of the psychoacoustics behind spatial hearing in a frontal stereo field, it proves also to be highly effective when positioning sounds in a circle around the listener. This mode of spatialisation gives an accurate pantophonic sound field depiction, assuming the number of speakers is high enough that one can avoid any serious 'holes' in the array (generally at least 8 speakers are needed, but this number is dependant on their distance from the listening point).

As Processing requires the Cartesian positional data to define positions to draw the elements of its sketch, it was at times necessary to revert back to Cartesian space. In certain circumstances, such positions may not be defined by the presence of a finger. Therefore, a

conversion from polar to Cartesian coordinates is necessary. When required, the follow-

ing equations are used:

$$x = r(cos\theta)$$
$$y = r(sin\theta)$$

For the above equation *r* equals radius and *θ* equals the angle. After applying the equa-

tions, a reorientation of the origin of the system back to the top left corner (Processing's

origin point) is necessary.

Figure 12 explains the interaction of the classes within the Processing code that occurs

when a user moves an audio object in VBAP mode.
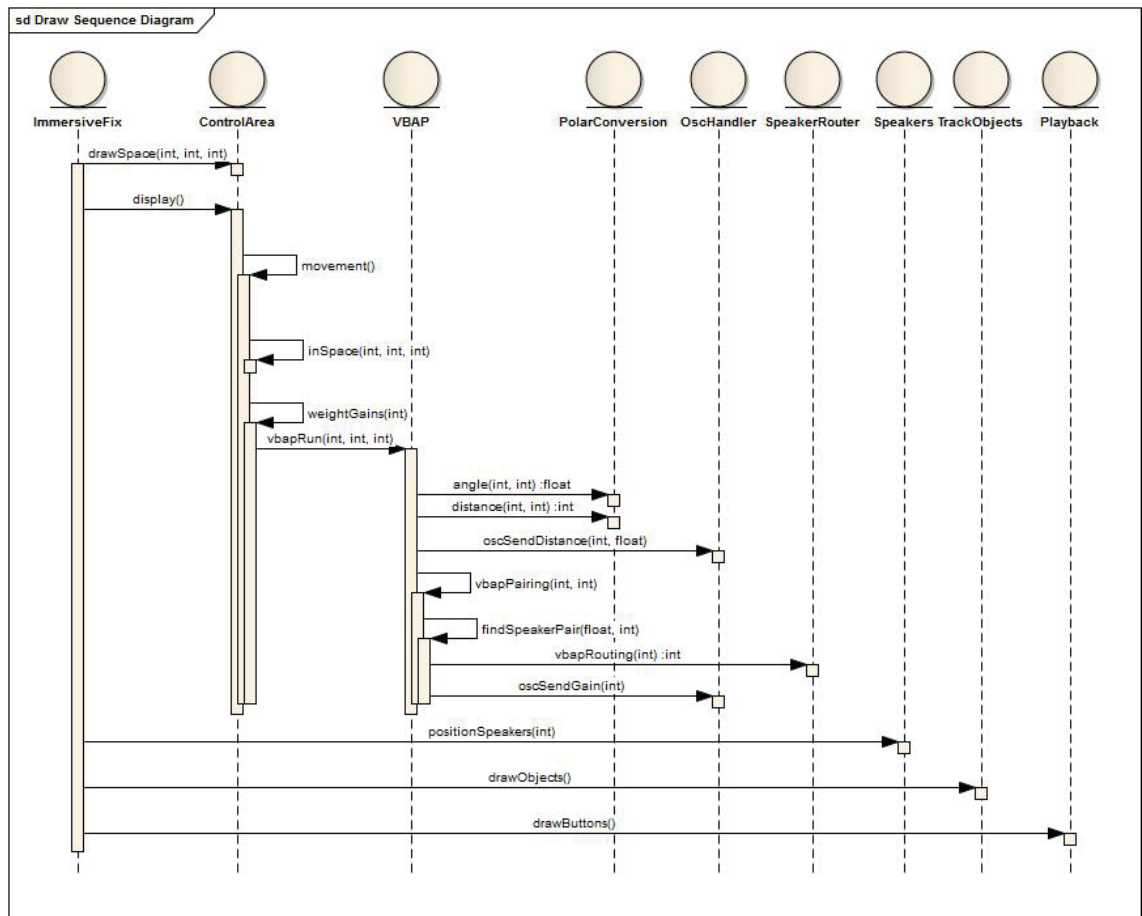


**Figure 12: Sequence Diagram For VBAP Mode Interaction**

## 4.3 Source Spreading



**Figure 13: User Spreading An Audio Object**

In addition to direct source positioning, *tactile.space* also offers an ability to control the spread of a sound source. This feature was developed in collaboration with Blake Johnston. By placing a second finger inside a track object and separating the two fingers, the square-shaped audio object will turn into an arc-shaped 'spreader object'. This will draw as an equally distant arc shape circling the listening position as shown in figure 13. A spreader object is drawn between the two fingers with three small white circles along the arc; one at each of the far points of the arc and a third that always represents the midpoint of the spreader. Once the spreader has first been set, a user can then update either end of the spreader, changing the size of the spread by placing a finger inside one of the circles at the far ends. This updates the gain factors for all speakers and redraws the arc, also updating the midpoint of the spreader. By placing a finger inside the circle in the middle of the spreader, a user can move the spreader as a whole object. This feature will allow both radial movement (changing the distance from the listening position) and spherical

movement (changing the spatial position of the spreader). If a user then wishes to make a certain spreader object back into a standard track object, they may do so by simply giving the object a small enough spread that it will be of similar width to a standard track object and it will automatically revert to a square audio object and update all settings to return to the original object behaviour. If the object is spread so the total size of the arc is larger than 345 degrees *tactile.space* assumes the user is aiming for full immersion of the sound filed and thus updates all gain factors accordingly.

The source spreading feature provides immense creative possibilities to the performer: it represents both a characteristic that is assumed in traditional diffusion (the ability to have a sound emanating from more than one direct position), but does so in a way that allows intuitive dynamic movement of this spread out source that would not be achievable on a traditional system. The physical action that would be required from a performer to achieve this sonic motion with a mixing console in adjusting each fader's gain dynamically is extremely difficult. Whilst conceived specifically for the VBAP mode of *tactile.space*, it is also something that could significantly add to the aesthetic possibilities of diffusion in Stereo Pairing mode if conceived as a linear spread across the horizontal speaker pairs rather than as an arc shape.

## 4.4 The GUI

All of the features that will be described in this section about the graphical user interface, such as the spatialisation modes, can be selected by the user before running the Processing sketch. In order to activate each feature, the user only needs to change that features Boolean to true, or update the appropriate integer for the feature; all corresponding calculations will be made accordingly.
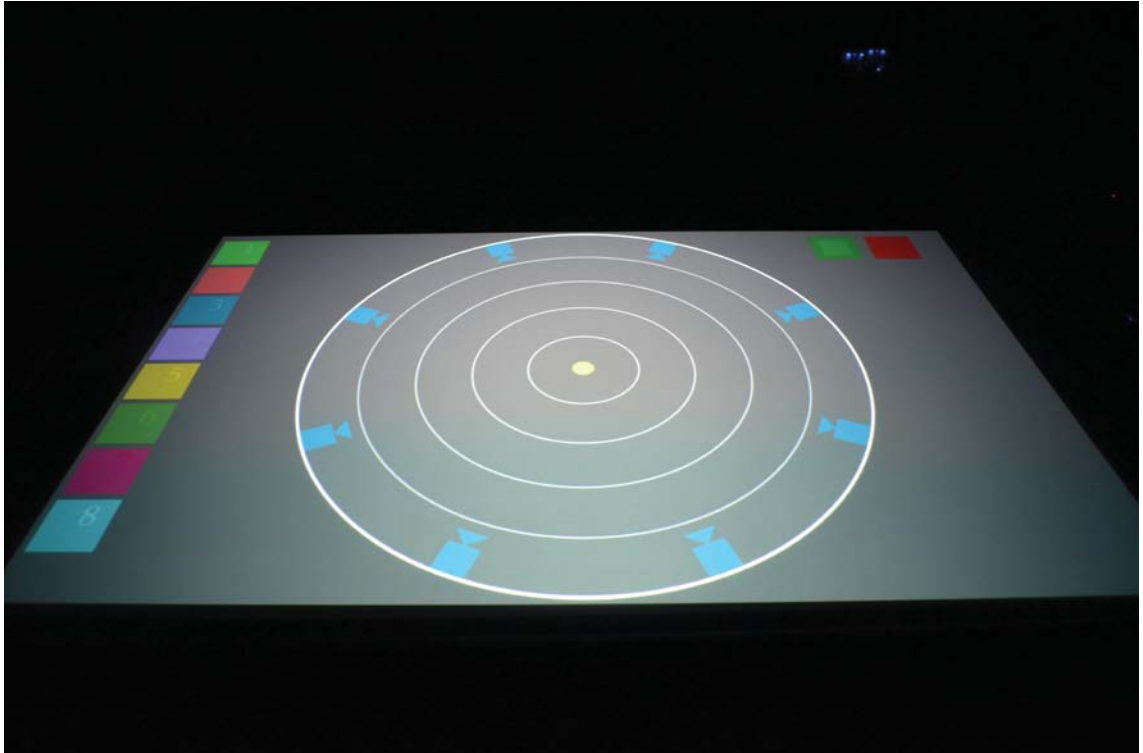
**Figure 14: *tactile.space* startup GUI**

The graphical user interface had to reconcile many considerations in its design. A balance between visual and sonic aesthetics was considered as well as a balance between aesthetics, user learnability, and ease of control. As the program was drawing to completion and the application began to be used by other performers, many aspects became apparent that were not originally considered. These mostly pertained to the ease of learning this new instrument. Given that performers (other than the designer) were going to have limited rehearsal time and space before concerts, some changes needed to be made to reduce the time required to learn the basics of the instrument, allowing rehearsal focus to be on the artistic and performative aspects that the *tactile.space* instrument affords.

