2013

Music and spatial verisimilitude

Etienne Deleflie

University of Wollongong

Recommended Citation
COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
Faculty of the Creative Arts

Music and Spatial Verisimilitude

Etienne Deleflie

This thesis is presented as part of the requirement for the
Award of the Degree of
Doctor of Philosophy
from
University of Wollongong

May 2013
DECLARATION

I, Etienne Deleflie, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Creative Arts, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The work has not been submitted for qualification at any other academic institution.

Etienne Deleflie
28 May 2013
ABSTRACT

Electroacoustic composers have developed a variety of approaches to the use of space within music. Relatively recent technological and scientific advancements have augmented synthesised space with the promise of verisimilitude: the sensation of spatially authentic audible reality. How does this promise correspond to musical concerns?

The thesis develops a new perspective in which composers' interest in spatial verisimilitude is characterised as more concerned with auditory illusion, than with notions of space. This perspective describes a departure from past compositional engagement with isolated spatial parameters, and signals a movement towards the pursuit of realistic illusion within music. The exploration of this pursuit, and its significant technological mediation, highlights fundamental and complex tensions that affect spatial music composition.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank Adjunct Professor Greg Schiemer who not only taught me how to write, but who has also been extremely generous with his help and guidance throughout the entire thesis.

I am greatly indebted to Associate Professor Brogan Bunt. Under his supervision, I felt a nurturing respect for intellectual freedom that gave me both the room and the support to explore at will. I have found Brogan’s broad command of concepts, and his fast intellect, both inspiring and engaging. It takes half as many words to say something to Brogan. I shall miss the leaps of understanding that result from even a short discussion with him.

Thanks also, of course, to Dr. Paul Doornbusch whose extensive knowledge of both compositional and spatial compositional practice has provided me with the necessary confidence to progress in the development of arguments.

Thanks to my children, Dylan and Wilhelmina, who provided me with smiles all along the process. I sincerely hope that they will somehow benefit from this effort, even if it is just in their father having a better ability to identify and question his own assumptions. And thanks, of course, to Kath for putting up with me at times when few others would. It has given me another kind of freedom: the freedom to continue, which is what this thesis is ultimately indebted to.
# TABLE OF CONTENTS

Declaration.................................................................................................................................................. 3
ABSTRACT.................................................................................................................................................. 5
TABLE OF CONTENTS .............................................................................................................................. 9
LIST OF FIGURES .................................................................................................................................... 11
Preface...................................................................................................................................................... 13
1 Introduction........................................................................................................................................... 17
  1.1 From space affects music to music effects space.............................................................................. 21
  1.2 The context for the research question .............................................................................................. 25
  1.3 Research question and argument .................................................................................................... 28
  1.4 The concept of spatial verisimilitude ............................................................................................... 30
  1.5 Approach........................................................................................................................................ 33
  1.6 Key Terms....................................................................................................................................... 35
    1.6.1 Reality-equivalent technologies ................................................................................................. 35
    1.6.2 Sound ........................................................................................................................................ 36
  1.7 Chapter Summary ............................................................................................................................ 36
2 Portfolio of Works................................................................................................................................. 39
  2.1 Broad contextualisation of my compositional process ................................................................. 40
  2.2 ImageSynth: a spatial compositional environment ......................................................................... 49
  2.3 Sound excerpts and compositions .................................................................................................. 62
  2.4 Summary of research conclusions .................................................................................................. 81
3 Spatial Information Encoding Modes: A Framework ....................................................................... 83
  3.1 Spatialisation techniques understood by perceptual mechanism targeted .................................. 85
  3.2 Existing frameworks for spatial analysis ........................................................................................ 91
  3.3 The framework .............................................................................................................................. 97
  3.4 A detailed example of encoding modes .......................................................................................... 101
  3.5 Spatial encoding modes illustrated by examples ............................................................................ 104
4 Referenced Space ................................................................................................................................. 113
  4.1 Peirce’s theory of signs ................................................................................................................... 115
  4.2 The Spatial Icon ............................................................................................................................ 120
  4.3 Schaeffer’s semiotics and musical meaning ................................................................................... 124
  4.4 Signs of space in music ................................................................................................................ 129
LIST OF FIGURES

Figure 1. Screenshot of ImageSynth ................................................................. 52
Figure 2. Each image acts as a map laid flat, with the listener positioned in the middle. ................................................................. 53
Figure 3. Each pixel can be thought of as a virtual loudspeaker. ...................... 53
Figure 4. File player allows auditioning completed spatialisation renders by streaming audio channels to AmbDec. ........................................ 55
Figure 5. The ‘play’ button instigates real-time rendering for a limited count of pixels. The ‘generate’ button instigates non-real time rendering. ............... 56
Figure 6. Different pixel maps can easily be chosen within the interface. ............. 57
Figure 7. A script-editing window allows editing images using scripts. ............... 59
Figure 8. Image Filters allow highlighting how different images are translated into sound. ............................................................................. 60
Figure 9. Sequences of images are displayed as thumbnails for easy access. ......... 61
Figure 10. Each frame is accompanied by a set of parameters that can be used to affect the rendering of pixels to sound ......................................................... 62
Figure 11. First image for Fluoresce ................................................................. 66
Figure 12. Second image for Fluoresce ............................................................ 66
Figure 13. Part of the sequence of images used to create Drivetrain ................. 68
Figure 14. Part of the sequence of images used to create Lost Corps ............... 70
Figure 15. A grid of just-intonation pitches designed by Kraig Grady and modelled within ImageSynth ................................................................. 79
Figure 16. A sound’s position can be represented as the combination of angular direction and distance ................................................................. 84
Figure 17. Notional outline of steps involved in the perception of sound .......... 87
Figure 18. A simple trumpet score, in which only the loudness changes .......... 101
Figure 19. All spatial encoding modes may co-exist in a single piece ............... 103
Figure 20. The difference in the referential mechanisms of: a recording of a bird and a recording of reverberation, in signifying spatial characteristics ....... 117
Figure 21. Karlheinz Stockhausen operating the Rotationstisch in the Studio for Electronic Music of the WDR, Cologne, in 1959 (Chadabe 1997, p.41) 163
PREFACE

I very quickly relinquished my obsessive interest in jazz in 1994 when, on the recommendation of a friend, I went to see a live performance by the then not-so-well-known electronic music producer Richard James (aka Aphex Twin). For years I had attempted to emulate John Coltrane’s fierce drive and sense of musical exploration, but had always been disappointed by the apparent absence of aesthetic adventure in contemporary jazz performance. On that night in 1994 Richard James engaged in a challenging exchange between filling the dance floor with danceable beats, and emptying it by gradually deploying ever-more complex rhythms. In Richard James, I heard John Coltrane: a sophisticated musical ear supported by an adventurous musical intellect. I realised I was not interested in a musical genre, but in aesthetic exploration.

I exchanged my tenor saxophone for digital technology and began to exercise this newfound interest. After a year of research I began producing live audiovisual performances with a high level of synchronicity between music and video. However, my aesthetic ideas quickly extended beyond the digital tools at hand and I became heavily engaged in technological experimentation. I eventually turned to a software platform called jMAX (jMAX 1996) developed by the Institut de Recherche et Coordination Acoustique/Musique (IRCAM). jMAX was a versatile platform that exposed sound processing as modular objects that could be easily customised and reconfigured. My idea was that if I could adapt jMAX to also process video then I would have an audio-visual environment not limited by the software developer’s ideas. I envisaged a general apparatus that would enable me to freely explore specific aesthetic ideas. I learnt how to program in the “C” language, which was not without its challenges, and eventually developed my ideal environment. IRCAM had decided to open-source jMAX and I became one of the contributing developers. My subsequent work, such as the Maria Island Set (Deleflie 2001) exercised a deeply digital aesthetic in which sound and video held equal status as rivers of digital bits.
I later discovered the work of 20th century composer Iannis Xenakis, “one of the great free-thinkers in the history of music theory” (Roads 2004, p.64). I found, in Xenakis, the same engagement I had heard in John Coltrane and Richard James. Xenakis, however, had developed a compositional approach that resonated with my ideas. His aesthetical explorations were realised by the invention of technique.

During this time I studied Architecture at the University of NSW and, after practising for a number of years, developed an interest in spatial audio. My first creative inclination in the exploration of the musical possibilities of spatial audio was to develop techniques that could expose new aesthetics. One early technique involved placing point sources of sound, which were spatialised, at each vertex of a virtual cube that was animated using the media language Processing (Fry & Reas 2001). The listener was positioned in the centre of the virtual cube, which was then rotated on different axes and at different speeds.

I quickly discovered that there was no clear relationship between technical invention and aesthetic interest. An interesting technical idea does not necessarily have interesting aesthetic implications. This made me realise that Xenakis’ talent lies not so much with his technical creativity, as it does with his intuitive understanding of the aesthetic potential of specific technical inventions. I also discovered that it was easy to project sounds from many speakers, but very difficult to create immersive spatial works. My attention thus turned to the invention of techniques equally informed by both the perceptual aspects of spatial audio and potential musical interest. One of these techniques was developed during the course of the thesis, and it is documented in Chapter 2.

During my engagement with spatial audio, I have been an active member of the surround-sound community. Some years ago I founded the website ambisonia.com (Deleflie 2005), a site dedicated to the dissemination of spatial audio recordings and compositions realised with a spatialisation technique known as ambisonics. This site became an important resource within the ambisonic community and is now hosted at the University of York, which generously offered to take over its management. Later, I also founded soundOfSpace.com (Deleflie 2009), designed to facilitate listening to
spatial audio by offering streaming multi-channel audio. The development of both these sites was initially motivated to facilitate the publication of my own work, but both sites also became community resources.

It is the diversity of the spatial audio works published on ambisonia.com that cemented my interest in spatial music. What is striking is that these works invariably reveal one of three distinctly different interests: firstly, the recording of classical music performances; secondly, field recordings unconcerned with music; and thirdly, the production of novel electroacoustic composition. The grouping of such disparate interests indicates a powerful overarching umbrella that is capable of hosting very different approaches to sound. It is the desire to identify that umbrella and understand its aesthetic movement that drove me to formalise my interests within this thesis. I found that the uniting umbrella is the sonic illusion of space, and that its relation to music is indeed complex.
1 INTRODUCTION

Over the past decade, since developing an interest in spatial music, I have found that there has always been something engaging about the idea of music in space, just as there has always been something disappointing with the experience of it. Entering a concert hall sporting multiple speakers arranged around the listening space sparks expectations of sonic immersion and transportation to new spaces and places. And yet, I have often found that music not explicitly concerned with space can be more immersive and transporting than its spatial counterpart. At first, I attributed this to poor ‘spatialisation technique’, but this does not explain how non-spatial music, such as composer György Ligeti’s orchestral work Atmosphères (1961), can create such a powerful sense of spatial immersion.

Perhaps I would not have been disappointed with the experience of spatial music had I first been exposed to seminal spatial compositions such as American composer Henry Brant’s Antiphony I (1953) in which five orchestras are spatially separated; or early pioneer of electroacoustic music Karlheinz Stockhausen’s Kontakte (1960) in which spatially-rich sounds pre-recorded to multi-track tape are routed to four loudspeakers (Stockhausen & Kohl 1996, p.92); or technical innovator John Chowning’s Turenas (1972) in which the movements of sounds are digitally simulated and also projected through four speakers. Each of these works, further discussed in this introduction, has been celebrated as a fine example of spatial music (Zvonar 2004b).

Despite these seminal works, my disappointment in spatial music is not unique. As highlighted shortly, the suggestion that the experience of spatial music can fall short of expectations does exist in contemporary critical enquiry. The key to understanding this suggestion lies in understanding what these expectations are. One testament can be found in the editorial to an issue of Organised Sound (Myatt 1998), an

---

1 Electroacoustic composer Denis Smalley (in Austin 2000, p. 20) describes some of Gyorgi Ligeti’s and Iannis Xenakis’ orchestral music as good examples of “really spatial orchestral music”.
international journal that “focuses on the rapidly developing methods and issues arising from the use of technology in music today” (Landy 2012). The issue focuses on the spatialisation of sound. Myatt discusses this issue specifically in relation to music:

When I listen to surround-sound or three-dimensional (3D) sound systems, I have to admit that there are only rare occasions when I am convinced that 3D sound images or 3D movements exist. When they do, the affect is startling but lasts only for a brief moment of time – rarely do I experience any consistent depth of illusion or the semi-immersive audio environment that I expect to be reproduced (particularly, though not exclusively, in concert performances). (1998, p.91)

This corresponds to my own experience of spatial music, and indicates an increasingly important conception of the relationship between space and music. In this conception the spatial characteristics of spatial music have assumed the form of illusion. Myatt’s use of the word illusion is clarified by his stating that he expects to be convinced that such things as 3D sound images and movements actually exist. In other words, there is the expectation that the spatial attributes of music take on the characteristic of verisimilitude: that is, the semblance of spatially authentic audible reality.

This is far from the kind of spatiality present in Ligeti’s Atmosphères in which the spatial immersion experienced is crafted through musical texture rather than technologically realised spatial projection. In Atmosphères, impressions of sweeping spatial movements are created through instrumental dynamics and orchestration of timbre. The notion of spatiality Myatt discusses is also not relevant to Brant’s Antiphony I where the spatial characteristic of what is heard is no illusion; it is real. The physical separation of the five orchestras does not serve to create an impression of space; rather, it uses the existing space to affect music in such a way as to highlight “the contrasts of timbre, meter, key, texture, and motivic content between the five musical streams” (Harley 1997, p.71). In other words, Brant uses the way spatial separation changes sound to help listeners perceive differences in the separate musical parts. As a composer of spatial music, Brant is not concerned with creating
impressions of space, let alone doing so with verisimilitude\(^2\), but rather he works with space directly.

Brant’s compositions are not the only works to use the *effect of space on sound*. Indeed “spatial relationships between musical performers has always been an integral part of performance practice” (Zvonar 2004b, para. 1) and its exploitation to compositional effect can be seen in the work of many 20\(^{th}\) century composers. Bela Bartok, Pierre Boulez, Karlheinz Stockhausen, Iannis Xenakis and John Cage, amongst others, all explored the effect of space on sound as a compositional device (Harley 1994b, 2000). In Bartok’s *Music for Strings, Percussion and Celeste* (1936-1937), for example, the proper placement of performers on the stage results in a connection between symmetries in musical form and symmetries in performance space (Harley 2000, pp.157-158). Cage (1961, p.39), for his part, sought to avoid the harmonious fusion of musical parts that he saw as characteristic of European musical history. In a similar way to Brant, he used the spatial separation of performers to support the perceptual isolation of different musical parts (p.39).

Myatt’s statement suggests an entirely different concern with spatiality; one in which spatial characteristics are authored *with* and *within* the music. These spatial characteristics likely have no relation to the listening space, and are manifest only as a result of hearing the music. On the one hand, this can be seen as an inversion of the above-described compositional engagement with space: instead of space supporting the perception of the music, it is music that supports the perception of space. On the other hand, the notion that music can create impressions of space is not new. Whilst Ligeti’s *Atmosphères* is not typically considered to be ‘spatial music’, it creates spatial impressions musically. The difference between Ligeti’s work and this new conception of spatial music is, of course, that the spatial impressions created by *Atmosphères* do not have the characteristic of verisimilitude. In Myatt’s discussion, the spatial impressions created occur as illusions, that is, we are to some extent tricked into believing that they are real.

\(^2\) For a detailed account of Brant’s approach to spatial music see (Brant 1967).
This thesis is concerned specifically with such spatial music that can be defined as having two key attributes. Firstly, spatiality is authored through the music that is heard, and secondly, the spatial characteristics are perceived, or perhaps more accurately, are designed to be perceived as real. The perception of being real can be understood by paraphrasing Myatt: we are ‘convinced’ that spatial characteristics ‘exist’. Both these attributes need to be defined because the spatiality in the work of composers such as Brant and Cage already has the appearance of being real since it is real. What this thesis is concerned with is the attempt to manufacture realistic perceptions of space exclusively within music. The emphasis on music is also important: it distinguishes this thesis’s concerns from other contexts, such as computer gaming or virtual reality applications, in which the pursuit of spatial verisimilitude in sound does not necessarily hold a musical dimension.

Whilst spatial music has a rich and diverse history (see Harley 1994a; Zvonar 2004b) the attempt to create highly realistic impressions of space is very recent. It is a development that, of course, coincides with technological advancements. The 1998 issue of the journal Organised Sound, described above, roughly correlates with the beginning of an era where greater spatial engagement is made possible by “cheaper sound cards, powerful software with automation tools and the availability of standardised multi-channel systems” (Otondo 2008, p.80). These expanding technological means, in combination with advancements in audio engineering and understanding of psychoacoustics, bring the promise of spatial verisimilitude in sound closer to the spatial composer. It is the composer’s transferral of this promise to music that introduces new and difficult questions.

One of these questions concerns the relationship between music and realism more generally: will all musical subjects and musical forms have equal capacity to be projected to a high level of realistic spatial illusion? Myatt (1998, p.91) considers the quality of spatial illusion independently of the musical material that carries it. Is it possible that certain musical material has a lesser capacity to be realistically projected in space? Another question concerns potential constraints: does the realistic spatial presentation of music impose any limitations on the music itself? Such
questions have received relatively little critical attention and, as is discussed later, the range of views is broad with consensus rare.

This thesis seeks to address these questions and lay the foundations for a scholarly enquiry that considers how compositional concerns are mediated by the pursuit of spatial verisimilitude.

1.1 From space affects music to music effects space

Whilst the composer’s concern with spatial verisimilitude has been catalysed by technological means, interest in authoring space in music is evident throughout the electroacoustic music of the 20th century. It is not the aim of this thesis to provide a detailed history of the evolution of this interest; however, a consideration of this trajectory, starting with the work of one prominent electroacoustic composer, Karlheinz Stockhausen, offers valuable insight into how spatial verisimilitude has come to be a compositional concern.

In Stockhausen’s work, there is no clear separation between the use of space to affect music and the use of music to author perceptions of space. Indeed, Stockhausen’s initial attempts to create impressions of space are motivated by the desire to have more practical ways to explore how space affects sound. Unlike Brant who was interested in spatial separation, Stockhausen was particularly interested in spatial movement. He saw movement in space as a compositional parameter equivalent in importance to other major compositional parameters:

And this movement in space of music becomes as important as the composition of its melodic lines, meaning changes in pitch, and as its rhythmic characteristics, meaning changes in durations. If I have a sound of constant spectrum, and the sound moves in a curve, then the movement gives the sound a particular character compared to another sound which moves just in a straight line. (1989, p.102)

In a 1971 lecture (Stockhausen & Maconie 1989, pp.101-102), Stockhausen describes how in 1953 he conceived the possibility of placing musicians in chairs that could be swung around. He subsequently enquired about wiring up speakers also to be swung around. Both options were motivated by the desire to explore the sound
of spatial movement and both were rejected on practical grounds. Several years later Stockhausen engaged in an entirely different technique for exploring spatial movement. In *Gruppen für Drei Orchester* (1955–7) and *Carré* (1959–60); “Stationary instrumental groups are placed around the audience and successively play sounds of the same pitch and timbre with similar dynamic envelopes (crescendo–decrescendo)” (Harley 2000, p.151). By carefully scoring changes in loudness of successive instrumentalists, Stockhausen could create something of an *illusion* that sounds are rotating in space. Of course, the effect is only a crude approximation of a moving sound, but it allows the exploration of spatial movement within compositional concerns, whilst avoiding the impracticalities of real movement.

This kind of approach signals a shift in the compositional engagement with space. Stockhausen is interested in how spatial movement affects sound, but these affectations are now artificially re-created. Perhaps the word ‘synthesised’ could be used to describe these re-creations, although ‘synthesised’ might best be reserved for a more mimetic representation of the original spatial affectations. Within the context of Stockhausen’s work, this shift is subtle but the engagement in re-presenting spatial characteristics launches the composer into a very different process.

For Stockhausen the motivation for attempting to re-create movement involves the exploration of the effect of space as a compositional parameter. The argument presented here is that engagement with the *representation* of spatial characteristics can ultimately lead to the kind of spatial concern indicated by Myatt; how convincing can the spatial representation be made? The pursuit of convincing spatial illusion involves a very different relationship to music composition. Within Stockhausen’s concerns representations of spatial movement are measured by their worth as compositional parameters. In Myatt’s concerns, however, representations of space are measured by their ability to produce convincing illusions. Within this later conception, spatial characteristics become a product of the music heard, and their contribution to the composition moves beyond musical parameter and introduces an element of *realism*. Here, another difficult question is raised: how can realism be used as a musical parameter?
In asking this question, perhaps the term ‘realistic representation’ might be more appropriate than the word ‘realism’, which is associated with 19th century visual and literary art movements and implies a concern with objective truth. The manufacture of convincing spatial illusions is not so much concerned with objective truth as it is with creating the semblance that what is heard is real. What this question highlights, however, is that by engaging in the pursuit of realistic re-presentation of space in music, the relationship between spatial concerns and the composition is altered. Spatial illusion becomes an end in itself whose role within the composition must be negotiated.

Stockhausen’s exploration of spatial movement eventually extended into the technological domain. In Kontakte (1960), a work that became a “reference for the organic relationship between sound synthesis and sound space” (Chagas 2008, p.191), a multi-channel tape recorder is used to capture sounds emanating from a rotating loudspeaker. Here, spatial movement is not so much ‘faked’ as it is technologically captured and subsequently technologically reproduced. This technique, discussed in greater detail in Chapter 6, deepens Stockhausen’s engagement with the representation of space in sound. Through the use of multiple microphones, Stockhausen succeeds in capturing not just the movement of the rotating speaker, but also some of the acoustic characteristics of the room (Harrison 1998, p.126). In other words, the spatial effect takes on extra dimensions, describing not only movement but also a specific acoustic environment. Here, Stockhausen’s engagement with space extends beyond his interest in spatial movement, and includes a range of represented spatial characteristics. Interest in the effect of space on music thus begins to be eclipsed by the pursuit of spatial representation within music.

Some years later, compositional interest in spatial representation escalated as a result of the work of John Chowning. Chowning’s Turenas (1972) is notable for its extensive use of synthesised spatial movement and long reverberant sounds. Chowning documents the techniques he employed in this composition in his article, first published in 1971, “The simulation of moving sound sources” (Chowning
Zvonar (2004a, para. 9) describes this paper as “a watershed between the intuitive and empirical spatialisation work of the 1950s and 1960s and subsequent rationalistic approaches to computer music systems design”. In his paper Chowning (1977) describes how a computer was used to simulate various aspects of the behaviour of sound in space. The language used in the paper itself highlights a very different musical concern with space. Words such as ‘simulate’ and ‘synthesise’ and phrases such as ‘convincing spatial images’ all confirm the shift towards spatial representation as a way to engage with spatial music. However, Chowning’s measure of ‘convincing’ still falls far short of Myatt’s characterisation articulated almost 30 years later. Chowning seeks to synthesise spatial images that are “comparable to a good stereophonic or four-channel recording” (Chowning 1977, p.48). In other words, he seeks to convince that the spatiality heard was real prior to being captured on tape; he does not seek to convince that it exists in the current listening environment.

This gradual shift, in the compositional engagement with space, from exploring space’s effect on sound towards the pursuit of ever-more realistic representations of space, is confirmed in a 2001 paper written by Dave Malham, a researcher in the Department of Music at the University of York. “Toward reality equivalence in spatial sound diffusion” (2001a) was published a few years after the 1998 issue of Organised Sound. In it, Malham discusses the challenges faced by sound systems that aim to achieve “reality-mimicking performance” (p.32). Again, as in Chowning’s paper (1977), it is Malham’s choice of language that characterises his concern with space. His interest is in the synthesis of “full reality” sound (2001a, p.32). Here, the concern with spatial representations in space has progressed to its ultimate conclusion and spatial sound systems are now measured by their ability to produce results that are “indistinguishable from reality” (p.37). Malham’s paper is important in that it captures and characterises composers’ contemporary expectations of spatial projection systems. Unlike previous key figures discussed thus far Malham is not a composer. The significance of his paper thus also confirms that the pursuit of realistic spatial illusion, in music, has become a goal in itself, considered independently of specific compositional concerns.
1.2 The context for the research question

This short historical thread on the evolution of the compositional interest in space has thus come full circle. Malham’s expectations of spatial projection systems fall into alignment with both Myatt’s expectations of spatial imagery in concert music, and mine. All anticipate a high level of realistic spatial illusion within a musical context.

Herein lies the first of three key characteristics that describe the context of the research question: Spatial verisimilitude is a quality that is anticipated in spatially projected music, but that anticipation is not directly concerned with the specific musical context. One indication of this disassociation is that poor spatial imagery in music is rarely attributed to the musical material itself; it is typically attributed, as Myatt does (1998, p.91), to poor spatialisation technique or poor understanding of psychoacoustics. This highlights a very important point. In the proceeding discussion a range of views on the relationship between music and spatial verisimilitude is outlined. There is an underlying implication common to many of these views that the concerns of spatial verisimilitude and music are independent of one another, yet somehow sympathetic to each other. I have found no evidence of this being the case. It is perspectives centred on the technological aspect of spatial verisimilitude that most clearly enunciate this position. Malham, for example, states that outside of cases where musical forms do “not rely on specific acoustical locations or timbres to achieve their compositional effect (for instance, a simple melodic line)” (2001a, p.32):

[The] more capable our sound diffusion systems are of mimicking reality, the more options composers will have for exploring the use of spatial elements within their music (2001a, p.32)

In other words, outside of certain musical forms that have no meaningful association with space, spatialisation technologies are seen as an empowerment to the composer.

This view seems eminently sensible, but is quickly challenged by arguments presented by Gary Kendall et al (2001, p.2460) who highlight that there are some specific incompatibilities between the accurate modelling of sounds in space and certain musical elements. Indeed, in what is almost an inversion of Malham’s view,
they criticise spatialisation technologies for adhering to acoustic reality at the expense of serving musical concerns:

Experience with such audio tools reveals some issues of special concern for music. For example, the rapid movement of musical sound sources can create Doppler shifts that produce harsh detunings of pitch (2001, p.2460)

Through these ‘detunings of pitch’, which are a characteristic of moving sounds, important musical devices such as melodic lines can be corrupted. In other words, contrary to Malham’s position, a melodic line does have a meaningful association to space; it presumes a static physical location. As such, Kendall et al argue that audio spatialisation tools should serve musical interests “without slavishly adhering to models of acoustic reality” (2001, p.2460). They do not suggest that the movement of melody should be excluded from spatial music, but rather that the spatial verisimilitude should be compromised to accommodate it. In such an example, spatial verisimilitude does not serve a musical intent, but rather challenges it.

Without directly identifying them, Malham alludes to such problems. He makes an important qualification to his proposition that the spatialisation technologies he discusses empower the composer. He explains that successfully mimicking reality effectively imposes restrictions on how sounds are presented:

For a system to produce reality-mimicking sound, it is necessary for all departures from reality equivalence to be below the relevant thresholds of perception. (2001a, p.32)

As is illustrated by the use of spatial movement with melody, this restriction has implications for composers. Refraining from modelling Doppler shifts in moving sounds is one example of a perceptible ‘departure from reality-equivalence’ that is motivated by musical concerns. Malham’s qualification is therefore musically significant, and leads to the question: what other aspects of music and musical form can be considered ‘departures from reality equivalence’? Of course, faced with such a situation, it is the composer’s prerogative whether a musical form is retained at the expense of verisimilitude, or rejected in the interest of maintaining reality-
equivalence. Here, however, it is clear that the composer must answer the question of what should take priority: spatial verisimilitude or musical concerns?

Thus the concerns of spatial verisimilitude and music are not independent of one another, and are not necessarily sympathetic to each other. Simon Emmerson, a writer and scholar of electroacoustic music, does not go so far as to state this point; but he does identify and reject the assumption that spatial verisimilitude might be of service to musical intent:

Over recent decades there have been persistent claims that ‘perfect reproduction’ of a soundfield is within our grasp. But it remains an open question as to whether this is actually an ideal [musical] aim. (2007, p.143)

Emmerson’s statement essentially suggests that the relationship between music and spatial verisimilitude, which is the central concern of this thesis, has not yet been adequately investigated. As is confirmed by the above discussion, there remains a range of views that approach this relationship from different perspectives. What is clear is that this area of critical enquiry requires further research, and that different perspectives require a level of normalisation.

The second key characteristic that describes the context of the research question concerns the practical accessibility of spatial verisimilitude. Here again, there is a correlation between my expression of disappointment in spatial illusions, Myatt’s observations, and Malham’s discussion of spatialisation technologies. Malham states that no systems are “yet” capable of achieving results that are “indistinguishable from reality” (2001a, p.37). Thus, whilst contemporary technologies promise access to realistic spatial illusion, there is the recognition that convincing spatial illusion in sound is presently evasive. Whilst I use the word ‘presently’, arguments presented in Chapter 5 suggest that realistic spatial illusion may never be more than a holy grail. As such, the compositional engagement with spatial verisimilitude may forever include an active struggle to capture realistic illusion.

The third key characteristic concerns the way composers approach the authoring of spatial verisimilitude. Realistic spatial illusion is achieved technologically, and this
technological dimension is complex both in terms of its understanding of psychoacoustics and its practice of audio engineering principles. This has two implications; the notion of spatial illusion in musical pursuits is largely limited to a technological encapsulation, and the compositional engagement with it is thus subject to an extensive technological mediation.

In summary: for the composer, the idea of spatial verisimilitude turns out to be not so much an attainable quality as it is a technologically mediated pursuit, with inconsistent results and, above all, with a little-understood relation to musical concerns.

1.3 Research question and argument

It is within this characterisation of the compositional engagement with realistic spatial illusions, as well as in terms of my own dedicated experimentation in this field, that my specific research question emerges. The thesis sets aside any suggestion that technology might eventually conquer spatial illusion, it rejects the assumption that verisimilitude is of clear musical interest and asks: what is the relationship between music composition and the pursuit of spatial verisimilitude?

In essence, the thesis is concerned with a compositional process in which an elusive and musically uncharted quality, spatial verisimilitude, is pursued technologically. The research question thus holds several dimensions. First is the relationship between music and representations of space, both realistic and otherwise. Second is the relationship between sound and illusion, which is related to the notion of fidelity in audio recording. Third is the role of spatialisation technologies in mediating the compositional process: how do these technologies influence both spatial conception and musical conception?

The structure of the thesis follows these dimensions, each explored in their own chapter. Of the three, however, it is the first that resonates throughout the entire thesis. The difficult dynamic between music’s general capacity for representation and the composer’s pursuit of realistic spatial representation underlies many of the ideas developed. It is the identification and articulation of this dynamic that is the main concern of this enquiry. The thesis argues that the musical material being spatialised
has a complex relationship with its own spatialisation. In effect, there is a tension between music and spatial verisimilitude, and this tension indicates the presence of competing concerns.

This argument clears the way for a new understanding of the relationship between music and space. In this understanding: firstly, the composer can no longer jettison the responsibility for realistic spatial illusion to technology; secondly, music is understood to already have a broad capacity for representing space outside of realistic illusion; and thirdly, the composer must answer the question, in musical terms, of how realistic representation might relate to the representation present in musical form. This last point is perhaps the most difficult one for composers to engage in. The issue of the musical meaning of realistic illusion is one that has attracted very little attention but has been identified by some, as is evidenced by the following discussion between (American composer) Larry Austin and British composer Ambrose Field, in which the spatialisation technology known as ambisonics is discussed:

Austin: That was what was distracting me. I was disconcerted when I first heard ambisonics, because it was too ‘‘real.’’
Field: This is a big issue. ‘‘Too real.’’ I absolutely agree with you. […] There is a problem there with the reality aspect of it. Reality is always a problem, though, isn’t it? (Field in Austin 2001, p.28)

The question of exactly why reality is a problem, within a musical context, is of key interest to this thesis. Here, the thesis contributes to this line of enquiry not just by characterising this issue, but also by attempting to uncover its cause: why should the illusion of reality pose a problem to musical endeavour?

It is worth noting that the problem of ‘reality’ within music is one not isolated to technologically realised illusion. Expressed from the concerns of certain 19th century composers, any association between ‘reality’ and music is seen as problematic:

Composers like Ferruccio Busoni, Arnold Schoenberg and Kurt Weill, none of them starry-eyed rhapsodists, all held the opinion that music is a fundamentally
unrealistic art, and that therefore the concept of musical realism represents an error in the thing so designated or in the judgment formed of it. (Dahlhaus 1985, p.10)

Of course, any 19th century conception of realism must consider the visual and literary movements of the same era, and this is touched on later, but this passage suggests two things: firstly, the difficult relationship between reality and music exists outside of the immediate concerns of this thesis; and secondly, that an exact definition of what is here meant by ‘reality’ is required. This is forthcoming.

1.4 The concept of spatial verisimilitude
The preceding discussion highlights that researchers employ a wide variety of terms to indicate interest in what is here called ‘spatial verisimilitude’. Emmerson’s ‘perfect reproduction of a soundfield’, Malham’s ‘reality-equivalent sound’ and Chowning’s ‘convincing spatial images’ all point towards a concern with the convincingly realistic representation of space in sound, but each term is employed in a different context and holds a slightly different emphasis. Within my own writing, the term ‘realistic spatial illusion’ is used synonymously with ‘spatial verisimilitude’.

At this point, a strict definition of the term ‘spatial verisimilitude’ is required to clarify the specific understanding of spatial projection of sound concerned, how it is achieved, and what is meant by ‘realistic’. This definition extends over the following pages but, to begin, it is Myatt’s description of his expectations of spatial audio that most concisely approaches the intended meaning. Myatt expects to be “convinced that [spatial characteristics] exist” (1998, p91). In line with this expectation, spatial verisimilitude is here defined as the illusion that the space or spatial characteristics heard actually exist.

In this definition, the word illusion is of central importance. It is used to indicate that the representation is, to some degree, actually taken to be the thing represented. In other words, the realism in the representation moves beyond a mere likeness and results in a suspension of disbelief. Through this suspension of disbelief the representation is itself dissolved and what is perceived is accepted as fact. In other words, when spatial verisimilitude is achieved, consciousness of the representation ceases. The notion of moving beyond representation can be understood in the terms
employed by 20th century philosopher Jacques Derrida in his article ‘Différence’ (1982). Derrida describes representations as involving both a spatial difference, and a temporal deferral (pp. 1-27) between the representation and the notional original. Spatial verisimilitude refers to the experience of space where this difféance, between the representation and the original, has ceased thus creating immediacy in that experience.

Spatial verisimilitude can be described as the impression that the spatial representation heard in the sound is real. One approach to achieving this impression involves cultivating realism within the representation. Here, and elsewhere in this thesis, the word realism is employed in the colloquial sense of the word, which is distinct from the visual and literary arts movement of the 19th century known as realism. The relationship between music and 19th century realism, which has a specific concern with objective truth, is discussed later in Chapter 5. Notwithstanding the inherent difficulties of comparing music with the other arts, spatial verisimilitude is perhaps closer to what is known as illusionism in the visual arts, in which spectators are momentarily deceived into believing that a pictorial representation is real. Within music the pursuit of this deception or illusion is related to the notion of fidelity, which, as discussed in Chapter 5, originates from the advent of audio recording technology. Fidelity in sound is not so concerned with what is recorded, but rather that it appears real when played back. In other words, the emphasis is on the perceptual fidelity, of the representation, to a ‘notional’ original.

The above definition of spatial verisimilitude resonates with the term presence as defined by Lombard and Ditton (1997) within the context of research on interactive mediated environments. Indeed, Lombard and Ditton’s understanding of presence is used throughout the thesis. For them, presence is the “illusion that a mediated experience is not mediated” (para. 1). This definition encapsulates the above discussion: what is represented is taken, in some form, to be real instead of mediated. Lombard and Ditton’s understanding of presence also defines the extent of the illusion, and this definition is here adopted and applied to the term ‘spatial verisimilitude’. For Lombard and Ditton presence does not occur in degrees; it occurs in instances that have a greater or lower frequency during the experience of
the mediated environment (section "Presence Explicated"). Similarly, spatial verisimilitude is understood to be a quality that occurs in ‘instances’. Thus, when the suspension of disbelief collapses, what is heard returns to the status of representation. This movement, between illusion and representation, is discussed further in Chapter 3. Lombard and Ditton stress that presence does not consist of an illusion that is consciously confused with the real world. Indeed, none of the definitions presented above propose that illusion reach identity with the experience of reality. What is of concern is not the approach towards reality, but rather that the representation is accepted as real by the perceiver. This acceptance might be fleeting, or might resonate and persist, but it does not require the loss of consciousness of the mediation.