The tracking objects caused the most concern to new performers. The first versions of the track objects were coloured as subtle variations of opaque greens. Whilst the variation between objects was distinguishable with the graphics precision of a computer display, these colours became slightly blurred and hard to tell apart once projected onto the table. This was an example wherein the visual aesthetics had to yield to the usability of the interface. For the most part, however, compromises could be reached, maintaining a strong visual aesthetic whilst allowing desirable user feedback.

Another object-related adjustment was the addition of visual representation when a touch event occurred inside an object, activating the object and therefore rendering it moveable. As the tracking on the table can be error-prone, the physical and practical characteristics and nuances of the table began to impact GUI features. The tracking objects were modified so they become highlighted as they are successfully 'grabbed,' allowing their free motion. In a similar way; when a finger is successfully placed within the circles of the spreading objects, the corresponding circle is highlighted so the performer can be confident the desired sonic output of their gesture will be audible. Another point that came from early user tests was that it would possibly be beneficial to have more than the difference in colour to distinguish between track objects. Not everyone remembers colours effectively; it is also very common for people to have trouble distinguishing between certain colours. Therefore, the track object's number was added to its representation in order to give more cues to the performer and make it a more transparent connection to differentiate between their original numerical stem organisation and the visualisation of these stems through the audio objects. As this feature creates a very different and possibly undesirable visual aesthetic users may choose to activate or deactivate this feature when running *tactile.space*.

A further addition to the GUI made for user feedback benefit was that of the distance representation. The maximum distance that can be set when executing the program generally represents the distance between the listening point and the speakers, though for aesthetic reasons this may not necessarily be the case. The maximum can range between 0 and 20 and is set in metres. In rehearsals, users found it hard to imagine the approximate relationships between the representation of the speaker space and the physical space. To counteract this problem when run in VBAP mode (the mode relating to the distance encoding), *tactile.space* now divides its control area space by drawing a circle to represent each metre of perceptual space. If the maximum distance is set to 2 meters, the GUI will show a division at the 1-metre mark halfway between the speakers and the listening point. If the maximum distance is set to 10 metres, the control area will be divided into 10 equidistant zones, each representing a metre of perceptual space.

It also proved beneficial to performers to incorporate playback controls within the GUI. These allow the user to start their piece in their own time rather than on a particular time constraint. Additionally, it gives the user the option of creating spatial positions for audio objects prior to the beginning of their piece, rather than having to introduce each object to the space once the piece has begun. This gives many more variational possibilities and control to a performer to create a dramatic start to their piece.

With the addition of a play button, it makes sense for the performer to also have a stop button. Whilst ideally this may not be necessary in performance, it helps increase the rehearsability of the instrument, which is an important aspect that will be discussed in section 6.2. The play and stop buttons send a 1 and a 0 respectively to Max/MSP to start and stop the audio files. With a separate button for each state, the issue of debouncing and false positive TUIO data interrupting the on/off data of a button is avoided.

## 4.5 The Audio Unit

The audio application for *tactile.space* has been built in Max/MSP and is described in the following subsections.

### 4.5.1 Max/MSP Programming

The Max patch is designed to be as simple as possible for computational efficiency. The patch uses the poly~ object to create instances of an audio player. Each instance handles a separate audio stem. Whilst the Processing sketch is set up to work with any plausible configuration, the amount of speakers and audio stems limitations had to be set inside the Max environment. In its current state, the Max patch limits the number of speakers to sixteen and the number of audio stems to eight. These values were chosen as the upper limitations because sixteen is the highest amount of speakers available for concerts at the NZSM and it was decided that most performers would find eight separate stems to be sufficient for creative freedom. Whilst these numbers are technically permanently set, the program was built in such a way that it would be a very easy process to add extra speakers

and audio tracks. However, it is important to note that these parameters in the Max environment have to be set and running all the time whereas in the Processing sketch, only what is needed for each performance is built. Therefore, one must consider the computational expense when considering the addition of extra speakers and audio stems.
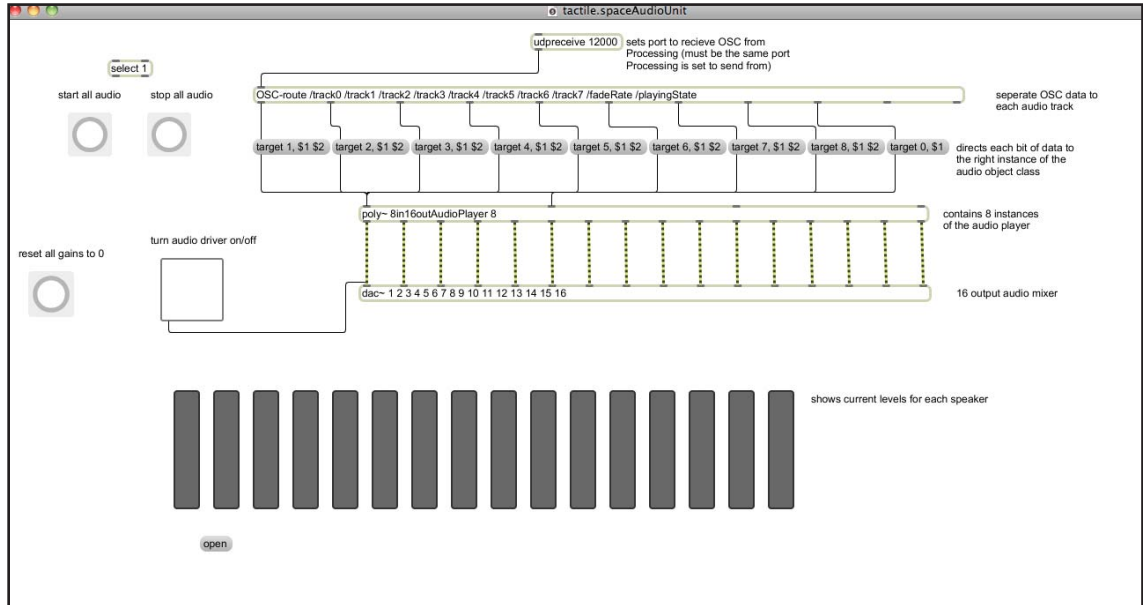


**Figure 15: Max/MSP GUI**

The Max patch receives a live stream of OSC messages sent from Processing. The messages are packed into the audio tracks for which they are destined and contain the gain factors for each of the speakers as well as the current distance for the distance encoding. The first thing the patch must do is to unpack these messages and send the data to the right place. The poly~ object is input data with a target message beforehand. This message can be used to direct information to the correct instance of the patch. Once inside the audio player sub-patch, the data is then unpacked further and sent to the correlating algorithm where it is multiplied by the input of the audio signal and sent to the digital audio converter for the appropriate speaker. When the Processing sketch is first run it sends a series of constant values for the particular performer such as the maximum distance and the fade rate.

### 4.5.2 Distance Encoding
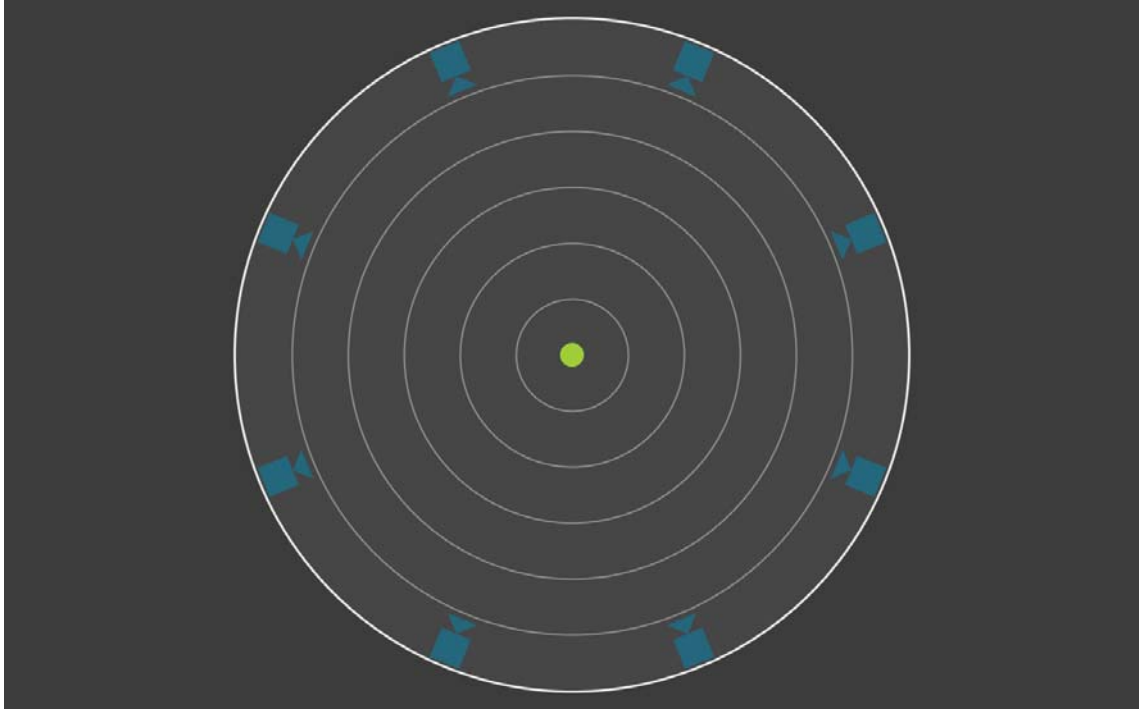


**Figure 16: Perceptual Distance Representation**

The distance encoding was built into VBAP mode in an attempt to reconcile the distance perception limitations within VBAP spatialisation. As the distance encoding involves direct manipulation of the waveform, this had to occur within the Max/MSP patch. The distance-simulating sub-patch was built in collaboration with NZSM student Jason Post. It is based on an understanding of the *1/r* law of atmospheric absorption. As sound waves travel through the air, they encounter friction caused by contact with air particles. The amplitude of a sound decreases with distance travelled (Blauert 1997). The high band frequencies are the first to be lost with distance travelled; the patch therefore applies discrete spectral filtering to the audio with each change in distance. The final effect of the distance encoding is the addition of a subtle reverberation effect that mimics the added reverberation that occurs naturally when a sound is heard from a greater distance. These processes are applied to the audio file before it is sent to the speakers and updates as an audio object is moved between the representation of the speakers and the listening position.