The characterisation of successful illusion as the **dissolution of the representation** leads to the last concern to be discussed here: how is this illusion to be manufactured? Contemporary spatialisation technologies aim to convince that spatial characteristics are real by, in the words of Malham, “mimicking reality” (2001a, p.32). This involves the scientifically informed and technologically mediated imitation of the behaviour of sounds in space. Again, such an approach resonates with naturalism in the visual arts, which involves strict adherence to the physical appearance of nature. However, Lombard and Ditton’s (1997) discussion makes it clear that mimicking nature is not the only way to achieve presence. Other approaches such as perceptual and psychological immersion (section “Concept Explication”), for example, can create presence. Here, an intriguing question arises: To what extent might the listener’s immersion in musical form itself be capable of producing ‘presence’?

This last question suggests one possible source for the tension between music and spatial verisimilitude. The notion of spatial verisimilitude, here defined, is very close to Lombard and Ditton’s (1997) notion of presence. But Lombard and Ditton’s notion of presence includes the result of techniques that are closer to the effects of music, than to ‘mimicking reality’. If both spatial verisimilitude and music aim to create presence, but in different ways, then the pursuit of spatial verisimilitude can be
seen as competing with, rather than supporting, music. I develop this crucial argument further in Chapter 5.

1.5 Approach

Insight is drawn into the research question by first engaging in an extensive spatial music compositional effort. This effort includes the design and development of powerful software tools for both spatialising sound, and composing with that spatialised sound. Many of the arguments researched and presented in the thesis thus stem from my own practice as a composer of spatial music. The issues identified during the realisation of the portfolio of works are then used to inform the direction of the research.

There are two principal challenges involved in this research. The first is that many disciplines are involved, and the second is that the relationship between spatial illusion, achieved technologically, and music is a relatively new field. Whilst the research question outlines a very specific line of enquiry, the breadth of disciplines involved require an interdisciplinary approach that brings together very different theoretical perspectives. In this sense, the thesis can be understood to be concerned with the identification of different possible approaches to understanding the relationship between spatial verisimilitude and music composition.

The primary discipline concerned is, of course, spatial composition. It is the addition of realistic illusion, to the concerns of the spatial composer, that results in the substantial expansion and integration of research areas. For example: the perception of illusions concerns cognitive psychology and is referenced in the study of interactive mediated environments and virtual-reality; the consideration of how to represent ‘reality’ involves philosophical debate on what ‘reality’ is; the relationship between illusion and sound references the history of sound recording devices; the behaviour of sound in space concerns the broad field of acoustics; and the use of technology to manufacture simulations of acoustic space requires that the relationship between technology and music be considered. In exploring these different disciplines, my approach has been to focus on the pursuit of insights that can advance the understanding of the compositional engagement with spatial verisimilitude, as opposed to scanning all of the literature related to the numerous
disciplines and fields noted above. As already stated, it is only the literature concerned specifically with spatial composition that is comprehensively covered.

Some insights are drawn from the work of other composers working with spatial verisimilitude, but these insights are limited for two reasons: the first is that interest in spatial verisimilitude in composition is relatively recent; the second is that access to spatial verisimilitude works is constrained by the technically rigorous demands of soundfield playback which, if poorly setup, may misrepresent the composer’s work especially in terms of its verisimilitude.

There is relatively little published research that deals specifically with the use of spatial verisimilitude in music. There is also relatively little composed material that can be accessed\(^3\). It is thus difficult to contextualise concerns. Two notable contributors to the area are British composers Ambrose Field and Natasha Barrett. Both have produced works that have a component of spatial verisimilitude, and both have published research discussing their compositional approaches to verisimilitude. Certainly, more and more composers have started working with spatial verisimilitude in recent years and contemporary spatialisation tools have proliferated. But, as stated by Barrett, “the understanding of spatial issues, at least among composers, is still not so advanced” (Otondo 2007, p.17). New conferences and events, such as the Ambisonics Symposium first held in 2009, are beginning to appear. Whilst compositional works are performed at these events, most of the research presented concerns the technical challenges surrounding spatial audio. The composer’s understanding of the issues involved in pursuing spatial verisimilitude is thus advancing slowly. Other contributors who have published research concerning the compositional aspect of spatial verisimilitude include David Worrall, Peter Lennox, Tony Myatt and Gary Kendall each of whose work is referenced in the following chapters.

Composers of very different aesthetic and stylistic concerns have explored space. Of course, the technological nature of the contemporary pursuit of verisimilitude will

\(^3\) One of my motivating concerns for founding both ambisonia.com and soundOfSpace.com was precisely to facilitate the distribution of and access to spatial audio work that is concerned with realistic illusion.
have an impact on aesthetic concerns. It is predominantly composers of electroacoustic music who will embrace a technological opportunity and, in keeping with this statement, the portfolio of works engages with the concerns of electroacoustic music. These concerns, and how they relate to the research question, are first highlighted in Chapter 2 and further explored in Chapter 5.

1.6 Key Terms

1.6.1 Reality-equivalent technologies

The term ‘reality-equivalent technologies’ is used to denote any spatialisation technology that pursues spatial verisimilitude. These technologies typically focus on simulating the most perceptually important aspects of the behaviour of sound in space. The term ‘reality-equivalent’ is taken from Malham’s 2001 paper ‘Towards reality equivalence in spatial sound diffusion’. Malham (p.37) states that there are no technologies yet capable of producing sound that is “indistinguishable from reality”. Others corroborate this view (Rumsey 2002, p.653) and argue that identity with reality may never be possible for both practical and technical reasons. As such, technologies that produce spatialised sound that is truly equivalent to reality may never exist. For the purposes of this thesis, the term ‘reality-equivalent technologies’ is used to denote a class of technologies that merely have a concern with creating realistic spatial illusions.

It is important to note that the use of reality-equivalent technologies does not guarantee any level of illusion. In this sense, when composers engage with reality-equivalent technologies they are exposed to a wealth of perceptual issues that they may not be equipped to negotiate. Indeed, this point forms a significant characteristic of the compositional engagement with reality-equivalent technologies. The specific ways in which reality-equivalent technologies mediate the compositional process is explored in Chapter 6.

An in-depth description of these technologies is outside the scope of this thesis. A good summary of the most commonly used ‘reality-equivalent’ technologies can be found in Malham (2001a). Some of these technologies are listed and briefly described below:
- **Ambisonics.** Ambisonics (Gerzon 1973) is a technique invented in the 1970s by Michael Gerzon, a British researcher who worked at the Mathematical Institute at the University of Oxford. It is a technique that uses the combined signals from multiple speakers to reconstruct sound waves at the centre of a speaker array.

- **Wave Field Synthesis (WFS).** WFS uses many speakers to create a soundfield that has some advantages over ambisonics; the reproduced sound image extends beyond the centre of the array. It uses Huygens’ Principle which states that any wave front can be synthesised by combining spherical waves from multiple point sources.

- **Binaural renderings.** Binaural renderings use Head Related Transfer Functions (HRTFs) to simulate the effect of the shape of the head, the ears and the body on sound.

The reality-equivalent technology employed in the portfolio of works is the first one mentioned above: ambisonics.

1.6.2 Sound

Whenever the word ‘sound’ is used, it refers to what is heard as opposed to any representation of sound as audio signal. In other words, a pure sine tone is not a sound. To become a sound, a pure sine tone needs to travel through space to reach the ear. As a result of travelling through space the sine tone will have been altered, and will no longer be pure. Given this definition of sound the statement ‘all sound is spatial’ holds true because, to be heard, all sounds must have travelled through space.

This definition alludes to a discussion developed in Chapter 3, in which it is argued that all *perceived* sounds must contain the signature of space.

1.7 Chapter Summary

The portfolio of works is presented first in Chapter 2. This chapter includes details not just on the compositions and sounds created, but also on the significant technical
and technological dimension involved in the compositional act. As each composition is described, the observations it contributes to the thesis’ arguments are identified.

Chapter 3 begins by examining the body of literature associated with the discipline of spatial audio composition. Following this examination I propose a framework that identifies the different spatialisation techniques employed by spatial music composers. The significance of this framework lies in contextualising how the concerns of composing with spatial verisimilitude differ to other spatialisation techniques. For example, spatial verisimilitude is shown to offer access to compositional attributes that are very different to those of more common spatialisation techniques.

The remaining chapters each deal with the three principal dimensions, iterated earlier, that underlie the research question. Chapter 4 seeks to understand the relationship between music and representations of space. Specifically, it attempts to outline how space comes to be perceived in music. To this end the ideas of logician Charles Peirce are adopted. The application of Peirce’s theory of signs exposes the different ways that space and spatial attributes can be referenced in music. This chapter highlights an important observation that is central to this thesis: references of space exist in music outside of the conscious compositional act of spatialisation.

Chapter 5 considers the creation of illusion in sound. It begins by exploring some important parallels between the notion of fidelity in sound recording and realistic illusion. Arguments about how reality is interpreted and perceived are presented. Key insights proposed from research in the use of sound to create mediated environments and virtual reality environments are described.

Chapter 6 considers the compositional impact of the technological dimension of the pursuit of spatial verisimilitude. Taking the lead from composer and researcher Agostino Di Scipio, the work of 20th century philosopher Martin Heidegger is referenced and a philosophical view on the relationship between art and technology is explored. The perspectives of some other important commentators in this field, such as Theodor Adorno, are referenced.
The concluding chapter, Chapter 7, gathers insights from the various approaches explored. Some of the key arguments presented by this thesis arise from the union of these disparate approaches, and therefore only crystallise in the conclusion.

The portfolio of works was developed prior to the extensive research effort. It is therefore presented first to respect its chronological relation to the remainder of the thesis. The remaining chapters, however, are presented in a logical, rather than a chronological one. The first chapter presented after the portfolio of works, Chapter 3, seeks to understand the relation between spatialisation technique and composition. Its aim is to shed light on how composing with spatial verisimilitude might differ from composing with other techniques.
2 PORTFOLIO OF WORKS

In this chapter I focus on the compositional engagement with spatial verisimilitude. Compositional engagement is demonstrated by a portfolio that includes both creative work and the creation of software for the realisation of that creative work. Whilst some of the creative work provides empirical support for the conclusions of the thesis, the portfolio mostly serves to catalyse challenges faced by the spatial composer. The portfolio is generative of the thesis, not a product of it. Its significance lies in how the realisation of the works has informed the direction of the research.

My approach to composition is first outlined in order to isolate any generalities concerning the pursuit of spatial verisimilitude from my compositional concerns. Indeed, the production of the portfolio has involved a broad range of compositional concerns including visual representations of musical abstractions, notions of gesture, approaches to sound generation and the mapping of non-musical parameters to musical ones. Many of these concerns pertain to the design and development of a custom compositional environment named ImageSynth. These concerns will first be contextualised within critical discussions relevant to the role of computers in music composition. This leaves the remainder of the chapter to focus on the identification of concerns specific to the research question.

The portfolio of works includes a collection of isolated spatial-musical gestures and several spatial compositions. The software tools developed are described in some detail since they encapsulate a specific compositional understanding of space. Observations drawn from the production of each work are then discussed.

The accompanying DVD-ROM offers access to all compositions as well as the code behind the software developed. The compositions are made easily accessible by way of the index.html file on the root folder of the DVD-ROM. All works, except for the first, are each available in their original ambisonic format, as well as Dolby 5.1 files (encoded as MP4s) and as stereo renders. The stereo renders are not binaural
renderings: that is, they are optimised for listening on loudspeakers more than headphones. These stereo renderings were created by deriving audio signals representative of two virtual microphones placed within the ambisonic sound field. The position of these two virtual microphones corresponds to a pair of stereo speakers with a sixty-degree separation. The stereo pair maintains the same forward orientation as the original ambisonic works. The only variable in this process concerns the directivity pattern of the virtual microphone, which can be configured as omnidirectional, cardiod, figure-8 or anywhere in between. Slightly different directivity patterns were chosen for each work but none strayed far from a cardiod pattern.

It is worth noting that ambisonic encoding implicitly includes the definition of the forward listening position. The forward orientation was preserved in the stereo decodes. When the works are listened to on a speaker array of four or more speakers the listener may rotate their head and explore different listening orientations. The works were composed to cater for voluntary head movements, and this aspect of the experience of the spatial sound field is notably lost in the stereo versions.

The localisation resolution of the original spatial audio files is optimally experienced when decoded over an eight equally spaced horizontal speaker array. That said, a 5.1 system setup with 4 equally spaced speakers should give a reasonable impression of the spatiality composed in the pieces.

2.1 Broad contextualisation of my compositional process

Several different compositional approaches were explored. These are, in part, related to my past compositional practice, in part informed by research into the working processes of composers I hold in high regard, and in part influenced by the characteristics of the compositional environment developed.

My past work has mostly adopted a compositional approach that Emmerson calls “empirical and critical” (1986, p.22). This is an approach that has its roots in the work of Pierre Schaeffer when at the *Groupe de Recherches Musicales (GRM)* in Paris. In his text *Traité des Objets Musicaux* (1966), Schaeffer “established rules for the combination of sounds, abstracted from an analysis of their perceived properties”
(Emmerson 1986, p.21). In other words, this compositional process involves the critical appraisal of existing sounds, and their subsequent empirical ordering in time. The composer begins with a set of chosen sounds and structures them into a composition. Emmerson (1986, p.23) describes this approach as the development of a syntax that is *abstracted from* the existent musical material. This stands in opposition to the approach of composers of the same era such as Boulez and Stockhausen who define *abstract syntaxes* prior to their expression through sound material (1986, p.23).

The first and primary compositional approach engaged in the creation of the portfolio differs from the work of early electroacoustic composers at the GRM, in that the sounds being used are generated through an independent process, they do not begin as recordings. Musical-spatial gestures are first designed and rendered by the software developed. These gestures are collected, appraised, and then ordered in time. It should be noted that this approach differs to contemporary approaches encapsulated by modern Digital Audio Workstations (DAW), in that the musical material is *not edited* while it is being organised into a composition. Due to the fragility of the illusion of space in sound, any additional processing that might occur in a DAW was found to have the capacity to erode spatial verisimilitude. Of course, this is one of the insights drawn from the production of the portfolio of works; for example, it is found that the adjustment of the relative volumes of different pre-rendered spatial gestures diminishes spatial verisimilitude. Such observations are identified in this chapter and explored in detail in later chapters.

The distinction between a sound generation phase, and a sound organisation phase is reflected in the compositional process of British electroacoustic composer Trevor Wishart (interviewed in Vassilandonakis 2009, pp.12-13). Wishart, however, describes another initial phase that concerns the creation of “some general overarching idea” (p.12). In so saying, Wishart endorses Emmerson’s concept of *abstract syntax*, since the ordering of the collected sounds is not entirely reflective of their inherent musical value, but owes something to a higher extra-musical idea. Wishart states that it is the genesis of this overarching idea that is “often the hardest part for me” (p.12). The conception of a musical idea that acts as a guide to the
structural manifestation of a composition is something that I also find difficult. This has led me to research discussions on the compositional processes of other composers I hold in high regard. Two such composers, whose approaches have influenced my work, are Iannis Xenakis and György Ligeti.

Ligeti does not talk of a distinction between a sound generation phase and a sound-ordering phase, but rather between what he calls naïve musical ideas, and structural order. For Ligeti, “Composition consists principally of injecting a system of links into naïve musical ideas” (1983, p.124). He clarifies, however, that the “structural potentialities are already contained in the primitive idea, and the act of composition consists mainly of developing these latent potentialities” (1983, pp.124-125). Ligeti’s notion that structural orders are already suggested within isolated musical ideas is one that I have attempted to embrace. As is discussed shortly, the majority of the musical ideas generated within this portfolio of works exist as spatial gestures. Several of these spatial gestures are here presented without any accompanying compositional structure, but reflection on how to develop a compositional structure emergent of the spatial gestures themselves has occupied much of my compositional thinking. Ligeti states: “I imagine the music in the form in which it will later be heard, and hear the piece from beginning to end in my inner ear” (1983, p.124). I have found the conceptualisation of the entirety of a spatial-verisimilitude compositional work, within my ‘inner ear’ very difficult. Of course, a high level conceptualisation of a work is somewhat at odds with my original ‘empirical and critical’ approach to composition, but arguments developed in further chapters indicate that this difficulty owes something to the tense relationship between spatial verisimilitude and abstract musical form. In the quest to find ways to contextualise independent spatial gestures into musical structural orders I eventually adopted a different approach informed by Xenakis’ compositional methods.

Xenakis calls for a “new type of musician”, that of “the ‘artist-conceptor’ of new abstract and free forms” (Xenakis 1985, p.3). Where his approach differs from the design of abstract syntax, as defined by Emmerson and introduced above, is that Xenakis does not seek to invent musical abstractions, rather he seeks to give musical expression to abstractions generally. He argues the need for a “new science of
‘general morphology’ informed by a breadth of fields such as “mathematics, logic, physics, chemistry, biology, genetics, palaeontology, the human sciences and history” (p.3). As is evidenced by Xenakis’ multi-media productions such as *Polytope de Cluny* (1972) these abstractions can also be expressed outside of a musical concern. Within the context of this perspective, the question I asked myself is this: what general abstractions can be found within the spatial organisation of sounding objects? The last composition presented in this chapter is the result of this mode of thinking.

It is necessary to clarify that many of the musical ideas developed are the result of processes that include an element of uncertainty. Some of the spatial gestures created do not originate as a *musical idea* but are empirically appraised as being *valid* as a *musical idea* once they are auditioned. The design of the compositional environment developed has, at least initially, favoured such an approach and this is detailed shortly. Of course, the elements of uncertainty and randomness within this approach have merely served to generate material possibilities that are then brutally culled. Like Xenakis, my use of randomness has no direct relation to compositional legacies such as that of John Cage, in which chance and indeterminacy feature as important compositional elements (Charles 1965).

The majority of the ‘naïve’ musical ideas created in this portfolio exist as spatial gestures. The notion of gesture, as it predates electronic technology, is historically tied to the physical means of production of sounds: “For many centuries, people learned to listen to sounds that had a strict relation to the bodies that produced them” (Iazzetta 2000, p.259). The introduction of electronic technology introduces a dissociation of sound and its means of production which forces a reinterpretation of the significance of *gesture* within a musical context. The arguments surrounding meaning in musical gestures are well summarised by Iazzetta (2000), Cadoz (1988) and Cadoz & Wanderley (2000) and are not discussed here. What is of significance is that the notion of gesture, whether related to or independent of a physical performer, is intimately tied to the notion of movement in space. In this sense gesture is a musical concern that will necessarily feature prominently within the compositional exploration of spatial music. This is confirmed by the portfolio of works in which the
initial musical ideas have manifested not as melodic, harmonic or rhythmic elements, but primarily as movements in space.

Iazzetta says that gesture “does not mean only movement, but a movement that can express something” (2000, p.260). This thesis is only partly concerned with how spatial movement might express something, and more focused on how the realistic illusion of spatial movement might express something. As such, what is of concern is not the compositional use of spatial gesture, but rather the compositional use of *illusion* as spatial gesture. Here, some difficult issues are raised. Chabot suggests that it is precisely the absence of a physical performer "that is naturally compensated for in music by reintroducing a concern for body, gesture and space" (1990, p.15). The use of musical gestures within non-spatial electroacoustic music can thus be understood as a way to *re-introduce impressions of space*. In this sense, spatial verisimilitude can be seen as *competing* with musical gestures since both have an interest in space, but manifest that space in very different ways. Indeed, the exploration of how musical gestures or constructs can result in the (non-realistic) representation of space, which is documented in Chapter 4, contributes key arguments to the research question.

Whilst the spatial projection of sound caters for spatiality previously inherent in physically performed gestures, it does not cater for the existence of the performer’s body. The spatial movement is thus disembodied from a performer. However, in pursuing spatial verisimilitude, a new connection is established with a different body. Spatial verisimilitude results in the creation of a sense of presence and immersion: qualities that are intimately tied with the body of the *listener*. In this sense, the realistic projection of sounds in space can be understood as a transferral of the gesture’s centre from the body of the performer, to that of the listener. This transferral resonates with arguments developed in Chapter 5: the notion of fidelity in sound originally consisted of convincing the listener that the performer was present within their company; this developed into a concern with convincing the listener that they are present in the performer’s company. In this later concern, the illusion operates on the listener’s body. The notion of gesture thus holds central importance.
in the development of the portfolio of works, and the issues it highlights feature prominently in subsequent chapters.

The centrality of body movements, within notions of musical gesture, have led to much research exploring the real-time mapping of physical movements to synthesised sound (Winkler 1995; Wanderley 2001; Wanderley & Depalle 2004). This thesis has no concern with real-time interaction, but how spatial movements are translated to sound is of central importance to the compositional environment developed.

*ImageSynth*, developed in the audio signal processing environment known as SuperCollider (McCartney 1996), translates images, and sequences of images, into spatial audio. It maps both movement and images into sound. As such, the realisation of the portfolio of work involves many issues relating to what is known, within the critical literature, as *mapping*. Hunt (2000, p.1) defines *mapping* as “the liaison or correspondence between control parameters (derived from performer actions) and sound synthesis parameters”. Within the context of *ImageSynth*, which serves primarily as a compositional environment, there is no specific ‘performer’ concerned. It is the composer who has complete jurisdiction over the ‘control parameters’ which, it should be noted, includes the possibility of designed randomness. Hunt (2000, p.1) describes two views on the role of mapping: the first sees mapping as an “integral part of the instrument” (p.1), and the second sees it as a “specific feature of a composition” (p.1). For the purposes of *ImageSynth*, mapping is best described as the second; a feature of the composition. Each composition, or spatial gesture, presented in this portfolio of works is accompanied by its own specifically designed mapping strategy.

A detailed description of the mapping strategies explored, within the production of this portfolio, will not be provided because the breadth of issues concerned lie outside of the immediate concerns of the research question. Some of these issues are detailed in a paper presented at the 2010 *New Interfaces for Musical Expression* (NIME) conference (Deleflie & Schiemer 2010). Generally speaking, as with many mapping strategies developed for compositional purposes, the mapping strategies
developed are the result of “a process of exploration” (Doornbusch 2002, p.155). In line with the primary concerns of this thesis, this exploration of different mapping strategies is guided not just by compositional concerns, but also by the quality of resultant spatial verisimilitude. Where specific insight into the research question can be drawn from the mapping strategies used, a detailed account is provided.

As a compositional environment, ImageSynth can also be considered within the critical enquiry surrounding the use of computers within music. Lejaren Hiller, a pioneer in the field of computer music, states that there are two different interpretations for this term:

The first, computer-composed music, involves composition, that is, note selection. The second, computer-realized music, involves conversion into electronic sound of a score that may or may not have been composed with the aid of a computer. (1981, p.7)

ImageSynth is concerned with both interpretations. Through the manipulation of sequences of images, and their content, ImageSynth allows musical abstractions to be designed and expressed independently of how those abstractions are eventually translated to sound. Of course, to expand on Hiller’s definition, those musical abstractions may involve musical parameters other than notes. Assayag et al. prefer the term Computer Assisted Composition (CAC) which they define as systems that “focus on the formal structure of music” (1999, para. 1). ImageSynth caters for the expression of formal structures of music through the careful design of images, and their sequences. Through the use of mapping strategies, ImageSynth also caters for the synthesis of sound from those formal structures.

The formalisation of compositional structures and their adaptation to computers can be traced back to work by Hiller and Issaacsion in a text named Experimental Music; Composition with an Electronic Computer (1959), and has been explored by 20th century composers such as Xenakis, and Gottfried Michael Koenig. Xenakis documented some of this work in his text Formalised Music (1992), first published in 1971. He eventually expanded on this work by designing a system that incorporated a graphical user interface: UPIC, which stands for Unité Polyagogique
Informatique du CEMAMu⁴, is “a machine dedicated to the interactive composition of musical score” (Xenakis 1992, p.329). Examples of more contemporary efforts that focus on the abstraction of musical structures, and also include a graphical user interface, include PatchWork and OpenMusic developed at IRCAM (Assayag et al. 1999). Of course, since the introduction of high-level digital signal programming environments such as MaxMSP (Puckette 1991), PureData (Puckette 1996), and SuperCollider, different approaches to composing music on computers have proliferated.

The UPIC system, first created in 1977, was initially conceived by Xenakis to “allow the composer to draw music”, thus freeing the composer from the constraints of traditional music notation (Marino, Serra & Raczinski 1993, p.260). One of its key characteristics, however, was the idea that the “composer could create autonomously” (p.260). Sound synthesis is therefore also an important feature of the UPIC system, allowing the composer to create completed works without engaging instrumentalists. Later versions, since 1991, are capable of real-time sound synthesis (p.260). UPIC thus holds significance as an early example of both computer-assisted composition and computer-realised music. Thiebaut et al. argue, however, that due to its inherent constraints UPIC is largely used as a sound design tool for the creation of original timbres (Thiebaut et al. 2008, p.2). In this sense, it is perhaps best compared to other modern sound design environments such as Metasynth, AudioSculpt and Sonos. This balance, between a focus on structural concerns and a focus on sound synthesis concerns, is one that has been central to the use and development of ImageSynth. Some of the works presented here are the result of using ImageSynth primarily as a sound design tool; others have used it primarily as a compositional tool.

ImageSynth also encapsulates an important third element not represented in the preceding discussion, which involves the control of the spatialisation of sound. How ImageSynth incorporates spatial design into both compositional concerns and the synthesis of sound is of significance to critical enquiry concerned with the

---

⁴ CEMAMu is the centre for studies in mathematics and automation of music, founded by Xenakis in 1965 (Xenakis, 1992, p. 329)
integration of spatial design within compositional tools. A broad overview of such efforts is provided by Schumacher & Bresson in “Spatial Sound Synthesis in Computer-Aided Composition” (2010). Many different approaches and software projects are reviewed, such as IRCAM’s Spat (Ircam Spat 2012), which attempts to expose spatial parameters in perceptual terms (Schumacher & Bresson 2010, p.271). Schumacher & Bresson state that most tools separate the spatialisation of sound from compositional concerns. They go on to propose:

[A] mode in which spatial sound rendering is regarded as a subset of sound synthesis, and spatial parameters are treated as abstract musical materials within a global compositional framework. (2010, p.271)

ImageSynth takes a similar approach but goes one step further. Spatialisation is not considered a subset of sound synthesis, but rather an inseparable and intrinsic part of it. In ImageSynth, it is not possible to audition synthesised sounds outside of their spatialisation. Similarly, it is not possible to create a high level musical structure without defining spatial attributes such as position. These characteristics, detailed shortly, are a function of the design of ImageSynth, which uses images to represent spatial sound maps. They make ImageSynth challenging to use, but ensure that spatial concerns remain central.

Another way in which ImageSynth differs to most of the projects described by Schumacher & Bresson is that it is a tool that is designed and developed in accordance with compositional ideas. Here, ImageSynth resonates somewhat with Xenakis’ UPIC, originally designed to support a particular approach to composition. Where ImageSynth differs from UPIC is that its development was an intrinsic part of the compositional process. In this sense, the work presented here also belongs to critical discussion concerned with the relationship between composition and the development of compositional tools (Hamman 2000, 2002; Manning 2006). One of the issues presented in these discussions concerns precisely the reduced involvement of composers in the design of compositional tools, Hamman states:
When experimental compositional concerns are not taken into account at this [design] stage, the resulting tools tend to reflect normative interpretations of compositional technique and musical dramaturgy. (2002, sec. 2)

It is for this reason that ImageSynth has been designed and developed from the ground up: by defining all aspects of the compositional environment, the exploration of the relationship between composition and the pursuit of verisimilitude largely escapes the “huge ideological and epistemological payloads” (Hamman 2002, sec. 2) of existing compositional tools. It should be noted, however, that the development of one’s own technology does not negate the technological mediation of the compositional act. This notion is explored in detail in Chapter 6, which begins with an exploration of the significance of Stockhausen’s design of a hardware device used within his composition Kontakte (1958-1960).

2.2 ImageSynth: a spatial compositional environment

ImageSynth abstracts spatial compositional structures by using images as spatial sound maps. It consists of three parts. Firstly, sounds are spatialised by a spatial synthesis library developed prior to the design of ImageSynth. This library pursues spatial verisimilitude by mimicking spatial acoustic reality for single point sources of sound. The spatial characteristics simulated, and the perceptual cues targeted, are detailed shortly. Secondly, spatial sound is abstracted as sequences of images. An extensive user interface exposes control over those images, which can be captured, edited or authored using scripts executed within ImageSynth’s user interface. Thirdly, sound synthesis is performed by way of a mapping layer that allows defining how images are converted to sound.

SuperCollider, the audio-programming environment used, has a built-in feature that is very amenable to spatialisation processing. Multichannel Expansion enables the ready expansion of audio processes to include multiple channels of audio without requiring extra lines of code to be written. SuperCollider is therefore well suited to the development of multi-channel spatialisation tools. It also offers a broad range of user-interface programming constructs, including a basic set of image processing capabilities. Lastly, SuperCollider offers good support for offline-rendering. This
enables rendering spatial scenes that include many sources of sounds, each treated using a variety of sound processing techniques.

Sounds are spatialised using ambisonic techniques. Ambisonics allows for a range of spatial resolutions to be used. That is to say, by increasing the number of audio channels used, and correspondingly the number of loudspeakers used, ambisonics can cater for different levels of localisation accuracy. These different levels of resolution are referred to as ‘orders’; a term borrowed from the mathematical foundation that underlies ambisonic theory: spherical harmonics. The spatial synthesis library developed uses 3rd order ambisonics, but the spatial information is limited to the horizontal plane. That is, the spatial sounds created have no height component. 3rd order horizontal-only ambisonic signals amount to seven channels of audio, and are well suited for playback over eight equidistant and horizontally spaced loudspeakers. This level of resolution was chosen to match the use of eight-channel speaker systems popular amongst spatial music composers (Otondo 2008, p.79).

The behaviour of sound in space is complex and subject to many environmental factors. Its simulation thus involves a trade-off between mimetic accuracy and available computational processing power. Choices need to be made about which spatial sound phenomena will be simulated. For the purposes of the spatial synthesis library developed, these decisions were informed by a range of literature (Chowning 1977; Blauert 1997; Malham 1998; Begault 2000; Malham 2001a; Anderson & Costello 2009). After the completion of the portfolio, further research led to a re-thinking of which spatial sound phenomena might best be simulated. This research is elaborated in Chapter 5.

The following phenomena are simulated in the spatial synthesis library:

- **Direction.** Encoded using 3rd order horizontal-only ambisonic encoding.

- **Change in sound intensity as a function of distance.** This partly simulates the effect of distance on sound.
• The loss of high frequencies due to humidity in the air. Implemented using a coarsely configured low pass filter. This partly simulates the effect of distance through high frequency absorption.

• Temporal distortion (delay in time) due to time taken for sounds to reach the listener. This partly simulates the effect of distance when multiple sounds are projected together.

• Early reflections of sound on walls: first, second and third reflections on up to 6 surfaces, including floor and ceiling. This supports the perception of both distance and direction.

• Diffuse reverberation. Implemented in 3rd order horizontal-only ambisonic encoding. Two different implementations were used to attempt to blur the spatial information encoded in the reverberant field.

• Correct proportion of direct to reverberant sound was respected. This helps support the perception of distance.

• Doppler effect simulated on moving sounds. This supports the perception of movement.

The software was run on an Apple Macbook Pro, whose processing capacity catered for the real-time simulation of around twenty point sources of sound. Any more point sources of sound begin to saturate the computer’s resources causing it to produce spatially imprecise sound artefacts. However, non-real time rendering allows for an arbitrary amount of point sources to be used.

The ImageSynth interface, of which the main view is shown in Figure 1, includes: the image currently being worked on; controls to manage the mapping layer responsible for synthesising sounds; one script window that allows editing the image; another script window that allows defining arbitrary parameters used in the rendering of the spatial scene and; an ambisonic player that allows for the auditioning of scenes recently rendered.
The large image in the top left hand corner is the image currently being rendered into spatial audio. The image acts as a map, laid flat, in which the listener stands in the centre, as indicated by the just-visible cross hairs. Each pixel acts as an individual point source of sound. Figures 2 and 3 illustrate this arrangement more clearly. In Figure 3, each pixel is expressed as a ‘virtual’ loudspeaker. In the spatial rendering, the pixels are allocated a virtual-separation such as ten metres. The separation of ten metres is chosen as the default because it allows a clear audible distinction between individual sounds that are very close to the listener, say one metre, and neighbouring sounds which are then necessarily close to ten metres away. At a separation of ten metres, the pixels in the corner of the image are around 424m away from the virtual-listener. Since the spatial synthesis library models the effects of distance on sound, the sounds generated by the pixels in the corner of the image will be heard over one second after the pixels immediately adjacent to the centre of the image. In the
screenshot shown in Figure 1, the image is made up of 3600 pixels thus representing 3600 point-sources of sound.

Figure 2. Each image acts as a map laid flat, with the listener positioned in the middle.

Figure 3. Each pixel can be thought of as a virtual loudspeaker.

Each pixel consists of numerical information that defines its colour. It is this information that is mapped to sound synthesis processes. Attributes such as colour, brightness, hue, and difference from adjacent pixels, are all accessible within the mapping layer. The sound synthesis techniques employed in this layer include: granular synthesis, a technique pioneered by Iannis Xenakis after seeing the value of Gabor’s research (Roads 2004, pp.65-66); stochastic synthesis, another technique pioneered by Xenakis which aims to create perceptually complex tones (Xenakis 1992, pp.289-294); and a range of additive synthesis techniques that include sine, saw and square oscillators.

The original concept for this interface originates from another spatialisation technique I developed which involves using granular synthesis techniques for creating sounding objects that have a physical dimension (Deleflie & Schiemer 2009). The focus on the dimensional aspect of sounding objects is a conscious attempt to avoid one of the common “ideological and epistemological payloads”

---

\[5\] In the 1940s British physicist Dennis Gabor explored the decomposition of sound into elementary sonic grains (Roads, 2004, p.57)
(Hamman 2002, section 2) of spatial compositional tools, that is; the encapsulation of sound in space as consisting only of point sources of sound. Apparent sound source size can be created by modelling many point-sources of sound, treated in such a way as to be perceived as a singular source of sound that has dimensional attributes (Potard & Burnett 2003, 2004). The technical term for this treatment is *decorrelation*, explained by Kendall as:

\[ \text{[A] process whereby an audio source signal is transformed into multiple output signals with waveforms that appear different from each other, but which sound the same as the source (1995, p.71)} \]

The original technique mentioned above involves the decorrelation of already-spatialised sound files. In this adaptation, pixels represent point sources of sound that, if decorrelated from their neighbours, act together to create the perception of size. In other words, groups of similar pixels are spatialised as singular sound sources with dimension. There are a variety of ways to implement decorrelation, but the technique employed here is a by-product of one of the sound synthesis techniques used, that is, granular synthesis (Rolfe & Keller 2000). As such, the attempt to create apparent sound source size depends on granular synthesis being used as the sound synthesis technique.

Spatial synthesis and ambisonic encoding is performed by the spatial synthesis library already discussed. Ambisonic *decoding*, which involves converting an ambisonic signal into audio signals destined for a pre-determined array of loudspeakers, is executed by the open-source software called AmbDec (Adriaensen 2009). AmbDec was used to decode ambisonic signals to both an eight speaker array and to stereo for audition over headphones.

ImageSynth incorporates an audio file player, labelled “Ambisonic Player” as illustrated in Figure 5, so that spatial audio renders can be quickly and easily auditioned after rendering. This audio file player simply streams the contents of the audio file to AmbDec via Jack (Davis & Hohn 2003). Jack is a background server application that facilitates streaming audio from one application to another.
As mentioned earlier, the spatial synthesis library is only capable of rendering around 20 point sources of sound in real time. The spatial rendering of 3600 point sources of sound must thus be done in non-real time (NRT). ImageSynth is capable of both real-time and NRT rendering. Real-time rendering is limited to the 20 point sources closest to the central listener and is used for a number of purposes including: debugging code, test-auditioning pixel maps and auditioning compositional structures.