The maximum distance is set before executing the Processing sketch, which then updates this information to the Max/MSP patch. The maximum distance can be any integer between 1 and 20 and is given as a distance in metres. Whilst at surface level this could be set at the approximate distance from the listening position to the speakers, it is important to remember that human sound localisation perceives relative distances much more than it does exact distances. Humans hear when a sound is closer to or further from another sound, and by approximately how much further it is in relation to themselves and each other sound source. However, humans struggle to put a metric distance on any given sound source beyond approximately one metre (Blauert 1997). Therefore it is the relative change in distance that is perceived by the audience rather than the exact metric distance of a sound source. It must be remembered that the monaural auditory cues used to perceive said distance are very dependent on the environment. The laws discussed were conceived for sounds within an open sound field. Therefore, the reasons for setting this variable are much more aesthetic: the higher this variable is set, the greater the difference between a source at the speaker array's edge versus one closer to the listening position. In assuming a higher maximum distance, a performer allows a great deal more perceptual space in which their composition can sit and move. As discussed, performers have an indication of this perceived spatial distance with visual lines dividing the control area into spaces of a metre each, allowing them to dynamically manipulate this distance perception.

## 4.6 Speaker Routing

Along with the problem of standardisation in configurations across multi-channel speaker arrays (discussed in section 4.1), there also exists a discord between conventions for routing speaker systems. Even amongst the commercial standards of 5.1 and 7.1, speaker routing can vary from system to system. A mistake in speaker routing within the *tactile.space* system will cause a complete breakdown of all phantom source positions. One of the standard speaker routing systems follows what would seem most natural for stereo pairing systems is shown in figure 17.1 As this is the most common occurrence for speaker setups at NZSM, it was decided that all modes of *tactile.space* should conform to this standard.
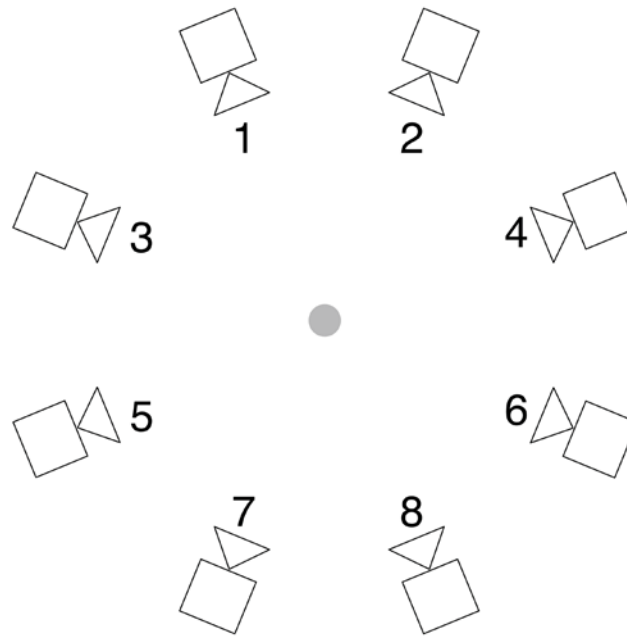
**Figure 17.1:** *tactile.space* **Speaker Routing Convention**

This caused concerns in the development of VBAP mode. A more intuitive way to route speakers in VBAP is a clockwise circular routing system as shown in figure 17.2. This is probably the second most common routing convention used for pantophonic arrays. If modes within *tactile.space* conformed to different speaker conventions, there would be a need for re-routing speakers between performances in a concert. Doing so would contradict the ease of modularity that has been a strong theme throughout the project.

It was therefore deemed necessary to add a speaker routing class inside the Processing sketch to ensure all modes conformed to the same routing convention. When VBAP mode is activated, whenever a gain factor is calculated it is first sent through the speaker routing class in order to send it to the correct speaker in Max/MSP. As mentioned, the coordinate system in Processing has the top left-hand corner of the screen as its coordinate origin (see figure 11.1). When the system is re-routed to the middle as the first step in the VBAP calculations, it then assumes a clockwise rotation starting to the right of this origin point. The first division of speakers into VBAP pairs therefore happens as shown in figure 17.3. The role of the speaker routing class is to take the speakers, as numbered in figure 17.3, and return their corresponding values as depicted in figure 17.1.
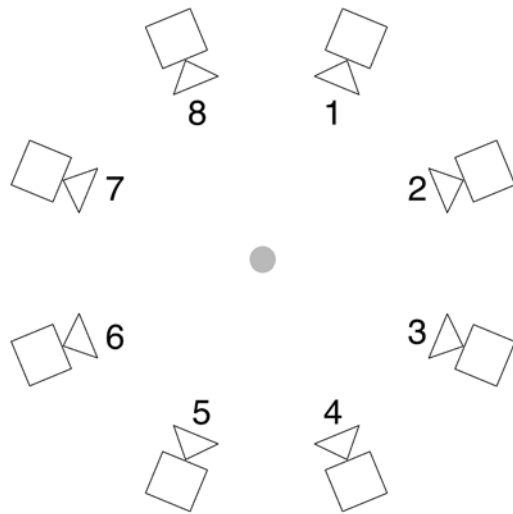
**Figure 17.2: Circular Routing
Convention**
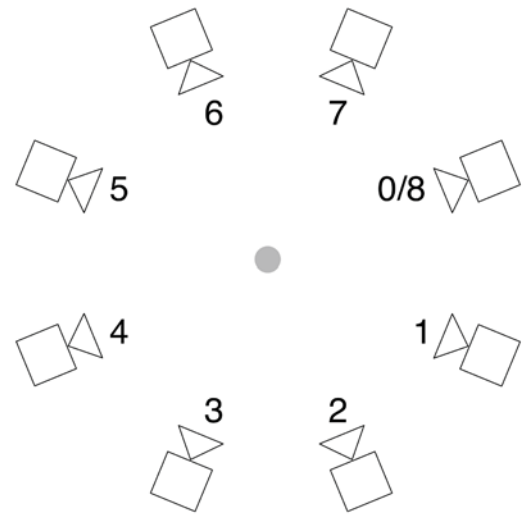
**Figure 17.3: VBAP
Mode Numbering**

The gain factors for the speakers are then sent to Max in the correct order no matter which mode is activated. The speaker router must take into account the number of speakers in use, as this will affect the changes that need to be made when rearranging the speaker numbering. When any integer is sent to the speaker routing class, the class first checks the value that is the current number of speakers in use. The number is then directed to a specific function according to the number of speakers and returns the value of the routed speaker.

## 4.7 Live Input

For a collaborative project with live electronic musician Blake Johnston, the *tactile.space* audio unit was given the addition of up to eight channels of live input. The project featured real-time spatialisation of Johnston's hyper-instrument, Ezither (Johnston 2012) (see portfolio DVD). This feature expands a traditional diffusion of fixed-media composition into the real-time diffusion of live electronics. Johnston's Ezither sends eight channels of processed audio into *tactile.space* to be spatialised in real-time. The project was featured at the opening of Massey University Wellington's College of the Creative Arts new building in June 2012. With spatial performances of this nature, a further level of interaction is achieved between two performers. With both musicians improvising, the instrumental

performer can react to the spatial attributes inherent in the piece. Additionally the diffusion artist can respond to the changing musical nature of the composition. The addition of live input increases the range of pieces that can be performed on *tactile.space*, and the spatial aspects of live electronic pieces. A performer may choose the number of channels ofr live audio input and the number of audio file playback channels.

## 4.8 Summary

Many of the features discussed in this chapter were designed not only to strengthen the relationship between performer and audience member, but also to increase the possibilities for creative expressivity in sound diffusion with an intuitive approach to the mapping of human gestural motion. In doing so, *tactile.space* has redefined the act of diffusion and expanded the possibilities of sonic trajectories affording a level of creative freedom that has previously not been available in diffusion performance. The modularity of *tactile.space* means that the interface can be easily configured for a performer's needs, making it a versatile and truly expressive diffusion tool, that is not only capable of traditional diffusion techniques but also allows for live spatialisation through methods previously restricted to rendering in the studio environment.

# Chapter 5 - Evaluation

This chapter presents an evaluation of the *tactile.space* diffusion interface. The first section summarises data collected from a user study questionnaire that evaluated the interface in many aspects and compares it to a traditional diffusion user interface. The second section outlines a new kind of interaction for the diffusion paradigm based on interactivity between performer and audience member, and how this was encouraged by a new performance setup.

The Victoria University Ethics Committee granted ethics approval for the study to take place. In line with their standards, all statements by participants are quoted anonymously.

## 5.1 User Study

A group of 18 students participated in the user study evaluating their experience using the performance interface and comparing it to a traditional diffusion user interface (i.e. a mixing console). All of the participants were studying, or had previously studied, electroacoustic composition and had varying levels of experience with traditional diffusion techniques.

The participants used common modes and speaker configurations to eliminate variation in results that may have been caused by differences within the interface itself. Performers collectively decided to engage with the VBAP mode as, after exploration of each mode, they felt VBAP gave them the most freedom and creative expression of space. The speaker configuration was an 8-channel European standard system, as shown in figure 6.1. As the users had each composed their own piece to diffuse, they were able to customise all other features of the interface individually for performance.

On May 30th 2012, a concert took place in the Adam Concert Room at the New Zealand School of Music. The concert featured students performing their compositions diffusing them with the *tactile.space* system. The concert made full use of the available modularity on offer through *tactile.space*. The number of audio stems ranged from three to eight depending on the aesthetic requirements for each performer. The maximum distance variable was also customised for each performer and ranged from one to eight metres.

In the first series of questions participants were asked to rank *tactile.space* on a scale of 0-10 with 0 being the lowest score and 10 being the highest and best possible result. The interface was ranked in terms of its intuitiveness, learnability, control and perception of spatial movement and positioning, and the creative expression it affords. Overall the interface rated quite highly with each category scoring an average of 8.0 or higher.

### 5.1.1 Intuitiveness

The first questions addressed the issue of intuitiveness. All of the participants rated the interface as a score of 8 or higher (see figure 18). The average score for intuitiveness was 8.8.
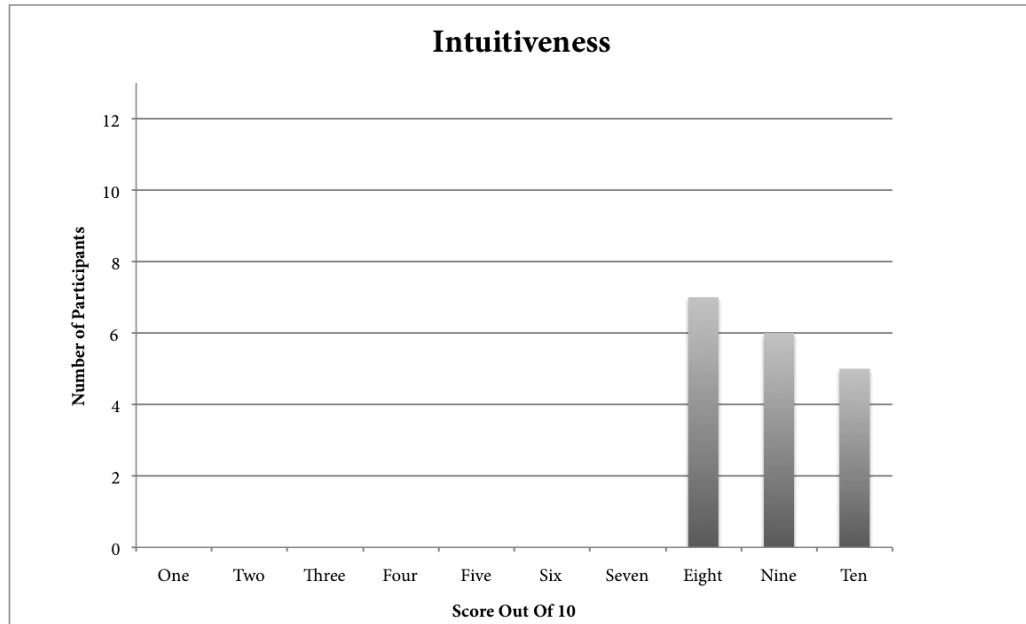
**Intuitiveness**

**Figure 18: Results for Intuitiveness**

Participants were also asked if there were any gestures in features on the *tactile.space* interface that were counterintuitive, all participants answered 'no'. Assessing this aspect of the interface was very important because gestural intuitiveness was one of the early goals for a new diffusion user platform. The gestural nature of interaction on a traditional diffusion interface has been pointed out as one of its major flaws, and as a contributing factor to the limitation of possible sonic trajectories. In comparing the two interfaces one participant stated

> *tactile.space* has much more parallels to actual movement and spatialisation of sounds.

The gestural intuitiveness of *tactile.space* allows for a wider range of sonic trajectories and greater transparency in gesture/output relationships for the performer and audience member.

### 5.1.2 Learnability

Closely linked to the intuitiveness of an interface is the ease with which a new user can learn how to use the interface. As discussed in Chapter Two, performers of sound diffusion rarely have adequate time to rehearse with a new diffusion system or in a new performance space. This increases the necessity for new diffusion performance interfaces to be very easy to learn. Mooney (Mooney 2005) suggests that the traditional interface of a fader bank is ideal for sound diffusion as general electronic musicians are already so familiar with the actions required that the time to learn the interface is minimal. In the same thesis, he is highly critical of the fader interface as it lacks performative aspects and limits possibilities for spatial movement. It is possible that, if a new interface was fast and easy to learn, artists may consider a change from the missing console with confidence to perform in a new style. *tactile.space* scored an average of 9 out of 10 for learnability, suggesting that it does have the potential for new performers to learn the system in a very short amount of time, therefore giving them time to focus on the musicality of their performance rather than struggling to get their ideas across.
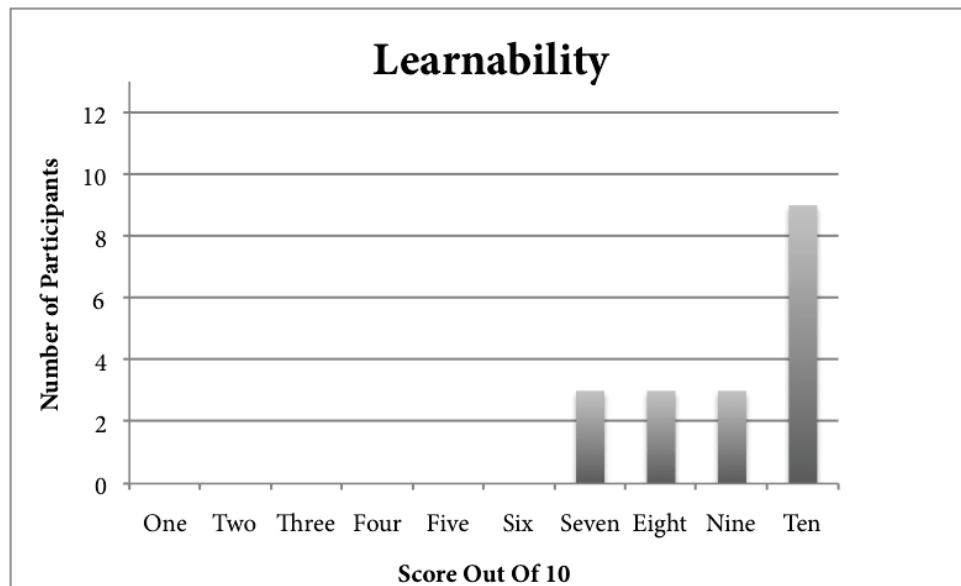


**Figure 19: Results for Learnability**

Participants in the user study had varying amounts of time experimenting and familiarising themselves with the interface, ranging from around 10 minutes to many hours. This may have affected their score for learnability. It is interesting to note that those who scored

the interface as 7 or 8 commented that is was getting used to the nuances of a large-scale touch surface itself that required a little extra time, learning the sensitivity of the table, how much pressure to apply and other factors. The majority of these aspects could be addressed by assessing the lighting, camera speed and calibration of the touch table itself. These issues regarding the learnability might therefore be rectified with advances in the hardware rather than the software.

### 5.1.3 Spatial Control

In order for any new musical instrument be capable of sophisticated and expressive performance, the artist must feel they have precise control over its musical elements. The VBAP mode of *tactile.space* (which was used in the user study), aims to present accurate and dynamic positioning of phantom sources. In order to assess this aspect of the interface, users were asked to evaluate the extent to which they felt they had control of the spatial positions and movements of their audio tracks. Position control rated slightly higher with an average score of 8.8, while movement control was rated 8.3.
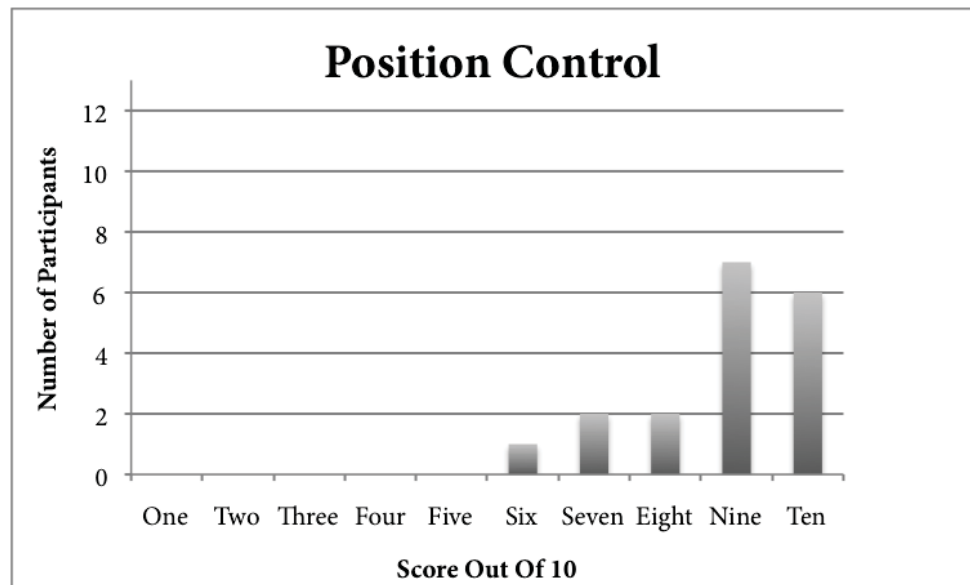