To enable NRT rendering, ImageSynth is designed such that a text score is first written to disk. This text score contains messages in the Open Sound Control (OSC) (Wright, M 2005) syntax that are then rendered into sound offline. SuperCollider’s internal architecture is very amenable to this way of working. SuperCollider’s language application is logically separated from the audio synthesis engine, which receives commands from the former. These commands employ subset of the OSC protocol and, as such, ImageSynth’s strategy of separating score from the eventual NRT rendering is relatively easy to achieve and manage.

Figure 4. File player allows auditioning completed spatialisation renders by streaming audio channels to AmbDec.
When using all pixels in a 60 x 60 pixel image, rendering several seconds of audio may take several hours. There is a significant amount of processing involved in calculating such things as sound’s reflections on walls.

As can be seen in Figure 5 a number of other parameters are exposed when using NRT: rendering is limited to the images numbered between the ‘Start’ value and the ‘go to’ value; the ‘duration’ field refers to the length of time, in milliseconds, that each image will be rendered for; the ‘volume’ field allows adjusting the output volume to avoid clipping in the resultant sound file. Rendering a limited range of images allows quick auditioning of limited parts of the composition or gesture. Changing the duration of each frame affects the temporal dimension of the rendered spatial sounds. Depending on the sound synthesis technique used, this can be more or less dramatic. For example, if the sounds are generated by granular synthesis techniques designed to emphasise timbre, then changing the duration of each frame can result in the perception of entirely different timbres.

Figure 6 shows how different mapping strategies can be chosen for the rendering of the current composition. Each strategy is encapsulated in its own file and defines exactly how the information contained within the image is translated into sound. Any
new mapping strategies defined can easily and quickly be incorporated into the compositional environment.

As expressed by Hunt, Wanderley and Paradis (2003, p.429), in discussing the importance of mapping layers within electronic instrument design: “by altering the mapping, even keeping the interface and sound source constant, the entire character of the instrument is changed”. Here, we are concerned with a compositional environment, not an instrument that is performed, but the observation is similar. By changing the mapping strategy employed to translate images into sounds, the resultant character of the sounds can change dramatically. More importantly, and of specific relevance to the concerns of this thesis, modification of the mapping layer can have a dramatic effect on the perception of space. This is of particular concern since the mapping layer has no control on the spatial synthesis of sounds. In other words, by merely changing the sounds that are spatialised, the perception of space is affected despite the fact that the spatialisation remains identical. The significance of this observation is explored in Chapter 4. Specific examples of how perceptions of space are affected by the sounds being spatialised are discussed shortly, within the context of the specific completed works.

One of the benefits of using images to create spatial design is that images can be sourced in a variety of ways. Still cameras offer a high level of control and video

Figure 6. Different pixel maps can easily be chosen within the interface.
cameras can source sequences of images that already have a temporal dimension. The implications of using sequences of images captured from a video camera are discussed shortly, within the context of the first spatial gesture rendered. Another way to create images is by manually authoring them using scripts. This allows a pixel-level control that, for ImageSynth, translates into a sound-source level control. To facilitate this method of working, a dedicated script window allows writing scripts that have complete control over the contents of images. Figure 7 shows that pre-written scripts are easily accessible via a drop down list, or they may be manually created directly within the ImageSynth user interface.

The script interface exposes control not just of the current frame of pixels, but also of the pixels in prior and subsequent frames. This allows scripting changes across sequences of images that can be used to model temporal structures. It is here that the careful combination of mapping strategy with temporal structures expressed in the sequence of images can be used to define structures of compositional significance. Of course, whether a temporal structure abstracted as a sequence of images finds expression as a perceivable compositional structure will be a function of both the original temporal structure and how it is mapped to sound. An exploration of how composers use mapping to explore compositional expression is provided by Doornbusch (2002). What is of specific significance here is how spatial temporal structures might translate to compositional expression. In the compositional exploration of the relationship between space and time, it was found that ImageSynth’s initial design of having all sound sources in a static physical location was unnecessarily limiting. During the development of my portfolio of works some new features were introduced that enable moving sound sources both relative to the listener and to each other.

Since the contents of all images can be accessed through scripts algorithmic patterns are easily expressed either across images, or within a singular image. Two of the compositions detailed below use simple statistical algorithms for the temporal distribution of parameters such as sound loudness and pitch.
Different pixel mapping strategies use potentially very different logics for translating pixel information into sound. A system of image filters has been created in order to facilitate visualising how a particular image might be translated into sound. These filters do not alter the mapping, the image data, or the resultant sound in any way. They merely allow a better visualisation of how the mapping strategy will affect the concerned image.

Figure 7. A script-editing window allows editing images using scripts.
Figure 8 shows an image filter that facilitates the visualisation of a mapping strategy that translates one aspect of pixel data to a scale degree. In this figure, red pixels represent the root or first degree, yellow pixels represent the third degree and violet represents the (major) seventh degree.

For easy access to various parts of the score of image sequences, a “show score” button allows displaying images in the sequence as a grid of thumbnails. Any image can be chosen, edited and auditioned.
A comprehensible set of programmed key bindings allows performing actions such as moving frames, copying and pasting frames and deleting frames. Some scores use several hundred images, others use only one. Examples of each are detailed in the following section.

As the exploration of the compositional potential of ImageSynth progressed, some significant limitations became apparent. The fixed position of sounding objects relative to the listener, already discussed, is one such limitation. In order to route around this and other issues, without altering the underlying design of ImageSynth, a set of arbitrary parameters can be defined on a frame-by-frame basis. These parameters can affect both the spatialisation of sounds, and their mapping. They are made accessible via a small text window on the bottom right hand side of the interface, as displayed in Figure 10. In this figure the parameter *frameTime*, for example, allows defining the time that each frame occupies in a composition. Other parameters allow creating coarse amplitude envelopes, rotating the image relative to the listener, and changing the position of pixels relative to each other and the listener.

![Figure 9. Sequences of images are displayed as thumbnails for easy access.](image)

A comprehensible set of programmed key bindings allows performing actions such as moving frames, copying and pasting frames and deleting frames. Some scores use several hundred images, others use only one. Examples of each are detailed in the following section.

As the exploration of the compositional potential of ImageSynth progressed, some significant limitations became apparent. The fixed position of sounding objects relative to the listener, already discussed, is one such limitation. In order to route around this and other issues, without altering the underlying design of ImageSynth, a set of arbitrary parameters can be defined on a frame-by-frame basis. These parameters can affect both the spatialisation of sounds, and their mapping. They are made accessible via a small text window on the bottom right hand side of the interface, as displayed in Figure 10. In this figure the parameter *frameTime*, for example, allows defining the time that each frame occupies in a composition. Other parameters allow creating coarse amplitude envelopes, rotating the image relative to the listener, and changing the position of pixels relative to each other and the listener.

![Figure 9. Sequences of images are displayed as thumbnails for easy access.](image)
2.3 Sound excerpts and compositions

The realisation of each of the works described below has contributed to the enunciation of the research question. The works are presented here in chronological order and can be understood as incremental experiments working towards the reconciliation of the relationship between music and spatial verisimilitude. Together, they raise a broad set of concerns that instigate the thesis research programme. These concerns include, but are not limited to: the impact of a significantly technological pursuit on the compositional process; the introduction of new parameters foreign to compositional thinking; and the apparently awkward relationship between music and realism in spatial representation.

All works are created using ImageSynth. Some of the works are post-produced in a Digital Audio Workstation (DAW). The DAW chosen is Reaper (Cockos 2012) because of its comprehensive support for multi-channel audio files. Its use is limited to re-ordering rendered spatial gestures. The use of Reaper holds its own implications in the pursuit of spatial verisimilitude, and these are detailed where appropriate.

One of the most significant restrictions imposed by the use of ImageSynth is the time required to render spatialisations. This restriction is by no means the only one but it
is mentioned here because it applies to all of the works discussed below. One minute of 3rd order ambisonic spatial sound can take up to one hour to render. A composition of five to ten minutes therefore requires an overnight render. This has an impact on the compositional process in that it practically limits the number of iterations when composing and rendering the work. This limitation favours a compositional approach in which musical structures are conceived outside of the software. This is yet another reason why, as my portfolio progressed, I sought to move away from the ‘empirical and critical’ compositional approach described by Emmerson (1986). The extensive rendering time dramatically reduces the amount of material that can be critically appraised and empirically organised. Of course, the pioneers of computer music experienced similarly significant delays in auditioning the results of computer processes. On the mainframe computers of the 1960s and 70s a minute of sound could take ten or twenty minutes of processing time, and more complex synthesis techniques required upwards of 100 or 200 minutes of processing time (Manning 1993, p. 224). These proportions are of the same order as those experienced when using ImageSynth. In the computer music of the 1960s and 70s, compensation for these extended delays lay in the greater level of control that composers could exercise (p. 224). Conversely ImageSynth allows for large numbers of point sources to be spatialised allowing for greater spatial complexity.

The following discussions are best read in association with listening to the material on the DVD-ROM.

*Torch video* is the first work to be produced by ImageSynth. It consists of the spatialisation of a sequence of images taken from a video of a torch being swung around in the dark. This video was chosen to test the software and establish exactly how spatial movement captured on a video might translate to spatial movement in sound. The video, which includes the sound subsequently generated by ImageSynth, is available on the DVD-ROM.

The video is short and its quality has been reduced to 60 pixels by 60 pixels to match the point-source rendering capacity of ImageSynth. In video terms, this resolution is extremely low. The video produced for the DVDROM is kept at this resolution to
give an indication of the graphical detail that is translated to spatial sounds. Whilst 3600 pixels do not make for a particularly clear or complex video, 3600 point-sources of sound do make for a very complex spatial sound scene. As a point of comparison, IRCAM’s Spat software only allows eight point sources of sound to be spatialised (Ircam Spat 2012). In other words, ImageSynth allows for 450 times more point sources of sound than SPAT. Of course, such a large number of point sources of sound means that ImageSynth’s spatial rendering cannot be done in real time using contemporary computing power.

*Torch video* demonstrates that ImageSynth works as designed, and also highlights a few interesting points concerning the isomorphism between what is seen and what is heard. ImageSynth processes each image as a flat plane, not at all in the same vertical orientation as when viewed. When the torch lifts in the air the sounds are projected forwards. When the torch drops to the ground, the sounds are projected rearward. When the torch is moved left and right, the heard orientation is the same as what is seen. The most isomorphic way to visualise the video would be to project it onto the floor. However, since ImageSynth interprets a distance of ten metres between each pixel, the video would have to be enlarged and projected down to the size of approximately a square kilometre. Of course, ImageSynth is designed as a spatial compositional tool. The images and sequences of images it employs are not designed to be projected alongside the composition. In other words, this lack of isomorphism in scale is not of any great significance.

The expanded scale of the sonic spatial rendering means that as the torch moves away from the centre of the video its loudness drops significantly. This can be seen in the lens flares. If a lens flare extends to anywhere close to the centre of the image it is heard loudly, if the lens flare occurs away from the centre of the image its loudness is much softer. Temporal delays are also modelled in ImageSynth and these also have a significant effect on the perceptual synchronisation between what is seen in the video and what is heard. When the lens flare is far from the centre of the image a time delay occurs before it is heard. As already mentioned, a pixel in the corner of the video will create a sound with a delay of over a second. The last lens flare in the video seems to hardly produce any sound because its distance from the centre of the
image means that its loudness is reduced, and a perceptually significant time delay is effected.

Whilst the video demonstrates that ImageSynth behaves as it has been designed it also highlights that there is only a moderate level of perceptual isomorphism between what is seen and what is heard. Of course, ImageSynth’s intent is to provide a spatial compositional environment in which musical abstractions can be represented as sequences of images; it is not to express videos as sound. Here, a distinction needs to be made between the use of visual data to generate sound and the field known as sonification. Sonification is defined by Kramer et al. as “the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation” (1997, p.4). They present the Geiger counter as a successful example of sonification (p.4). For the purposes of ImageSynth, there is no requirement for the data relations implicit in the images to be perceived in the acoustic signal. A spatial composition may hold musical interest despite the total lack of successful transformation of data relations into sound. Similarly, types of images may be appraised as holding musical interest without the presence of any successful sonification. Lastly, certain characteristics of images might be mapped to sound, to compositional effect, without any revelation of the data relations implicit in them. Of course, what is of concern is not the data relations within the images, but rather how the images are used as abstract musical structures. Hence, as the portfolio progresses, greater emphasis is placed on authoring images using scripts. In this sense, the data relations contained within these images are specifically designed to suit the mapping strategy used. Here, however, the images can be understood as nothing more than a spatio-musical ‘score’.

Whilst sonification is not the aim of ImageSynth, it might still occur. For example, the resultant sound of Torch Video has an organic feel that can be seen as the expression of a hand gesture. In this sense, an aspect of the movement captured in the video is translated into the acoustic signal.

What Torch Video highlights is that if the composer wishes to translate aspects of captured video into compositional structures, then a detailed understanding of the
image-to-sound mapping strategy used is important. This would allow the composer
to predict, or project, how a particular video would ultimately sound once translated
to spatial audio. Within the context of the pursuit of verisimilitude, the resultant
sound captures aspects of realistic movements, such as an organic character, but the
spatiality produced is not particularly realistic: the scale of movements, their speed,
and the abstract identity of the sounds deny a believable sense of reality.

_Fluoresce_ is a spatial gesture created using only two images, shown in Figure 11 and
Figure 12. Both images were randomly generated using scripts written in
SuperCollider. Sound is synthesised exclusively through the manipulation of
granular parameters processing a recording of a fluorescent light being turned on.
The first image sets the starting parameters that are progressively transitioned to
parameters set by the second image. In other words, _Fluoresce_ represents a sonic
transformation of the first image into the second.

![Figure 11. First image for Fluoresce.](image)

![Figure 12. Second image for Fluoresce.](image)

The way in which sound is manipulated by each pixel is very simple. Each pixel acts
as a granular synthesiser where the colour of the pixel determines the position in the
sound recording where grains will be sourced. The first image determines the initial
positions where grains of sound are sourced, and the second image determines the
final positions. The progression takes around 40 seconds.

This logic equates to playing the same audio file 3600 times simultaneously but with
different start and end points, and each playback is spatialised in a different, but
fixed, position surrounding the listener. Pixels of similar colour will start playing the
file in the same place, but end in different places according to their corresponding
colour in the second image.

What results is a broad range of sounds that have a significant spatial dimension, and
outline a number of spatial gestures. Sounds are heard in all horizontal directions and
several different sounding objects appear to be identifiable despite only one sounding
object being involved in the original recording. Different spatial movements can also
be perceived. Approximately two-thirds of the way through the rendering (25sec –
30sec) a large reverberant space can be heard, somewhat like an underground cave
with droplets of water falling into a pool at the bottom.

Whilst *Fluoresce* demonstrates that, with only two small images and one source file,
ImageSynth can generate sonically complex and spatially rich material, it also
highlights some concerns for the pursuit of spatial verisimilitude. Firstly it is very
difficult; if at all possible, to pre-
determine what any combination of image and
source-sound file will sound like. Few of the spatial characteristics of the resultant
render were designed. A reverberant cave-like space was certainly not intended. In
other words, there are some spatial elements whose perception can only be the result
of incidental artefacts, in the spatial rendering of granular textures, and not a result of
the present mimetic simulation. This fact in itself seems to throw the efficacy of
mimetic simulation of space to the mercy of the complexity of the sounds that are
being spatialised. This theme, of significant concern to the spatial composer, is
explored in Chapter 5.

Secondly, *Fluoresce* creates the impression of movement in height whilst in effect
there is no height modelled at all. These impressions are synchronous with changes
in pitch that are caused by how pixel information has been mapped to granular
synthesis processes. This further places the simulation of acoustic spatial behaviour
to the mercy of the sounds being spatialised. The interpretation of spatial information
from common musical constructs, such as changes in pitch, is also of concern to
composers, and is discussed in Chapter 4.
Fluoresce thus demonstrates that both sonic complexity and pitch can cause the perception of spatial characteristics not simulated let alone designed\(^6\). The pursuit of spatial verisimilitude thus turned my efforts towards the spatialisation of simple recognisable sounds. By using recognisable sounds, unintended spatial perceptions can be minimise, thus allowing greater control and design over the spatial characteristics present within the composition.

Drivetrain aims to reduce the complexity of sonic information in order to regain control over the spatial verisimilitude. By avoiding abstract sounds, it is hoped that incidental spatial characteristics can be avoided. Drivetrain uses a recording of a drummer playing a simple rhythm. The choice of a drummer is motivated by practicality: the recording can easily be looped such that its start and end points are not perceptible. The recording is played at normal speed. Granular synthesis is still being used to play the loop, but the granular parameters have been tightly controlled so as to allow some decorrelation without losing the clear recognisability of the sound source. No videos, captured images or randomly generated images were used to produce this work. Instead, images were authored pixel-by-pixel. The sequence of images designed models a line of around 200 drummers, performing the rhythm, slowly advancing towards the listener. Figure 13 shows three images, each several seconds apart in the sequence. The light pixels indicate that a new loop is being triggered, and the dark pixels indicate a loop is currently playing.

![Figure 13. Part of the sequence of images used to create Drivetrain.](image)

\(^6\) The capacity for perceptions of space to arise from timbral changes is discussed in Chapters 3 and 4. The association between pitch and perceptions of height is discussed in Chapter 4.
*Drivetrain* successfully regains control over the spatial verisimilitude and highlights that the exploration of distance can be used to compositional effect. When the sounds are simulated to be far away (approximately 300m) their source is not identifiable, but as they move closer to the listener, the spectral effect of distance incrementally reduces until the sound sources can finally be recognised as a drum rhythm. This procedural movement, from far away where the sounding objects are obscure; to near where they are finally recognised represents a temporal gesture with compositional potential. Here, a compositional structure based on tension and release is created by the transition from abstract to recognised sounding object. This transition is exclusively created as a result of spatial circumstances.