**Figure 20.1: Results for Spatial Position Control**

As *tactile.space* employs direct source positioning in a two-dimensional space that is represented visually on the user interface, the control of spatial positions of sound source is an easy relationship to make. In a traditional diffusion system the spatial positions are an

abstraction of the controlled amplitude of numerous speakers; with *tactile.space* this relationship is flipped. Instead, the amplitudes of each speaker are an abstraction of the direct positioning of sound sources.
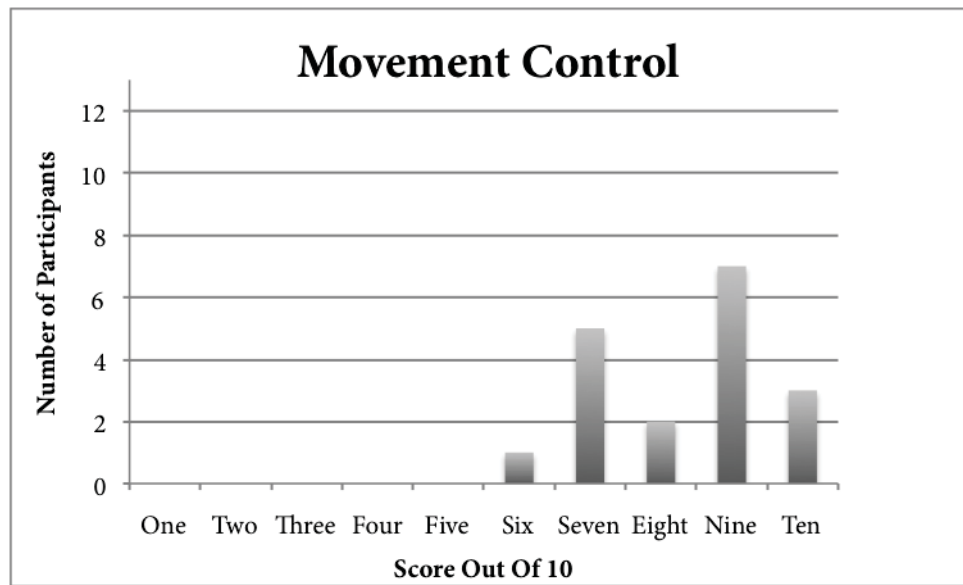
**Movement Control**



**Figure 20.2: Results for Spatial Movement Control**

One comment stated

easy to move sound around the table and reposition in a new area.

The results suggest that users felt they had adequate control of spatial manipulations on the interface. The spatial movement that *tactile.space* affords varies greatly to that of a traditional diffusion system. Figure 20.2 shows that users felt controlling their sonic trajectories directly from the user interface rather than as an abstraction gave them a high level of control of the sound field's depiction.

**5.1.4 Spatial Perception**

Participants were also asked to evaluate the extent to which they felt the spatial positioning and movement were perceivable. This is relevant for both performer and audience member. As an interface that encourages real-time spatial positioning and source spreading, it is important for the expressivity and control of the instrument that these gestural movements are accurately perceived by the performer and the audience member. Average results for perceiveability of both position and movement were 8.0.

**Perceivable Position**

Number of Participants

Score Out Of 10
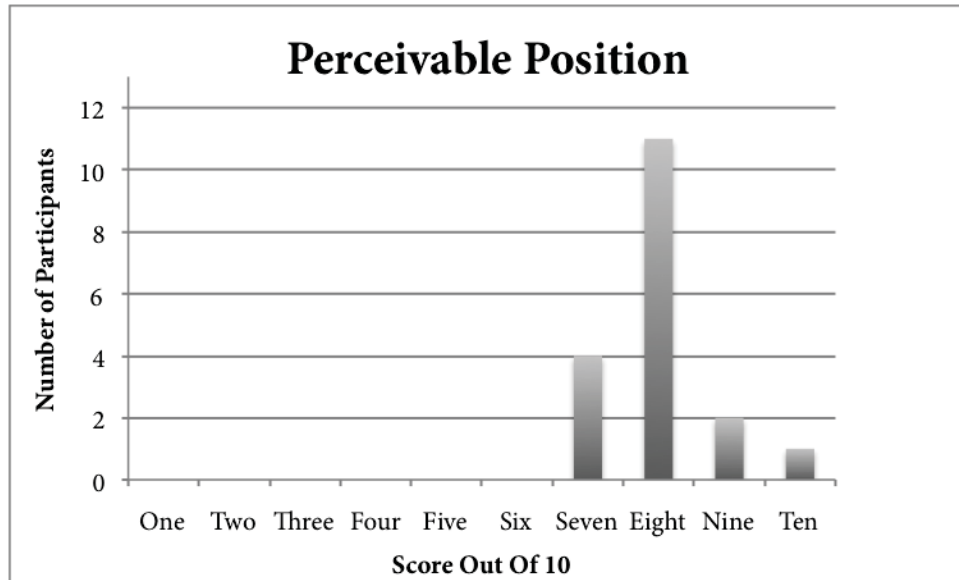
**Figure 21.1: Results for the Perception of Spatial Position**

Some users commented that they felt they were less able to perceive the position and movement of bass frequencies in their pieces. This problem could be attributed to the psychoacoustic phenomena by which lower frequencies resonate omnidirectionally, therefore making it incredibly difficult for humans to localise their positions (Blauert 1997).

**Perceivable Movement**

Number of Participants

Score Out Of 10
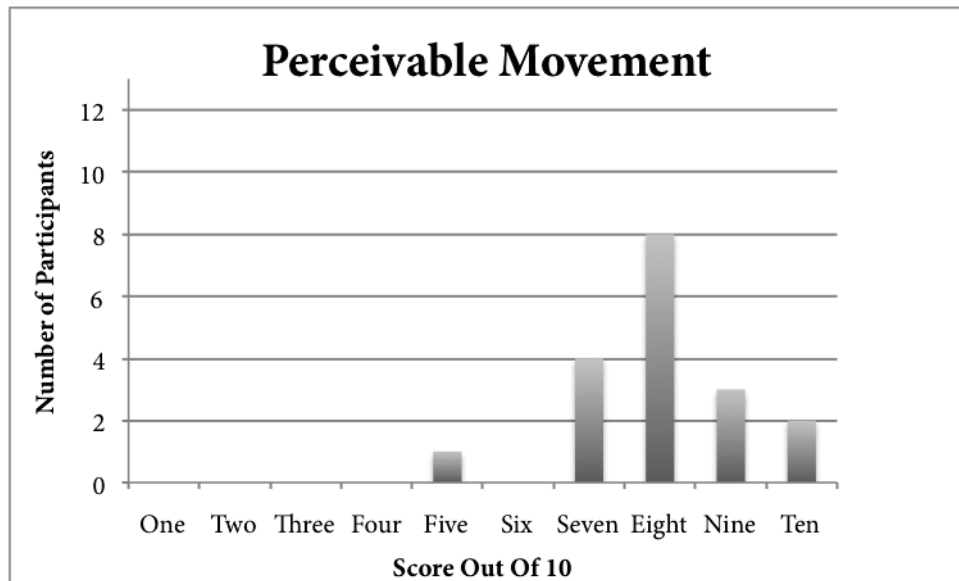
**Figure 21.2: Results for Perceivable Movement**

The aspect of perception of both position and movement of sound objects is closely related to the issue in any surround sound format of the sweet spot. As discussed in Chapter Two, the depiction of the sound field and the accuracy of this depiction based on the listener's position is a variant of many factors. The spatialisation technique used and the

number and positions of speakers are all relevant as well as the position of the listener and the direction they are facing. Participants in the study were both performers and audience members. Therefore, they were exposed to varying listening perspectives at different times. The survey did not specifically ask for the difference between the two perspectives. This might account for the large variation in results for movement perception however, which, whilst averaging 8.0, had 6% of people score it only 5. In contrast to this, 28% of people scored movement perception as 9 or above; suggesting spatial field perception is highly subjective.

### 5.1.5 Creative Output

As a new tool for sound diffusion and as an interface for artistic musical performance, *tactile.space* must give a performer creative freedom and expressivity in order to allow the creation of an engaging performance. One factor with a traditional diffusion system as discussed in chapter two, is the problematic coupling of performative gesture to sonic output, affecting the significance of performative gesture of the diffusion artist. An important goal for the *tactile.space* interface was to address these performance issues in diffusion, encouraging a more active and engaging gestural performance.

### *5.1.5.1 Expressivity*

The results for questioning the extent to which performers felt the interface gave them creative expression varied widely. Over 50% of performers scored the interface with nine or above, however, expressivity also received a 4 which was the lowest score given across the entire survey. Comments from those who gave lower scores for expressivity suggested that whilst the interface gave them spatial freedom the expression in the music came from the composition of the piece itself rather than the diffusion. It is also interesting to note that those who ranked the expressivity as 4 or 5, then went on to rank the questions regarding feelings of performance as 8 or higher. This suggests that whilst they felt more expression was achieved through composition, the act of diffusion on *tactile.space* was also a significant creative outlet.
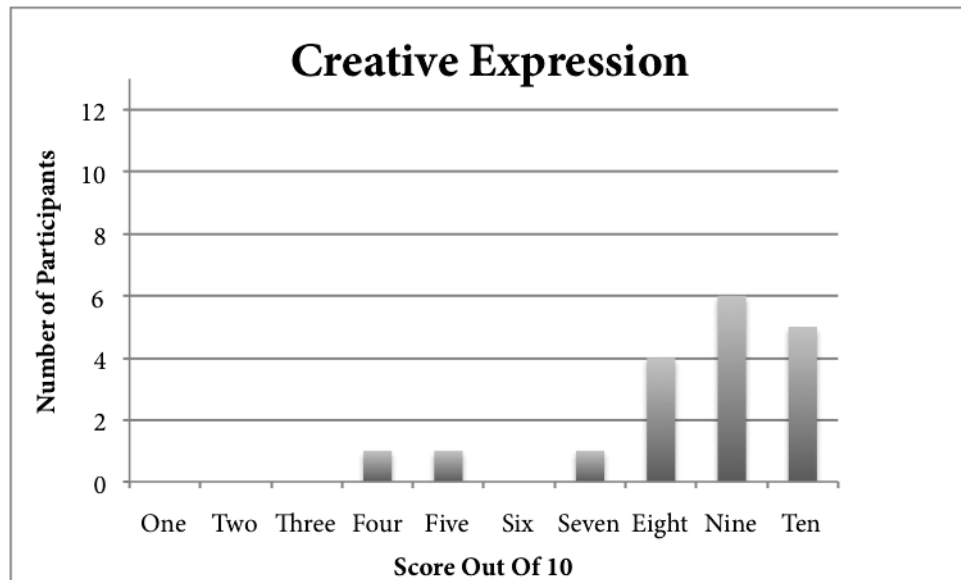
**Figure 22.1: Results for Creative Expression**

*5.1.5.2 Performativity*

Participants were asked to evaluate the extent to which they felt their live interaction with the interface gave them the feeling they were actively performing their music. The average score for performativity was 8.6.
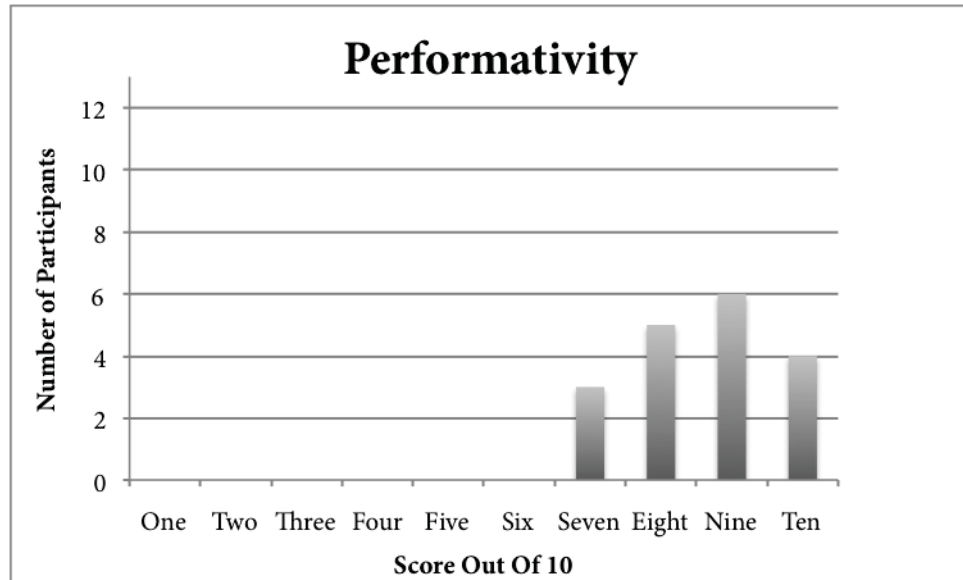
**Figure 22.2: Results for Performativity**

The transparent relationship between gesture and output exhibited by *tactile.space,* allows the diffusion artist an interaction with the space and the audience that previously has been abstracted by the mixing desk console. One performer commented,

> [*tactile.space*] allowed for actual bodily integration as opposed to a mixing desk which without the visual aid of the touch table is not as intuitive and performative.

### 5.1.5.3 Real-time Addition To The Piece

The final question relating to performativity asked the extent to which performers felt they were making a significant contribution to their piece in the live setting. The average score was 8.6, suggesting that the diffusion on *tactile.space* was integral to the success of the performance.

**Contributing To The Piece Live**
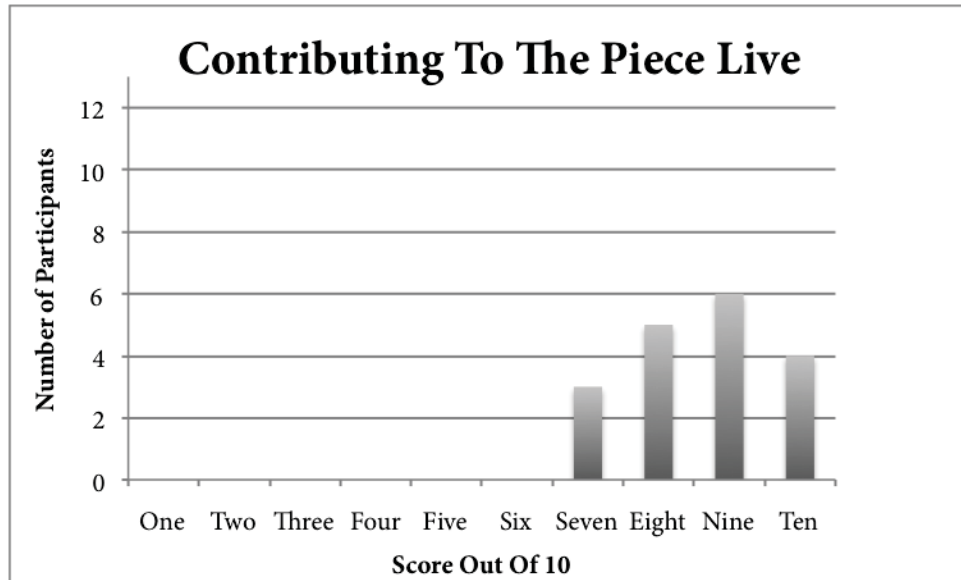


**Figure 22.3: Results for Active Contribution to the Piece**

Sound diffusion has always been said to be an active process (Mooney 2005). To achieve this the performer must feel that their live contribution significantly affects the reception of the piece, or else there is no need for a performer's presence at all.

**5.1.6 User Interface Comparison**

Another series of questions asked the participants to give a direct comparison between *tactile.space* and the traditional diffusion user interface of a fader bank. All of the participants stated they preferred diffusion by *tactile.space*, and expressed a desire to use the interface again for performance of electroacoustic works.

Performers compared both the intuitiveness and performativity of the two interfaces: *tactile.space* was rated higher than the mixing desk by all performers in both areas. Comments comparing the gestural intuitiveness included,

> a mixing desk is being a mixer, *tactile.space* is being a performer, a musician.

> [*tactile.space*] is better, you can see individual elements in [the] work, and gesturally move them.

Comments comparing the performance practice included,

> The performance felt more 'live' in terms of being distinct from typical associations of a mixing desk. In terms of practice I found the positioning aspects of surround sound most useful and connected.

> *tactile.space* is far better as it allows for live spatialisation of individual stems, thus composing elements of the piece live. This would also make it more intuitive in terms of performing other peoples' work.

Users were also asked to list their favourite features of *tactile.space*. The most common answers were the source spreading and the distance encoding. These are both features that represent a shift in the diffusion performance paradigm. The spatial effect that occurs from both the distance encoding and the source spreading are unachievable from performance gestures on a traditional diffusion system. This may suggest that after sixty years of a performance tradition, musicians are ready to embrace the expanding possibilities of live sound diffusion that systems such as *tactile.space* afford. This is not to suggest that *tactile.space* is the only user interface affording such complexity of interaction, rather that composers and performers are ready for the occurrence of said shift in performance practice, and are therefore willing to embrace these aspects of any new technologies or interfaces being integrated into diffusion performance.

## 5.2 Concert Setting

As section 5.1.6 has shown, the use of *tactile.space* changed the traditional interaction between performer and audience member for a diffusion concert. In order to embrace this new level of interactivity and performance gesture in diffusion concerts, a new concert setup model was formed. In a traditional diffusion concert the performer sits in the centre of the space, the sweet spot, with optimal perspective of the auditory field. The audience is then seated as closely behind the performer as the space allows, all facing the same direction as the performer, to the front of the space (see figure 23.1).
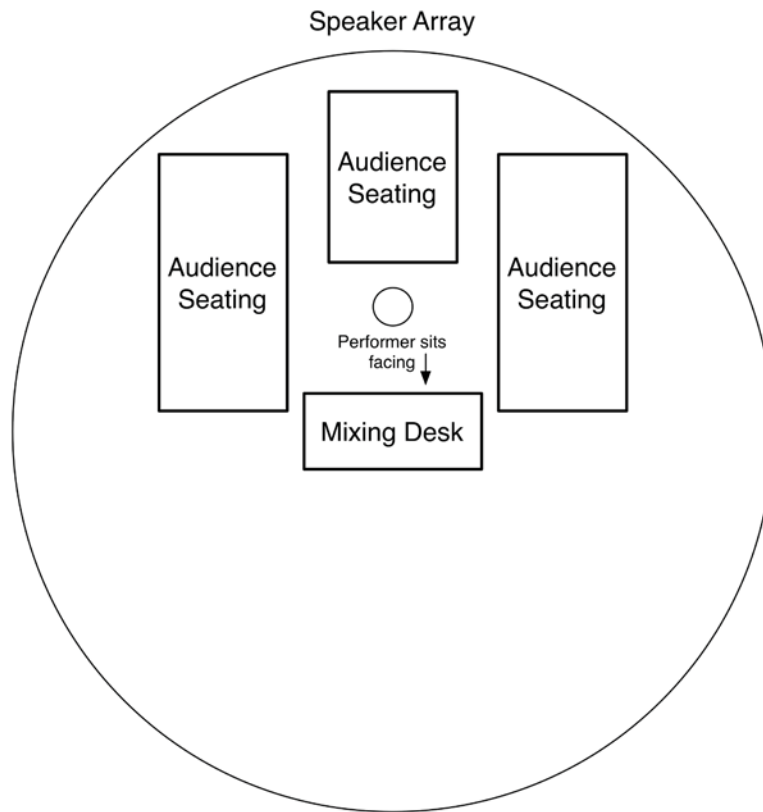
**Figure 23.1: Diffusion Concert Setting Model**

This physical spatial situation outlines one of the major barriers in diffusion performance with traditional systems: that in order to project an optimal spatial field the performer must face the same direction as the audience, removing any opportunities for facial communication with an audience, as well as hindering the ability of the audience member to watch the performer's actions. Some purists may argue that this situation is actually ideal for electronic art music, as it somewhat removes the stage presence of the performer and allows the music to take place in an acousmatic setting (Barrett 2007). Diffusion practice dictates a more Cagean approach (Cage 1961): that a performance of music should be just that, a performance. By merely placing a performer into the concert scene, an expectation of a level of performative action is asserted, requiring the attention of the audience.