*Drivetrain* also demonstrates that there is a spatial relationship between abstract and recognised sounding objects. Abstract sounds can thus be consistent within a realistic spatial scene because they can represent recognisable sounds obscured by spatial circumstance. Here, however, it is the composer who must manage spatial circumstances to support that an abstract sound is, in fact, just an obscured recognisable one. The difficult relationship between perceptions of distance and abstract sounds is known to composers. Stockhausen (1989, pp.106-107), for example, explains that to use abstract sounding objects in spatial composition, the composer must first expose the listener to how those objects sound at distances both far and near. What is described here is different. *Drivetrain* shows that within a realistic spatial setting, sounds may suggest spatial obfuscation, such as distance, by virtue of their unrecognizability. Here, abstract sounds can represent spatially obfuscated recognisable sounding objects. In other words, the transitioning between a recognisable sound and an abstract sound can represent spatial movements.

Based on the objective of pursuing spatial verisimilitude, *Drivetrain* caused a re-evaluation of the ideas behind ImageSynth: if each pixel, and hence sound source, must be explicitly defined and scored to achieve any level of control over spatial verisimilitude, then sequences of *captured* images do not make good abstractions for spatial sound.
*Lost Corps* takes the direction of *Drivetrain* and continues further. In the interest of optimising control over spatial verisimilitude, *Lost Corps* also uses a readily recognisable sound but chooses one in which pitch is controllable. All sounds in *Lost Corps* are sourced from a recording of a fluegelhorn playing a chromatic scale over several octaves. Fluegelhorn sounds are positioned in the score image in the same way as the drum loop was in the previous work, but now access to pitch is also exposed through the pixel colour.

![Image](image_url)

**Figure 14.** Part of the sequence of images used to create *Lost Corps*.

Figure 14 shows some of the images used to produce *Lost Corps*. Fluegelhorn sounds are grouped into sections of rows, columns and square clumps that move through the video, towards and away from the centrally positioned listener. Here, simple statistical algorithms were used to evenly spread sounds over rectangular sections of image sequences.

Choosing a musical instrument, as the recognisable source sound, also means that changes in pitch can be used without creating accidental representations of height. Since an instrumentalist is understood to be typically at or near ground level, the sound of high pitches performed on a fluegelhorn can escape the perception of height. Arguments that I present in Chapter 3 indicate that perceptions of space are heavily contingent upon knowledge of the spatial circumstance of the identified sounding objects. *Lost Corps* confirms this. However, it also highlights another complex issue. By associating high pitches to a physically performed musical instrument, the perception of height is only inhibited at the level of the acoustic modelling. Within the context of a musical composition, as is elaborated in Chapter 4, pitch can still be used to create musical representations of height. If the listener
focuses on the realistic spatial representation, then the perception of the instrumentalist’s sounds will remain tied to the ground, but if the listener focuses on changes in pitch with respect to the context of the musical composition, then height may well be perceived.

In other words, in this case there are two levels of spatial representation; one is musical, the other is simulated. In this particular context, where the music heard is tied to a musical performer, there is not necessarily any conflict of spatial information. This raises a key question: if music is understood to already have a capacity to create representations of space, then what is brought by the realistic illusion of sounds in space? Here, realistic spatial modelling can be seen as a parallel, rather than a support, to the expression of the music. This issue is identified by the production of *Lost Corps*, and is explored throughout the arguments developed in the later chapters. It is worth noting that this insight, into the relationship between music and spatial verisimilitude, does not concern technical challenges. It does not involve the *pursuit* of spatial verisimilitude, technological or otherwise, but rather highlights that music’s powers of representation continue to operate irrespective of a parallel *realistic* representation. The question the composer must address is how the two might coexist.

Ultimately, *Lost Corps* did not evolve beyond the initial compositional sketch. As an idea it highlighted that the rendering speed of ImageSynth has a significant impact on the composition process. Whilst it took a quarter of an hour to write scripts that could place notes in pre-determined places and move them as required, it would take several hours to render them into an audible work. As such, the ‘empirical and critical’ approach to composition, in which possibilities are generated then appraised, was strenuously challenged. The harmonic component of *Lost Corps* was conceived outside of the ImageSynth interface, but the conception of how it is projected in space required an iterative dimension; testing different ways of moving groups of instrumentalists around to understand how groups of tones responded to spatial movement. It is at this point in the production of the portfolio of works that a range of mechanisms was developed, within ImageSynth, to reduce the time between a spatial render and its subsequent appraisal. Ultimately, however, the technical
difficulties in maintaining control of all the parameters involved suggested that it might be worth exploring a different compositional approach that is less dependent on an iterative cycle where possibilities are first generated then appraised.

Strictly speaking, it is the pursuit of spatial verisimilitude that has introduced significant delay in producing spatial scenes. Computational resources are consumed as a result of the mimetic simulation of sounds in space. As such, the restrictions placed on the compositional process by ImageSynth can be regarded as much an effect of the pursuit of verisimilitude as an effect of the design of ImageSynth.

*Hammer* attempts to simulate the sound of many drummers moving in groups. The compositional interest in exploring rhythmic structures in space involves the differences in time delay caused by sounds generated at different distances from the listener. Thus, as groups of drummers move through space they will fall into and out of synchronisation with each other. This concept owes something to Charles Ives’ *Three Places in New England* (1911-1914) in which:

> [There] is a scene which represents the meeting of two marching bands in the village, and they play the same marching rhythm but at different tempi because they are coming from different directions (Nicolas Slonimsky in Perlis 1974, p.148).

In the second movement of *Three Places in New England* two marching bands are represented by the opposition of different musical forms performed simultaneously but at different tempi. ImageSynth is not restricted by the practical limitations of instrumentalists and so temporal differences between rhythmic parts can be explored with a level of accuracy that introduces subtle variations in rhythmic alignment.

Having found great difficulty in intimately controlling large numbers of pitched instruments in *Lost Corps*, I implemented a concept that allows the algorithmic control of large numbers of drummers. The concept involves casting the 32 binary bits, used to represent a single pixel’s colour, as 32 on/off switches representing 32 beats in a sound file. By choosing different colours, many combinations of rhythms could thus be determined. Each combination is created by playing back different
parts of a sound file that contains a recording of a drummer performing for a length of 32 beats.

The realistic modelling of the concurrent sound of many drummers proved to be a significant exercise, which was only partly fulfilled. As a result, the compositional idea of having different rhythms fall in and out of synchronisation with each other barely appears in the final work. Much like Lost Corps, compositional concerns were strangled by the effort absorbed in the pursuit of verisimilitude. There are significant difficulties in simulating large numbers of drummers from a single recording. Each playback of the recording needs to be sufficiently decorrelated from others such that it is perceived independently, yet the process of decorrelation can audibly affect sounds. This is particularly the case for drum sounds, which contain many audio transients. Different decorrelation techniques were tested, from the use of low frequency oscillation to introduce variable subtle time delays during the playback of drum sounds, to subtle manipulations of different granular parameters such as grain size, grain length and grain pitch. Manipulation of granular parameters needed to be particularly conservative to avoid significant audio artefacts. Other decorrelation techniques exist, and further efforts in simulating groups of drummers would warrant their exploration. Within the context of the exploration of the relationship between composition and the pursuit of spatial verisimilitude, however, Hammer did not warrant further exploration. Its compositional idea would be more easily realised using entirely separate recordings of rhythmic sounds, and its development has already contributed two significant insights.

The first insight consists of recognising the viability of an entirely different strategy for pursuing spatial verisimilitude: the empirical appraisal of the quality of verisimilitude. Many minutes of sound were generated, and the convincing illusion of large groups of drummers spread out in space remained elusive. Only a tiny fraction of the sounds generated were used within Hammer. It was found that by carefully choosing small fragments of the resultant rendered spatial scenes, the impression of realism could be significantly increased. Within each spatial render, isolated parts provided a more convincing illusion of reality. Emphasising these fragments resulted in an augmentation of the perception of realism. Here,
verisimilitude is pursued not by the technological simulation of acoustic reality, but rather by the empirical judgement of whether what is heard ‘appears’ real or not. This seems like a reasonable strategy but it raises a question about the correlation between mimetic accuracy and the preconception of the appearance of reality. As is explored in Chapter 5, what ‘appears’ real only needs to match our expectations of what reality sounds like, and those expectations might not necessarily mirror realistic acoustic modelling. Such an insight challenges the approach taken by reality-equivalent technologies that understands verisimilitude primarily as the mimesis of acoustic reality.

The second insight brought by *Hammer* concerns changes in verisimilitude as a result of the editing and re-organisation of sound snippets in a DAW. It was found that any reduction or increase in loudness of one sound snippet, relative to another, would very quickly degrade the sense of convincing spatial verisimilitude. This insight, in which standard audio processing techniques available in DAWs are understood to have a spatial effect, is explored in Chapter 4.

In both *Lost Corps* and *Hammer* compositional concerns are largely eclipsed by the effort required to pursue verisimilitude. The following work, *Gorged*, thus intentionally focuses on satisfying compositional aims, at the expense of verisimilitude. This approach holds interest for two reasons: firstly, the compositional possibilities of ImageSynth can be explored without being side-tracked by issues of verisimilitude; secondly, ImageSynth’s encapsulation of realistic acoustic modelling can be appraised since it will be exclusively responsible for the resultant spatiality.

*Gorged* took three months of full-time effort to create, and uses a score comprised of over 600 images. Its realisation involved a minimum of software development effort. *Gorged* uses a range of spatial and timbral gestures that are born out of the exploration of the idiosyncratic characteristics of ImageSynth. In other words, while the composition is a product of ImageSynth, the aesthetic parameters it explores are not something that the design of the software deliberately anticipated. The compositional approach employed here is exclusively the ‘empirical and critical’
process described earlier. The images employed were cyclically edited, rendered, appraised and discarded or retained. Different spatial gestures are explored; from the juxtapositions of near and far sound to the grouped movement of sounding objects. One spatial gesture employed approaches Xenakis’ technique of modelling the movement of gas particles via statistical algorithms (Xenakis 1992, pp.60-61). In this case, several hundred sounding objects are organised in a grid around the listener, then quickly moved to random positions. This produces the set of short gestures with rich timbres that dominate the first half of the composition.

Working on Gorged highlighted three key points. Firstly, the release from the concern of spatial verisimilitude provides a musical freedom in which the possibilities inherent to the compositional environment developed can be fully explored. This freedom is perhaps akin to compositional developments in the 20th century where, released from restrictions such as harmonic structure, composers have been free to explore new musical potentials. Here the pursuit of spatial verisimilitude is clearly characterised as a restriction to musical exploration. Secondly, the design of ImageSynth imposes significant limitations as to how sound can be scored and controlled. Thirdly, I found that working within those limitations is more liberating than developing more software to work around them. In other words, the creative exploration of the limitations of the technology is more compositionally empowering than the development of technology to explicitly serve a compositional aim. These last two points engender broad critical discussion concerned with the relationship between technology and music making; particularly in how the technology used leaves its trace on the resultant music. Within the specific concerns of this thesis, however, it is the understanding that technology has an important impact on music making that is of significance. This line of enquiry is explored in Chapter 6, which examines how reality-equivalent technologies mediate the compositional act.

The spatial verisimilitude in Gorged is limited. The use of a reverberant field means that distal movements, such as the opening spatial gesture, are well perceived. Directional information is clear when the hundreds of sounding objects are grouped, but it is obfuscated when each sounding object moves in different directions. Granular synthesis techniques are deliberately parameterised to produce a wealth of
different timbres, which are further enriched through the random and quick panning of sounds around the listener. Here, synthesised spatial movement contributes more to timbral complexity, than to the perception of space. The composition holds more detail and complexity in its use of textures than in its spatial dimension, but the textures created owe some of their complexity to the spatial distribution of sounds. This demonstrates that spatial gestures are intimately linked to the perceived timbre of sounds. Research discussed in Chapter 5 suggests that the use of abstract sounds reduces the plausibility of spatial sound scenes. Thus, the limited spatial verisimilitude in *Gorged* is partly attributable to its extensive use of abstract sounds.

As a composition, *Gorged* introduces a wealth of complex timbres intimately linked to spatial movement. However, despite the effort that went into producing this work, I feel that its temporal development would benefit from further development. The criticisms I hold, concerning the compositional structure of *Gorged*, are attributable to the difficulty in controlling the sounds produced. Due to the complexities of the spatial gestures and timbres created, it is extremely difficult to predict how a render will sound. The results of even minor adjustments to image sequences, or the mapping strategy used, would often be unpredictable. As such, the compositional act consists of rendering as many spatial scenes as possible and subsequently choosing, empirically, the ones that best fit the compositional direction of the piece. Since one render requires around ten hours of computation, even small adjustments involve copious amounts of waiting, which may not always produce satisfactory results.

The last composition produced aims to avoid the challenges caused by the ‘empirical and critical’ approach to composition. *First Flight Over an Island Interior* is the result of the conception of a compositional structure that is finalised *before* being rendered by ImageSynth. The composition does not aim to explore the inherent limitations of ImageSynth, but rather to work with them. The compositional process is inspired by Xenakis’ approach, in which a general abstraction is adapted to musical concerns. This abstraction, however, needs to fit within ImageSynth’s encapsulation of the spatial disposition of sounding objects.
In *Lost Corps* pixels represent sounding objects that have pitch. The impression of movement is created by sequentially changing the colours, and therefore pitches, of adjacent pixels. Another way to create the impression of movement is to fix the colour, and pitch, of each pixel and simply move the entire grid of sounding objects relative to the listener. In this technique, all pixels move in the same direction and at the same speed. Only one movement is modelled. It might be better to describe this organisation as the listener moving through a field of sounds, rather than sounds moving around the listener. This simple inversion restricts the movements possible but also expresses an abstraction of the experience of moving through stationary objects; a solitary car driving through a remote landscape, a bushwalker walking through a still forest, or a train travelling through consecutive towns. It is a very simple abstraction, but holds strong associations with human experience. It is thus deemed a good starting point for an approach where the principal structural order of the composition is determined prior to rendering any spatial scenes. The musical expression of this abstraction introduces other dimensions. By expressing each sound as a fixed pitch, sequences of pitches are crafted by moving through the grid using different trajectories. Here, a parallel can be drawn with serialism since the sequential relationships of all pitches remains fixed, irrespective of the direction of the sequence. Of course, unlike serialism, the choice of pitch sequence can be explored in two different dimensions thus allowing a variety of sequences. Also unlike serialism, all the pitches present on the grid can always be heard at any one time; individual pitches can only be emphasised by spatial proximity to the listener.

This abstraction was discussed with Australia-based, American composer Kraig Grady. A student of tuning theorist Erv Wilson, Grady immediately recognised the compatibility between this spatio-temporal structure and the just-intonation techniques he uses in his work. Just-intonation refers to the technique of using pitch relationships that are mathematically pure. The ‘tempered’ twelve-note tuning common to western music requires the adjustment of pitches to ensure symmetry at different octaves and keys. One advantage with just-intonation is that all notes within the key have a “true”, as in mathematically pure, harmonic relation to each other and they therefore sound more consonant. Within the tempered tuning system the simultaneous presence of all twelve notes in the chromatic scale creates a tone cluster...
(Carter 1965, p.10) in which there is a finite number of difference frequencies produced between any pair of notes. By using just-intonation, however, the entire grid of notes produces an infinitely complex harmonic sonority. Different aspects of this sonority can be emphasised through controlled movements in which different combinations of notes receive different emphases. As the listener moves through the grid, pitches that are physically closer are emphasised through loudness, and their harmonic relationship to other proximate pitches is highlighted. Movement through the grid thus creates a complex harmonic progression, with melodic highlights, that is always underscored by the complex sonority of the entire grid.

Grady designed two just-intonation grids that were subsequently rendered in ImageSynth. Each grid is comprised of a mathematical series of frequencies that is expressed in two dimensions. The series were designed to ensure that all frequencies remain related by ratios of small whole numbers.

Each grid requires only one image, and represents 2500 sound sources. Figure 15 shows one of these grids: the brightness of each pixel is used to determine the resultant just-intonation pitch.
Much work was done in testing different sound synthesis techniques to voice the just-intonation pitches. Due to the spatio-temporal structure of the piece pitches are heard at vastly different distances, from over 400m away to 2m away. This posed a challenge in that the synthesis technique used had to provide sonic interest at both near and far distances. It was found that different synthesis techniques created very different results at different distances. For example, SuperCollider’s implementation of Xenakis’ GENDY algorithm (Xenakis 1992, pp.289-322) produced excellent results when heard from far: the resultant sound mass has a textural complexity with a realistic physicality that is engaging to the ear. Heard from close, however, the sound created by the GENDY algorithm gives itself away as a coarse digital synthesis of real-world sonorities. It is worth noting that SuperCollider’s implementation of the GENDY algorithm exposes parameterisation that allows for

Figure 15. A grid of just-intonation pitches designed by Kraig Grady and modelled within ImageSynth.
the very clear perception of pitch, which is important for working with just-
intonation. The use of simple sine oscillators has different qualities. Heard from mid-
distance, the sine oscillators provide a warm and listenable harmonic mass, but when
heard from close the lack of transients in the signal means that the perception of
direction is significantly reduced. As such, when the sine tones are heard close to the
listener there is a loss of the perception of proximity and movement. The use of saw
waves also has its own set of qualities. Heard from close the saw waves create a very
strong sense of proximity and movement, but the tone does not have the same
warmth as the sine waves. Neither the saw nor the sine waves have the timbral
complexity and interest of the GENDY algorithm when it is heard from far. Some
spatial renderings created employ a finely tuned mix of the above synthesis
techniques. Of course, many other synthesis techniques exist, such as Karplus-Strong
synthesis and Chowning’s frequency-modulation synthesis. The testing of these and
other synthesis techniques would be the subject of further research.

Whilst no singular synthesis technique proved satisfactory, in the rendering of First
Flight Over an Island Interior, the combined presence of different renderings
provided a satisfactory balance between the timbral complexity of far sounds, the
harmonic clarity in mid-distance sounds, and the localisation accuracy and timbral
interest of near sounds. No further testing of synthesis techniques was thus deemed
necessary. The sequential presentation of different passes through the grid, using
different synthesis techniques, introduces another compositional dimension that
juxtaposes the specific qualities of each synthesis technique. A singular pass through
the grid takes around ten minutes; multiple passes thus create a significantly longer
work, and extend the abstraction of moving through a scene into an extended
journey. The resultant composition is the combination of half a dozen passes through
Grady’s two different grid designs, organised in time by Grady and myself.

First Flight Over an Island Interior contributes two important insights to the
relationship between music and spatial verisimilitude. Firstly, the appropriation of
movement to pitch and harmonic relationships causes some inconsistencies. Since
the Doppler shift effect is modelled, the movement in the grid causes the pitches to
be altered. As such, the purity of the mathematical relationships between pitches is
compromised. The Doppler effect is most pronounced in pitches that are very close to the listener, so the background harmonic sonority characteristic of just-intonation tunings is largely unaffected. However, the melodic component of the work, created by pitches in close proximity to the listener, is audibly affected. In other words, *First Flight Over an Island Interior* confirms that certain compositional elements, such as pitch based musical constructs, mostly assume a static spatial presentation in which changes in perceived pitch do not occur.

As in *Lost Corps*, *First Flight Over an Island Interior* is also subject to perceptions of height caused by the extensive use of high frequency pitches. Whilst there is no vertical modelling involved in this work, height can be perceived in parts of the composition where incremental increases in pitch feature prominently. Here, unlike the previous insight, it is a musical construct that has the capacity to corrupt the physical spatial modelling; not the physical modelling that corrupts the musical construct. In other words, spatial verisimilitude and music challenge each other.

### 2.4 Summary of research conclusions

The realisation of the portfolio of works highlights that the pursuit of spatial verisimilitude is difficult. It is difficult because the simulation of the behaviour of sound in space is complex, it is computationally intensive and, in terms of the quality of the illusion achieved, its results are often disappointing.

Individual works demonstrate that spatial illusion has a dependence on the choice of sound material, which is in turn informed by compositional considerations. Conversely, musical constructs have the capacity to themselves create perceptions of space outside of any spatial modelling. In other words, compositional concerns and spatial verisimilitude are not clearly distinguishable.

In light of this apparent interdependence, the first research direction seeks to understand whether other spatialisation techniques have a similar relation to music. In the following chapter, spatialisation techniques are first understood by the perceptual mechanisms they engage. This understanding is then contextualised with examples of spatial music. How the pursuit of spatial verisimilitude thus differs from other spatialisation techniques is outlined.
3 SPATIAL INFORMATION ENCODING MODES: A FRAMEWORK

In 1954 Jacques Poullin, Head of Works at the Groupe de Recherches de Musique Concrète (GRMC) in Paris, wrote an essay (1999) detailing the achievements of the GRMC in developing new techniques for the creation of new musical forms. Poullin was a colleague of Pierre Schaeffer, the inventor of Musique Concrète who is also responsible for many early observations on the perception of sound. The essay presents a variety of new instruments and techniques developed, including the Phonogène; a variable speed tape player operated by a one-octave keyboard. What is of interest here, however, is a short passage where Poullin describes recent experiments (1999, section 5.2) that have demonstrated that a sound can be positioned anywhere, on a straight line, between two speakers. The technique described, effectively a form of stereo panning, is significant to Poullin because sounds are freed from being tied to the position of speakers. By manipulating the relative volumes of a single sound distributed between two speakers, a phenomenon that contributes to how we experience and perceive the direction of sound, now known as interaural level differences (Blauert 1997, pp.155-164), is replicated. This technique, in which the direction of sounds is perceived, represents an early example of the synthetic simulation of a spatial attribute of sound.

Poullin then goes on to explain how the addition of a second but very different technique allows an unprecedented level of control of the locations of sounds:

[The] theory of sounds in space in combination with our radiophonic listening habits, which have familiarised us with the subjective localisation of sound sources within the dimension of depth (or, more exactly, the notion of proximity without the clear distinction of direction), have enabled the use of stereophonic processes, using two channels of sound, to reconstitute with sufficient detail sound sources within a spatial perspective which would otherwise only be possible with a three channel system. (1999, section 5.2, my translation)

Poullin is essentially describing how sounds can be positioned on a two-dimensional plane by combining carefully chosen recordings with stereo panning. In this
technique the perception of depth is created in a different way to the perception of direction. Instead of being simulated, depth is referenced through recordings that contain sounds and qualities of sounds recorded at certain distances. As examples of such recordings, consider the sound of a tractor working in a faraway field, compared to the sound of a child whispering a secret. As recorded sounds, both the tractor and the voice might be of identical loudness but their distances will be perceived very differently. The recognition of the object that is making the sound, in combination with the recognition of certain qualities within the sound, allows a perception of how far away these objects are. The tractor, for example, is heard at a low volume compared to the subtle whooshing of the wind; its low rumbling sounds are more audible than its higher frequency sounds; and its loudness seems to quiver as the wind changes direction. These sound qualities act as cues that reveal that the truck is far away. A whisperer, on the other hand, must be close if they can be heard. This is the spatial encoding of distance by reference to experiential memory. Having acquired knowledge about the relative loudness of tractors and their environments, and the softness of whispering, the listener can immediately deduce an approximate measure of distance.

![Figure 16. A sound’s position can be represented as the combination of angular direction and distance](image)

Figure 16. A sound’s position can be represented as the combination of angular direction and distance
In Poullin’s proposed technique, stereo panning *simulates* the direction of the sound, whilst distance is *referenced* through the contents of mono recordings. As Figure 16 illustrates, these two parameters; direction and distance, work together to represent the position of a sound relative to the listener. The location of the sounding object is therefore communicated using two separate techniques each encoding a spatial attribute in a significantly different way. The result, however, is the perception of a singular attribute: the position of a sounding object.

Some 23 years after Poullin explained how depth could be represented through sound recordings another music researcher described how it could be simulated. John Chowning (1977), as introduced in Chapter 1, has described how the perception of a sound’s location can be created by simulating both direction and distance. Within Chowning’s technique, three distance cues are simulated. That is to say, Chowning simulates three aspects of how sound behaves in space, which contribute to the perception of distance. The cues chosen by Chowning are the ratio of the loudness of the direct sound of the sounding object compared to the loudness of the diffuse field; a change in the directionality of the diffuse field where nearer objects have a more immersive diffuse field; and the gradual loss of low-intensity frequency components as the sound moves away from the listener (1977, pp.48-50). Delivered over four speakers positioned around the listeners, these simulations give rise to a perception of distance in a very different manner to that discussed by Poullin. Chowning’s technique *simulates* changes in the physical attributes of sounds caused by their travelling through physical space. Poullin, on the other hand, relies on the recognisability of recorded sounds to suggest certain spatial characteristics. Whilst the resultant spatial attributes perceived are similar, the perceptual mechanisms involved in creating those perceptions are significantly different.

### 3.1 Spatialisation techniques understood by perceptual mechanism targeted

This chapter introduces a new perspective for understanding the relationship between music and space: one that is focused on the *perceptual dimension* of spatialisation techniques. The key notion explored concerns how different spatialisation techniques engage different perceptual mechanisms, which in turn affects the kinds of spatial attributes available to compositional thinking. This perspective challenges the notion that spatialisation technique can be read as a model of composer’s conception of
space. Indeed, it is argued that spatial compositional approaches reflect the intrinsic qualities of different spatialisation techniques to a greater extent than spatialisation techniques might reflect composers’ conceptual models of space. The framework presented in this chapter, and its subsequent illustration through examples of compositional work, suggests that the understanding of how space has been used in music is perhaps most clearly understood as a function of the perceptual characteristics of spatialisation techniques, rather than as a function of musical-conceptual models of space. In short, to understand how composers use space, we need to understand the perceptual dimensions of their spatialisation techniques.

This perspective subsequently suggests that compositing with spatial verisimilitude necessitates a deep understanding of the perceptual dimension of reality-equivalent technologies. Indeed, the examination of one scientific approach towards spatial fidelity in reproduced sound reveals a concern with a set of spatial attributes that barely registers in the critical discussion surrounding spatial compositional strategies. It is shown that frameworks for the analysis of space, as proposed within literature concerned with composition, tend to consider space independently of the biases of spatialisation techniques. As such, if the compositional engagement with reality-equivalent technologies adheres to these frameworks, then some important spatial attributes unique to reality-equivalent technologies will be ignored.

The categorisation of spatialisation technique by their perceptual dimensions highlights another key concern for reality-equivalent technologies: the perceptual mechanisms targeted by reality-equivalent technologies do not exclude the activity of other perceptual mechanisms. This holds important implications for composers working with spatial verisimilitude, since they must consider perceptions of space created beyond the reach of reality-equivalent technologies.

The different ways in which a spatial attribute comes to be perceived in sound will here be referred to as perceptual pathways. This term encapsulates everything that happens from the creation of a sound to the perceptual event. It includes the mechanical phenomena of sound waves travelling in space, sensory phenomena occurring at the ears, and finally the neurological event of perception. In short, the
term "perceptual pathway" includes combinations of the action of the physical environment, the ears and the brain. This includes the presence of any cognitive processes, defined by Begault (2000, p.29) as the “higher-level processes of memory, understanding and reasoning”, that might be involved in a listener being conscious of a space, or a spatial attribute. To illustrate an example of different perceptual pathways engaged by different spatialisation techniques, consider the difference between Poullin and Chowning’s techniques for encoding distance. Poullin’s technique, which relies on knowledge of the relative loudness of sounds, effectively targets cognitive processes. Chowning’s technique of simulating the physical attributes of sound in space is concerned primarily with recreating auditory stimuli. Whilst different techniques may be concerned with specific aspects of our perceptual mechanisms, it is important to note that all other aspects are still at play. Herein lies a key disadvantage of all spatialisation techniques described in this chapter: they tend to focus on one perceptual pathway despite the continuing presence of others.

![Diagram of perceptual pathways](image)

Figure 17. Notional outline of steps involved in the perception of sound.

Figure 17 illustrates how certain spatialisation techniques might target different perceptual pathways. The first row of captions provides a notional outline of the steps involved in a spatialised sound being perceived. The second row of captions...
describes examples of spatialisation techniques that target the associated step. Each component described in the top row is always active, irrespective of the spatialisation technique used. Sound is always transmitted in a physical space\(^7\), it is always received by the ears, and is always interpreted by our cognitive faculties.

The consideration of the role of spatialisation technique within spatial composition is intimately related to critical enquiry examining the role of technology within electroacoustic music. Composer and researcher, Agostino Di Scipio, argues that the \textit{techné} of electroacoustic music, which he explains is ‘the realm of techniques and technology captured in the creative process of music composition’ (1995a, p.371), defines aspects of ‘the cognitive and aesthetic potential specific to the medium’ (1995a, p.380). This line of inquiry, explored in detail in Chapter 6, is related to the concerns of this chapter but considers a broader perspective on the technological dimension of music making. Here, the primary concern is in outlining the relationship between the perceptual dimension of different techniques, and compositional concerns. We are not concerned with how the technologies that embody those techniques might bring further mediation to the compositional act.

Examples presented below illustrate the approach taken in this chapter. These examples outline three key points. Firstly; the spatialisation technique used affects both the kinds of sounds that can be used and how they are used, secondly; different spatialisation techniques can coexist, thus leading to multiple sources of spatial information, and thirdly; different spatialisation techniques can expose significantly different spatial qualities.

Poullin’s approach of relying on the contents of sound recordings to reference distance imposes restrictions on the sound material; other sounds also captured by the recorder may be present; distances and movements will be limited to what can be captured within recordings; and other spatial information may also be present. Chowning’s technique, which relies on the presence of a reverberant field to contextualise perceptions of distance, will favour sounds modelled within

\(^7\) Even if listening to sounds on headphones, the listener will always be present in a physical space whose acoustic will contextualise what is heard through the headphones.
reverberant environments. Indeed, Chowning’s composition *Turenas* (1972), in which sounds are panned on a two-dimensional plane, makes extensive use of reverberation. In other words, different spatialisation techniques summon different classes of sound material.

Different techniques for encoding the same spatial attribute may also result in entirely different sounds being used. Whilst a simple recording of trees swaying in the wind can reference the experiential memory of an outdoor space (Barrett 2002, pp.315-316), the acoustic simulation of an outdoor space may involve modelling a scene that does not include the presence of wind.

Different spatialisation techniques can coexist. This second point requires some more detailed discussion because it is of significance to reality-equivalent technologies. In Poullin’s technique, recordings are chosen specifically because of their ability to reference distance. In Chowning’s technique, distance is simulated. The simulation of distance involves the application of audio processing techniques that operate on sounds regardless of their contents. In other words, the technique of simulation introduces dissociation between a sound and its spatialisation. The implications of this dissociation are clear when one considers that distance can be simulated on a recording that already contains the sound of a distant truck. In such an example distance is encoded using two different techniques: Chowning’s simulation and Poullin’s reference. Whilst these two different techniques can technically coexist the result may give rise to conflicting spatial information. Speculation might ensue on which cue might take precedence in determining the distances perceived: the simulated or the referenced. In the real world, visual cues may contribute to the resolution of conflicting perceived information but, limited to auditory stimuli, how the mind of the listener might resolve such contradictions is difficult to foresee.

Chowning specifically states, in his paper, that the sounds being panned in the illusory acoustical space are synthesised (1977, p.48). One might speculate that this style of simulation is therefore only pertinent for the projection of abstract sounds which have no direct association to known sounding objects, and therefore to their spatial circumstance. However, even the random processing of synthesised sounds
can give rise to perceptions of space (Smalley interviewed in Austin 2000, p.14).
Indeed, Worrall (1998, p.95) explains that “timbral modifications often suggest
spatial modulations”, thereby highlighting that processing techniques operating on
the timbre of sounds can also result in perceptions of space. In other words, the
perceptual pathway targeted by the simulation of distance leaves the door wide open
for the activity of other perceptual pathways, whether it be through referential cues
present in sound recordings, or timbral changes present in processed abstract sounds.

The spatial attribute of distance is known to be significantly dependent on familiarity
with the sound source (Blauert 1997, p.45). This suggests that perceptions of distance
are favoured by spatialisation techniques that target the cognitive dimension of sound
perception. It also suggests that the success of the simulation of distance, as
described by Chowning, remains subject to the nature of the sound being spatialised.
Reality-equivalent technologies have no intrinsic understanding of cognitive
processes and thus leave a wide potential for the action of other cues. Begault (2000)
notes, in his text *3-D Sound for Virtual Reality and Multimedia*, that spatial cues
“can be rendered ineffective by the summed result of other potential cues” (p.69).
Thus, the strict adherence to acoustic reality as pursued within reality-equivalent
technologies does not guarantee the accuracy of the listener’s perception.

In his text *Listening: An Introduction to the Perception of Auditory Events*, Stephen
Handel (1989, pp.111-112) suggests that our perceptual abilities are more likely
limited by our ability to *select the relevant information*, than by insufficient
information. Contextualised to spatial concerns; Handel’s suggestion says that a
listener’s ability to perceive the designed spatiality will more likely be limited by
their ability to select the right spatial information, than by enough spatial
information. This view lends further weight to the issue of co-existing spatialisation
techniques. If our perceptual faculties are limited by their ability to choose the right
information, then the presence of any inconsistent or conflicting spatial information
born of different perceptual mechanisms is a problem. The pursuit of spatial
verisimilitude might therefore be better served by the detection and suppression of
extraneous spatial information, than by the creation of more detailed acoustic
simulation.
The third example of how the perceptual dimension of spatialisation techniques has an impact on compositional concerns involves the resultant spatial attributes perceived. Different spatialisation techniques expose access to, or favour, different spatial qualities. This can be illustrated by considering Chowning’s technique of using reverberation to simulate distance. By depending on the presence of a spatial reverberant field to encode distance, an entirely new attribute is produced: a sense of immersion. Begault (2000, p.69) states that reverberant fields contribute to both perceptions of distance and perceptions of environmental context, which work together to engender a sense of immersion. Thus, the simulation of distance can result in the creation of the spatial quality of immersion. Immersion is one of a class of spatial qualities, discussed shortly, that have received very little attention within the critical enquiry concerned with spatial composition strategies. An understanding of the perceptual dimension of spatialisation techniques thus opens up new compositional opportunities not considered within known spatial compositional strategies.

All of the above insights originate from the characterisation of spatialisation technique as targeting specific perceptual pathways. This characterisation provides a new perspective for understanding the relationship between music and space, and underpins the definition of a new framework that is defined to have the following intentions: identify aspects of spatial composition that are mediated by spatialisation technique; facilitate the identification of conflicting or contradictory spatial information and; identify the spatial qualities favoured by different spatialisation techniques.

This framework is formalised shortly, in Section 3.3. To contextualise its significance existing frameworks for understanding the relationship between space and music are first examined.

3.2 Existing frameworks for spatial analysis

Frameworks for the analysis of space within sound occur in different disciplines. Within musicology, concern is focused on the compositional meaning of space, for example: how spatial design can assist in the temporal structure of the music. Within
scientific endeavours frameworks for spatial analysis are more concerned with the perceptual aspects of spatial audio. Both are of interest to the concerns of this thesis, and a selection of the relevant literature is reviewed below.

Within the electroacoustic music literature, paradigms for describing space often involve the identification of abstract spatio-temporal structures; that is, abstractions of space and time that are independent of the biases of spatialisation techniques. Emmerson (1998a, 1998b) and Smalley (2007) have both proposed scene-based paradigms that purport to delineate space into compositionally meaningful components. Emmerson deploys the notion of frames at different scales to identify areas of interest: in increasing order of scale these are; event, stage, arena and landscape (1998a, p.138). Smalley’s framework extensively categorises both spatial and spatio-temporal constructs as intersected with many non-spatial properties of sound. Categorisations such as agential space\(^8\), vectorial space\(^9\), and arena space\(^{10}\) display the breadth of Smalley’s framework but also suggest that all classes of sound have a categorical relationship with space. Whilst both Smalley and Emmerson discuss, in some depth, the complexities of the different ways space is perceived in electroacoustic music, the abstraction of space into compositionally meaningful components is considered largely outside of perceptual phenomena. A spatio-temporal structure, as conceived in the mind of the composer, might involve contrasting an indoor space with an outdoor space, but the choice of spatialisation technique employed to create the perception of those spaces will bring a significant mediation that may result in very different compositions with very different aesthetic dimensions.