To address this issue, for performances on *tactile.space*, the diffusion artist stands in the middle of the space and the audience is situated in a semi-circle facing the performer, as shown in figure 23.2.
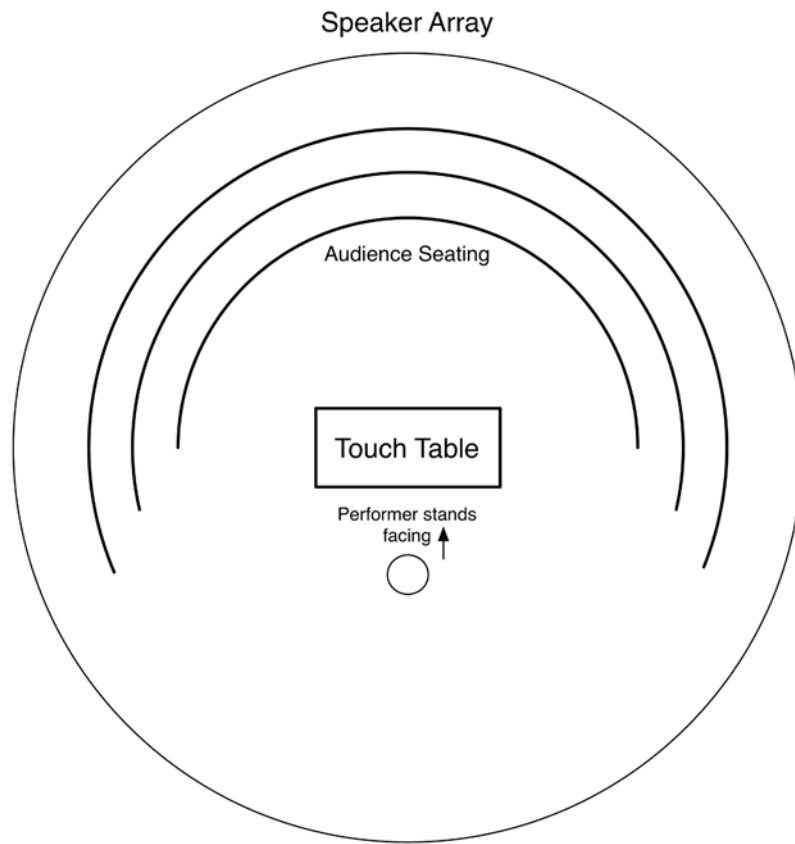
**Figure 23.2:** *tactile.space* **Concert Setting Model**

The audience has view of both the performer and the gestural interaction with the interface. Although what is then a frontal spatial field for the performer differs to that of the listener, the direction any listener faces is irrelevant as the spatialisation occurs through interaction with a direct representation of the physical space. By giving the audience a view of the interface as well as the gestural action of the performer, barriers in transparency between gesture and output are also broken by a visual link accompanying the sonic movement.

Early evaluations of the new diffusion tool, *tactile.space*, show not only the potential the instrument exhibits for live performance of sound spatialisation but also a desire by diffusion artists to gain access to and explore new gestural possibilities in their art form. An increase in intuitive gestural input from the diffusion artist has afforded a greater complexity and transparency in multiple interactive contexts within a sound diffusion concert.

# Chapter 6 - Conclusions

To conclude the exegesis, this chapter will outline directions for future research both specifically for the *tactile.space* project and for sound diffusion in general. There will be a discussion of how performance research such as this project, affects the diffusion paradigm. Finally, some suggestions for avenues outside the specific diffusion paradigm that may benefit from research of this nature will be made.

## 6.1 Summary

This thesis has presented a critique on the performative interaction that occurs in the context of a live diffusion concert. The suggestion has been made that by incorporating research from a wider field of electronic music performance, specifically that of interface design and live electronics, diffusion performance could become a more engaging and spatially complex art form.

The first section introduced the wider paradigm of sound spatialisation, describing its developments through history and how they have affected the field. In focusing on sound diffusion, there was a discussion of both diffusion as an aesthetic practice and current systems built for sound diffusion, looking at their individual positive and negative aspects.

The discussion of sound diffusion systems showed a gap in the field for research into developing user interfaces specifically for the performance art of diffusion.

The second section presented the idea that the sub-community of interface designers, developing musical interfaces for multi-touch tables, had exhibited potential for development of new musical instruments. The gestural and configurability aspects of such surfaces lend themselves ideally for future development of diffusion user interfaces.

In introducing one solution to the problems outlined, the author's diffusion interface research project *tactile.space* was described and evaluated as a successful tool for performance of sound spatialisation.

## 6.2 Future Research

The user study evaluations and success of performances proved the merit of a diffusion interface like *tactile.space*; however, with further research and developments the interface could encourage more increases in performative gestures in the diffusion domain.

One area that the interface could explore is that of gestural complexity. Currently, the program recognises three main levels of interaction: moving an object, spreading an object, and moving an already spread object (see figure 4). It is possible that the application could also introduce ways to group objects so the user may move one object and have selected other objects mimic, mirror or synchronise their movements concurrently. By increasing the application's recognition of performance gestures, the program could then allow a performer a series of gestures that react in certain ways for computer-assisted diffusion. An example of this might be if a performer moves an object in a circular motion for a certain amount of time, the program could learn this motion and have it continue until the performer commands the motion to stop. This would increase the complexity of spatial trajectories afforded by gestural user interaction, and further enhance the spatial segregation of multiple stems, as the performer could potentially have all of their stems

moving independently of each other in real-time. This same concept could be extended to incorporate some physics of the circular control area. For example, the program might recognise a gesture that triggers a ping-pong-like effect so an audio object appears to be bouncing around the speaker array, getting its new trajectory from the shape of each curve it hits.

As a new instrument it is important that measures be made to increase the rehearseability of the instrument to encourage sophisticated and virtuosic performance. These additions could be as simple as more playback controls (for example, buttons for pause, fast forward and rewind) to allow performers to rehearse specific sections of their piece with ease. This could also include graphical representations of the waveform that allow interactions directly with the audio file. The performer might then select and highlight specific sections to loop, or start playback from certain sections of the piece. This could be the waveform of the whole piece or of individual stems. It could potentially act as somewhat of a score in performance, giving performers a graphical indication of what is happening in the piece and allowing them to foresee certain sonic events. It would aid particularly with performance and rehearsal of other people's works. A significant addition that will be explored is the possibility of a real-time binaural mixdown. A binaural mixdown, would allow performers to practise their diffusion with headphones and still get the effect of a full spatial field. The advantage of this would be the possibility to rehearse effectively without always requiring access to a full multi-channel speaker system.

The strength of Ambisonics has so far been in the accurate reproduction of rendered sound fields. If *tactile.space* was to incorporate an Ambisonics decoding system it would be possible for Ambisonics to be reproduced on the system and for sound field operations such as sound field rotations, or tilts, to be manipulated through gestural live performance.

It has also been suggested that if there was a way that the gestural interaction could be recorded, with some additions to the software, *tactile.space* could be used as a studio tool.

Whilst it was always intended as a performance interface and the concept was to extend gestural complexities and give unrestricted sonic trajectories without need for computer assisted diffusion, the intuitive nature of such interaction could benefit those wishing to record sonic trajectories for rendering surround sound. Further to this, there are avenues outside the diffusion domain that might benefit from real-time spatial movement in this fashion. In particular with sound for film where Dolby 7.1 systems are very common, if the engineer was able to track characters and other sound-emanating aspects of the film, the intuitive nature of a system like *tactile.space* would make this type of work faster, simpler and more accurate. The same could be said about the theatre and dance industries, in this case the engineers could benefit from the real-time spatial control, with the ease of linking the graphical representation of speaker arrays to those in the performance spaces and the audio object representations to characters and sound effects on stage.

The success of *tactile.space*, even while still in its early stages of development, indicates that a marriage of the fields of diffusion and interface design could see a heightened performative atmosphere in diffusion concerts. *tactile.space* is, to the best of the author's knowledge, the first research of its kind. The research has discovered not only the benefits of gestural performance for sound spatialisation, but also a strong desire from those working within the art form to explore new possibilities that research of this nature might afford. By building performance interfaces that unlike the mixing desk are designed specifically for the art of diffusion performance, the real-time spatial possibilities can be extended. This is not limited to multi-touch applications: development of diffusion interfaces could take place on any number of sensor-based musical controllers, wearable interfaces or with extensions of more conventionally conceived fader consoles. As in other fields of electronic music, advancing technologies could be judiciously embraced to increase musicality and sophistication of more traditional performance models. With collaborative research and the creative application of technological advances, such as that demonstrated by *tactile.space*, live diffusion, until recently a relatively stagnant performance art, has a dynamic and exciting future.

# Bibliography

Austin, Larry. 2000. "Sound Diffusion in Composition and Performance: An Interview with Denis Smalley." *Computer Music Journal*: MIT Press 24 (2): 10–21.

Barrett, Natasha. 2007. "Trends In Electro-acoustic Music." In *The Cambridge Guide to Electronic Music*, ed. Nick Collins and Julio d' Escrivan. New York, United States of America: Cambridge University Press.

Blauert, Jens. 1997. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Revised Edition, English Translation. Cambridge Massachusetts: MIT Press.

Cage, John. 1961. *Silence; Lectures and Writings*. 1st ed. Wesleyan University Press.

Carrascal, Juan, and Sergi Jordà. 2011. "Multitouch Interface for Audio Mixing." In *Proceedings of New Interfaces For Musical Expression*. Oslo, Norway.

Clozier, Christian. 2001. "The Gmebaphone Concept and Cybernephone Instrument." *Computer Music Journal*: MIT Press 25 (4). Winter: 81–90.

Croft, John. 2007. "Thesis On Liveness." *Organised Sound*, Cambridge University Press 12 (1): 59–66.

Desantos, Sandra, Curtis Roads, and François Bayle. 1997. "Acousmatic Morphology: An Interview with François Bayle." *Computer Music Journal*: MIT Press 21 (3): 11–19.

Emmerson, Simon. 2007. "Diffusion Projection: The Grain of the Loudspeaker." In *Living Electronic Music*. Hampshire, England: Ashgate Publishing Limited.

Gayou, Evelyne. 2007. "The GRM: Landmarks on a Historic Route." *Organised Sound,* Cambridge University Press 12 (3): 203–211.

Gerzon, Michael. 1977. "Design of Ambisonic Decoders for Multi Speaker Surround Sound." *In Audio Engineering Society Convention Paper*. New York, United States of America.

Griesinger, David. 2002. "Stereo and Surround Panning In Practice." *In Audio Engineering Society Convention Paper*. Munich, Germany.

Harrison, Jonty. 1999. "Sound, Space, Sculpture: Some Thoughts on the 'What', 'How' and 'Why' of Sound Diffusion." *Organised Sound*, Cambridge University Press 3 (2): 117–127.

Harrison, Jonty, and Scott Wilson. 2010. "Rethinking the BEAST: Recent Developments in Multichannel Composition at Birmingham ElectroAcoustic Sound Theatre." *Organised Sound*, Cambridge University Press 15 (3): 239–250.

Hochenbaum, Jordan, and Owen Vallis. 2009. "*BrickTable*: A Musical Tangible Multi-Touch Interface." *In Proceedings of Berlin Open Conference 09*. Berlin, Germany.

Jérôme, Daniel, Nicol Rozenn, and Moreau Sébastien. 2003. "Further Investigations of High Order Ambisonics and  Wavefield Synthesis for Holophonic Sound Imaging." *In Audio Engineering Society Convention Paper*. Amsterdam, The Netherlands.

Jérôme, Daniel, and Moreau Sébastien. 2004. "Further Study of Sound Field Coding with Higher Order Ambisonics." *In Audio Engineering Society Convention Paper*. Berlin, Germany.

Johnston, Blake. 2012. "Ezither: Extended Techniques For Customised Digital Bowed String Instrument." *In Proceedings of Australasian Computer Music Conference*. Brisbane, Australia.

Johnston, Blake, Owen Vallis, and Ajay Kapur. 2012. "A Comparative User Study of Two Methods of Control on a Multi-Touch Surface for Musical Expression." In *Proceedings of New Interfaces For Musical Expression*. Ann Arbor, United States of America.

Jordà, Sergi. 2007. "Interactivity and Live Computer Music." *In The Cambridge Guide to Electronic Music,* ed. Nick Collins and Julio d' Escrivan, 89–106. New York, United States of America: Cambridge University Press.

———. 2008. "On Stage:  The *Reactable* and Other Musical Tangibles Go Real." *International Journal of Arts and Technology* 1 (3/4): 268–287.

Jordà, Sergi, and Marcos Alonso. 2006. "Mary Had a Little scoreTable* or the *reacTable** Goes Melodic." *In Proceedings of New Interfaces For Musical Expression*. Paris, France.

Jordà, Sergi, Martin Kaltenbrunner, Günter Geiger, and Ross Bencina. 2005. "The *Reactable.*" *In Proceedings of New Interfaces For Musical Expression*. Vancouver, Canada.

Kaltenbrunner, Martin. 2009. "reacTIVision and TUIO: A Tangible Tabletop Toolkit." *In Proceedings of ITS' 09*. Banff, Canada.

Kaltenbrunner, Martin, and Ross Bencina. 2005. "reacTIVision: A Computer-Vision Framework for Table-Based Tangible Interaction." In Proceedings of the 1st International Conference on Tangible and Embedded Interaction. Baton Rouge, United States of America.

Kaltenbrunner, Martin, T Bovermann, Ross Bencina, and Eric Costanza. 2005. "TUIO - A Protocol for Table-Top Tangible User Interfaces." *In Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation* (GW 2005). Vannes, France.

Matthews, Max. 1999a. "The Auditory Brain." *In Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics,* ed. Perry Cook, 11–20. Cambridge, Massachusetts, United States of America: MIT Press.

———. 1999b. "The Ear and How It Works." *In Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*, ed. Perry Cook, 1–10. Cambridge, Massachusetts, United States of America: MIT Press.

Milner, Greg. 2009. *Perfecting Sound Forever: The Story of Recorded Music*. London, Great Britain: Branta Publications.

Mooney, James. 2005. "Sound Diffusion Systems for the Live Performance of Electroacoustic Music". Ph.D, Sheffield, United Kingdom: University of Shefield.

Moore, A, D Moore, and James Mooney. 2004. "M2 Diffusion - The Live Diffusion of Sound in Space." *In Proceedings of International Computer Music Conference.* Miami, United States of America.

Newton-Dunn, Henry, Hiroaki Nakano, and James Gibson. 2003. "Block Jam: A Tangible Interface for Interactive Music." *In Proceedings of New Interfaces For Musical Expression.* Montreal, Canada.

Patten, James, Ben Recht, and Hiroshi Ishii. 2002. "AudioPad: A Tag Based Interface For Musical Performace." *In Proceedings of New Interfaces For Musical Expression*. Dublin, Ireland.

Pierce, John. 1999. "Hearing In Time and Space." *In Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*, ed. Perry Cook, 89–103. Cambridge, Massachusetts, United States of America: MIT Press.

Pulkki, Ville. 1997. "Virtual Source Positioning Using Vector Base Amplitude Panning." *Journal of Audio Engineering Society* 45 (6): 456–466.

———. 1999. "Uniform Spreading Of Amplitude Panned Virtual Sources." *In IEEE Workshop on Application of Signal Processing to Audio and Acoustics. New York*, United States of America.

Reas, Casey, and Ben Fry. 2007. *Processing: A Programming Handbook for Visual Designers and Artists*. Cambridge, Massachusetts, United States of America: MIT Press.

Rumsey, Francis, and McCormick. 2006. "Auditory Perception." In *Sound and Recording: An Introduction*, 25–40. Fifth ed. Oxford, United Kingdom: Focal Press.

Scheil, Kevin. 2010. "Relationship-Based Instrument Mapping of Multi-Point Data Streams Using a Trackpad Interface." In *Proceedings of New Interfaces For Musical Expression*. Sydney, Australia.

Traux, Barry. 1999. "Composition and Diffusion: Space In Sound In Space." *Organised Sound*, Cambridge University Press 3 (2): 141–146.

Wright, Matt, Adrian Freed, and Ali Momeni. 2003. "Open Sound Control: State of the Art 2003." In *Proceedings of New Interfaces For Musical Expression*, 153–159. Montreal, Canada.

# Appendix 1: Related Publications

Johnson, Bridget, and Ajay Kapur. 2012. "*Tactile.space*: A Multi-touch Tool For Live Sound Diffusion." In *Proceedings of Australasian Computer Music Conference*. Brisbane, Australia.

# Appendix 2: Processing Class Model



**class Class Model**

**«entity» Spreader**
+ updateSpreader(int) : void

**«entity» SpeakerRouter**
- routedSpeaker: int
+ vbapEight(int) : int
+ vbapRouting(int) : int

**«entity» Playback**
- play: boolean
- playColour: color = 22,147,22
- stopColour: color = 147,22,22
+ drawButtons() : void
+ pressedPlay() : void
+ pressedStop() : void

**«entity» ImmersiveFix**
+ numOfTracks : int
+ draw() : void
~ fingerResponse() : void
+ setup() : void
~ spaceCheck() : void

**«entity» TrackObjects**
- objectColour: color[]
+ drawArcObjects(int) : void
+ drawObjects() : void
+ initObjectPos() : void

**«entity» PolarConversion**
+ angle(int, int) : float
+ cartesianX(float, float) : int
+ cartesianY(float, float) : int
+ distance(int, int) : int

**«entity» OscHandler**
- myRemoteLocation: NetAddress
- oscP5: OscP5
+ init() : void
+ oscFadeRate(int) : void
+ oscSendDistance(int, float) : void
+ oscSendGain(int) : void
+ oscSendState(int) : void

**«entity» ControlArea**
- objectInSpace: boolean[]
- x: int
- y: int
+ display() : void
+ drawSpace(int, int, int) : void
+ fingerInObject(TuioCursor, int) : boolean
+ getRadius() : int
+ insideSpreader(int, int, int, TuioCursor) : void
+ inSpace(int, int, int) : void
+ movement() : void
+ resetObject(int, int, int) : void
+ spreaderInitiate(long) : void
+ spreaderRemove(long, int) : void
+ spreaderUpdate(int) : void
+ weightGains(int) : void

**«entity» VBAP**
- firstGain: float
- pairWidth: float = 360/numSpeakers
- secondGain: float
- vbapPairs: int[][]
+ findSpeakerPair(float, int) : void
+ overOrigin(boolean, float) : float
+ spreaderWeightGain(int, float, float) : void
+ trigPanningLaw(float) : void
+ vbapPairing(int, int) : void
+ vbapRun(int, int, int) : void

**«entity» Speakers**
- euroOffSet: float = speakerDivision/2
- euroStandard: boolean = true
- speakerColour: color = 36,103,124
- speakerDivision: float = 360/numSpeakers
+ positionSpeakers(int) : void

**«entity» TUIO**
+ addTuioCursor(TuioCursor) : void
+ removeTuioCursor(TuioCursor) : void
+ updateTuioCursor(TuioCursor) : void