Contemporary British composer Trevor Wishart proposes a simple abstraction of the characteristics of ‘aural landscapes’. He defines three key components; ‘I: the nature of the perceived acoustic space; II, the disposition of sound-objects within the space; and III, the recognition of individual sound-objects’ (1986, p.45). Such an outline of

\(^8\) Agential space is space defined by the activity of a human agent (Smalley, 2007, p.41).

\(^9\) Vectorial space is the “space traversed by the trajectory of a sound” (Smalley, 2007, p.37).

\(^{10}\) Arena space is “the whole public space inhabited by both performers and listeners” (Smalley, 2007, p.42).
spatial concerns does not specifically acknowledge the biases of different spatialisation techniques, but within Wishart’s carefully chosen language is the acknowledgment of the complexities of perception. Referencing the importance of cognitive processes relying on memory and experience, Wishart illustrates that technologies such as portable radios have introduced new modes of presenting sounds. These new modes effectively expand the range of plausible spatial interpretations. For example, the sound of an orchestra playing may suggest landscapes very different to the orchestra’s native concert hall:

[It] becomes possible to question, for example, whether we hear the sounds of an orchestra, the sounds of a radio (playing orchestral music), or the sounds of a person walking in the street carrying a radio (playing orchestral music). (1986, p.46)

Here, the subject of the recording remains the sound of an orchestra, but the physical spaces potentially perceived are significantly different. Prior to our familiarity with the sound of portable radios, the spatial interpretation of the sound of an orchestra would be different. Whilst Wishart largely considers perceptual pathways that contain a significant cognitive component, he concludes the chapter by mentioning that developments in the control of “virtual acoustic space” open new musical possibilities (1986, p.60). The control of virtual acoustic space, which is also the domain of reality-equivalent technologies, primarily involves the simulation of auditory stimuli and excludes a strict concern for cognitive processes. In other words, Wishart recognises the complexity and importance of the perceptual dimension of working with space in sound, but his abstraction of ‘aural landscapes’ does not consider the differences between spatialisation techniques.

Taking an entirely different approach Barry Truax (2002) characterises soundscape composition, a form of spatial composition pioneered at Simon Fraser University, by observing that “most pieces can be placed on a continuum between what might be called ‘found sound’ and ‘abstracted’ approaches” (p.6). Found sounds will typically contain strong associative powers, thereby engaging perceptual pathways that reference qualities of space through recognition of a sound’s identity. Abstract sounds and abstracted temporal forms will have less recognisability. At the
‘abstracted’ extremity of that continuum, Truax questions whether a soundscape is at all possible:

Where few if any of the environmental sounds used in a piece are recognisable, the listener will probably hear it as an abstract sound organisation, not as a soundscape composition with real-world associations. (2002, p.6)

By questioning whether unrecognisable sounds have the capacity to form a ‘soundscape’ Truax is clearly not concerned with the simulation of space. His use of the words ‘real-world associations’ references a conception of soundscape focused on acoustic ecology; ‘the study of the interrelationship between sound, nature and society’ (Westerkamp 2002, p.52). Acoustic ecology is an area of study initiated by composer Murray Schafer, a colleague of Truax, who founded the World Soundscape Project also at Simon Fraser University. To question that a lack of real-world associations can create a soundscape at all indicates a conception of space focused on perceptual pathways that favour cognitive dimensions. Whilst Truax’s discussion makes little mention of the role of technique within soundscape composition, he does observe that the use of an eight channel playback format has “achieved a remarkable sense of immersion in a recreated or imaginary sonic environment” (2002, p.1). In so saying, Truax has identified a particular spatial quality, which he calls immersion, created as a result of a particular spatialisation technique. However, he does not specifically comment on the potential role of immersion within soundscape composition.

One notable exception to the musical conception of space as abstract spatio-temporal structures can be found in the work of Natasha Barrett: a British electroacoustic composer. In a paper entitled Spatio-musical composition strategies (2002) Barrett outlines the qualities of space made accessible by different spatialisation techniques. She uses the term Spatial Illusion to classify sonically produced spaces that have a likeness to reality (p.314). Spatial Allusion refers to spaces perceived as a result of the listening imagination (p.315) and Simulation of the 3D soundfield refers to what can be achieved with what is here categorised as reality-equivalent technologies (p.317). Barrett’s approach, which is in fact informed by her familiarity with and use
of reality-equivalent technologies, is mostly analogous with the ideas presented here. How her scheme differs from the framework outlined shortly is discussed later.

Frameworks for analysing auditory scenes are also common within the scientific literature concerned with the perception of sound. Bregman’s text *Auditory Scene Analysis* (1990) seeks to analyse auditory scenes into their most perceptually significant component parts. Bregman uses the concept of *Auditory Stream Segregation* to refer to the perceptual isolation of streams of sound based on spatial, spectral, timbral and other factors. Such characterisations of perception pertain to all sounds, including those that make up a compositional work. Indeed, Bregman’s concepts surrounding stream segregation have been successfully applied to the study of music (Wright, J & Bregman 1987), and in some cases to the study of spatial music (Harley 2000). Proposed frameworks as developed by composers and music researchers such as Emmerson and Smalley differ in that they are concerned with the identification of the most *compositionally significant*\(^{11}\) component parts, as opposed to the most *perceptually significant* component parts.

Francis Rumsey, author of *Spatial Audio* (2001), and chairperson of the committee that developed the ITU 5.1 standard (Barbour 2002, p.2), proposes a framework (2002) for the subjective analysis of spatial qualities. Rumsey’s characterisations specifically concern the spatiality of *reproduced* sounds. In other words he has an underlying concern with the spatial fidelity of recorded sound. Fidelity is intimately related to the notion of verisimilitude, and Rumsey’s work is discussed further in Chapter 5. What is of concern here is Rumsey’s conception of a framework for analysing space. He adopts the common scene-based approach as a skeleton upon which to hang definitions of measurable subjective qualities. He does, however, also identify a new ‘class’ of spatial attributes. Rumsey (2002, pp.658-660) centres his scene analysis on the width, depth and distance of sounding objects, ensembles of sounding objects and the sounding environment. These metrics contribute to the first

---

\(^{11}\) *Compositionally significant* component parts refers to categorisations of sound that serve the composer’s goal, rather than reflect how sound is perceived. For example, Smalley distinguishes between enacted spaces; ‘spaces produced by human activity’ (2007, p.28) and naturally occurring spaces. Such distinctions may not have significance for categorisations based on *perceptually significant* component parts.
class of spatial attributes Rumsey calls “dimensional attributes”. But a second class is required to cater for the qualities that “are not perceived directly as linear quantities but more as semiabstract and multidimensional impressions" (p.661). Rumsey names this second class ‘immersion attributes’. He stresses that immersion attributes suffer the fate of multiple and subtly differing uses within the scientific literature. Terms such as listener envelopment, spaciousness and spatial impression all overlap and differ slightly in their use within the literature. Rumsey proposes to limit notions of immersion to two terms; envelopment and presence. Envelopment can be created on different levels; by individual sources, ensembles of sounds and by the environment. Rumsey’s definition of presence is quite different to that of Lombard and Ditton, introduced in Chapter 1. He defines presence as a singular quality that measures a listener’s sense of being inside an enclosed space.

Within the present context, what is of key importance in Rumsey’s research is that the class of spatial attributes he calls ‘immersion attributes’ is essentially not considered within frameworks proposed for the composition of spatial music. The abstractions proposed by Emmerson, Smalley and Wishart all consist of interpretations and extensions, including temporal extensions, of what Rumsey labels ‘dimensional’ attributes, without directly identifying notions of immersion, presence and envelopment. Smalley does use both words immersion and presence (2007) but their meanings are largely unrelated. Truax notes the effect of immersion, but does not consider it as an attribute of composition. The identification and compositional consideration of this class of ‘immersion attributes’, discussed in greater detail in Chapters 5 and 6, is one of the key contributions this thesis makes to the field of spatial composition.

There is little evidence of the understanding of the importance of spatialisation technique within composer’s models of audible space. Emmerson, Smalley and Wishart all propose abstractions of space that reflect a conceptual understanding of space, not a perceptually mediated one, nor a technically mediated one. Natasha Barrett’s work is one notable exception; she classifies spatial attributes by technique and discusses the compositional opportunities they offer. Barrett does not explicitly identify what Rumsey calls the class of ‘immersion attributes’ but her writing alludes
to an awareness of them. This thesis builds on Barrett’s line of enquiry by explicitly identifying new spatial attributes exposed by reality-equivalent technologies, and specifically uncovering how they might relate to compositional concerns.

3.3 The framework
The proposed framework identifies three spatial audio encoding modes formalised as a broad grouping of the strategies composers use to create spatial music. Each mode favours certain perceptual pathways and exposes certain spatial attributes. Each mode mediates how the composer thinks about sound. Each mode holds aesthetic implications. Generally speaking, whilst compositions will tend to be conceived within one mode, the other modes will often be present. These modes will first be described, and then illustrated by way of key compositional works.

The three modes are: the intrinsic, the referential, and the simulated. As discussed shortly, the referential mode is further subdivided into icon, index and symbol.

The intrinsic mode concerns the spatial encoding of sound caused by real physical space. It is most clearly present in spatial music where the spatial separation of instrumentalists is used to compositional effect. The acoustic of the performing/listening space, and how it changes sound that is ignited within it, is the domain of the intrinsic mode. The practical considerations of working with real physical spaces means this mode offers a limited palette of spatial gestures. Conversely, the quality of the spatial effect cannot be paralleled since it is not manufactured; it is real.

The intrinsic mode is also ubiquitous. It is always present, even when its effect is not sought in the projection of a spatial composition. For example, when listening to a work projected on stereo loudspeakers the spatiality composed within the stereo work is compositied with the spatiality of the present listening environment, that is, the intrinsic mode. Similarly, the intrinsic mode is readily captured within sound recordings. Depending on the recording technique used and the physical environment within which the recording takes place, the sound captured will typically have been affected by the present physical space. This spatial information belongs to the intrinsic mode of spatial encoding.
The referential mode refers to the encoding of spatial information through signs. It includes any spatial characteristic that is perceived outside of the veridical experience of sound in space, which includes both the intrinsic mode and the simulated mode. The referential mode is the broadest and most complex category, and it includes the effect of cognitive processes. It is explored in depth in Chapter 3, where it is shown that a typography of signs developed by the 20th century logician Charles Peirce provides a basis for the identification of the diverse ways in which spatial information can be referenced. Peirce defines three types of signs in which the subject is referenced by similarity, association, or agreed meaning. Within Peirce’s writing, these are labelled icon, index and symbol. As an example: the use of a whispering voice to encode distance can be described as a Peircian index of proximity. In words similar to those used by Peirce, it follows that when a whispering person can be heard, the speaker must be relatively close. The distinction between the different sign types, and their significance, is detailed in Chapter 4.

The simulated mode could almost be considered a sub-group of the referential mode. It is very close to Peirce’s icon type where a thing is referenced through similarity with it. Where the simulated mode distinguishes itself from the referential mode is that the similarity it pursues is so strong that the reference is mistaken for its referent. In other words, the simulated mode concerns the spatial encoding of sounds that seeks to be mistaken for being real. One might say that the simulated mode is a type of reference mode that attempts to create perceptions indistinguishable from the intrinsic mode.

Both the intrinsic mode and the simulated mode concern spatial encoding that results in the veridical experience of sound in space. The simulated mode, however, fakes this effect. As such, the simulated mode distinguishes itself from the intrinsic mode by the fact that it is an illusion: a deceit. Reality-equivalent technologies operate within the simulated mode.

---

12 The word ‘symbol’ might have also been used here. The word ‘sign’ is preferred because within Peirce’s semiotic framework the word ‘symbol’ is reserved to denote a specific type of sign.
The simulated mode can also be described as aspirational: when it succeeds it approaches the effect of the *intrinsic* mode; when it fails it approaches the effect of the *referential* mode in that the ‘likeness’ it has created has failed to convince, and thus becomes merely representational. Thus a characteristic of the *simulated* mode involves a kind of movement between moments when illusions are successfully created, and moments when its effect falls back to producing a mere likeness of space. Such a characterisation falls in line with Myatt’s description (1998, p.91) of his experience of spatial music as discussed in Chapter 1.

The approach of categorising spatial concerns by spatialisation technique follows a similar approach to that discussed by Natasha Barrett (2002). The framework presented here is a little broader and differs in its grouping. Barrett’s outline of strategies does not explicitly consider the conscious use of the spatiality of the listening space, here labelled the *intrinsic* mode, as a spatialisation technique. She does acknowledge its importance in tempering the sounds produced in the other categories: “The space in which the music is heard and the way it is performed have important consequences for the spatial aspects in the music” (p.321). Both Barrett’s *Spatial Illusion* and *Spatial Allusion* are here considered *referential*. *Spatial Illusion* refers to perceptions of space that appear real without actually being spatially veridical. Such spatiality is here categorised as *referential iconic*; it is its similarity with reality, without actually creating a successful illusion that distinguishes it from the *simulated* mode. Barrett’s *Spatial Allusion* is here categorised as *referential index* and *referential symbol*. Both of these types of the *referential* mode contain cognitive elements that are dependent on the recognition of sounds heard to represent spatial properties.

Where the ideas proposed here differ from Barrett’s approach is in the conception of the role of the different spatialisation modes within a compositional intent. This thesis is concerned with identifying the spatial attributes exposed specifically by reality-equivalent technologies, and how they might relate to compositional intents. Barrett does not go to such detail; she sees the simulation of 3D soundfields as an addition to the toolset of spatial composers and considers its use as complementary to other spatialisation techniques:
However, by including aspects of allusion and the advantages of the 3D synthesised sound field, the composer’s spatial options begin to increase. Musical examples from the repertoire where spatial allusions are important to the music can be found. However, this paper presents a closer look at how these devices can be used as a primary and not subsidiary part of the sound and compositional development. (2002, p.322)

Barrett states the need to explore in greater depth how such things as the simulated mode can serve a compositional intent, rather than just act as secondary concern. However, she does not discuss in any great detail how any specific attributes, such as the quality of realistic representation or Rumsey’s ‘immersion attributes’, might serve compositional interest.

The aesthetic implications of the compositional engagement with the different modes can be illustrated by turning to some of Rumsey’s analysis on data collected on the subjective interpretation of spatial quality:

[Room] size and reverberant level do not appear to require a strong sense of presence to judge them, whereas envelopment requires a strong sense of presence (although it may be the other way around). Also the ability to judge room width (in this paper’s terms, environment width) requires a degree of presence. The factor 2 (presence) attributes appeared to be strongly dependent on surround modes of reproduction, whereas a number of the room acoustics attributes loading factor 1 [which concerns room size and reverberant level] could be judged even using mono reproduction. (2002, pp.663-664)

In this quote, ‘factor 1’ and ‘factor 2’ relate to spatial properties extracted from a data set using a technique that allows identifying strong relationships between measured attributes. Rumsey has identified some relationships that effectively illustrate some of the differences of the spatial encoding modes. Our ability to judge the size of a room and its reverberant level do not require a high degree of presence in the space. This is already intuited by composers who know that the application of a single channel of reverberation to a sound will help communicate the size of the room that sound is in, without simulating any spatial experience. Rumsey is
essentially identifying a difference between the *referential icon* mode, and the *simulated* mode. Whilst room size can be encoded using the *referential icon* mode, the sense of presence is strongly dependent on a spatialisation technique that has the ability to envelop the listener, such as the *simulated* mode. A characteristic of the *simulated* mode is the reinforcement of the spatial attributes classed, by Rumsey, as ‘immersion attributes’. Thus the composer’s choice of the *referential* mode, to create a sense of the size of a room, will not result in a sense of presence. The choice of the *simulated* mode, however, can result in a strong sense of presence. In other words, different spatialisation techniques expose different spatial characteristics. The *simulated* mode introduces a new class of attributes that is largely unconsidered within the discussions surrounding spatial compositional strategies.

3.4 A detailed example of encoding modes
Consider a composition in which a recording of a trumpet is panned within a reverberant space such as a church. The composition is realised within the *simulated* mode. In other words, it seeks to convince that the spatial illusion created is real. Software is used to model the acoustic of the church and also pan the trumpet recording within it. The trumpet plays a very simple score, illustrated in Figure 18, which essentially consists of a single note incrementally increasing in loudness.

![Figure 18. A simple trumpet score, in which only the loudness changes.](image)

Once the completed work is performed in a concert hall, listeners will be exposed to spatial information present in all three of the spatial encoding modes. As the sound travels from the multiple speakers surrounding the audience to their ears, it will collect the spatiality that is *intrinsic* to the listening space: the present listening space always imposes its acoustic character to whatever is projected within it. The illusion
of a reverberant church is attempted in the *simulated* mode; this reverberation may sound very different to the acoustic of the concert hall. The recording of the trumpet may also contain spatial information. If recorded within a small room, spatial artefacts such as early the reflections of the trumpet’s sound bouncing off surrounding walls may have been captured. These spatial artefacts are the result of the *intrinsic* mode associated to the recording room. Of course, a properly damped recording room may minimise evidence of the *intrinsic* mode present during recording. If, however, the room was not properly acoustically managed, then the acoustic character of a small room would be present and perceivable within the sound of the trumpet. Once captured on tape and played back within on reality-equivalent systems, this spatial information may be perceived. This is the encoding of spatial information in the *referential icon* mode. The spatial information contained within the recording is the result of the activity of the *intrinsic* mode, but once heard, all that is left is a likeness to the sound of a small room.

Lastly, the musical figure performed by the trumpet also has the capacity to describe spatial attributes. The capacity for musical figures to reference attributes of space is explored in detail in Chapter 4. The trumpet’s score includes two qualities of sound that are characteristic of an object moving towards the listener. The first is an increase in loudness, the second is a change in the trumpet’s spectral profile; that is, the balance of its component frequencies. As the trumpet becomes louder, its timbre changes and overtones become more apparent in the sound. This change in spectral character broadly reflects one aspect of how sounds change as they move closer to the listener. The presence of humidity in the air causes the attenuation of high frequencies (Blauert 1997, pp.126-127). As sounds move closer towards the listener the attenuation diminishes and high frequencies become more apparent. Whilst the combination of changes in loudness and changes in spectral balance, as created by a trumpet, may be only a rough approximation of reality, it has the capacity to *reference* distance and movement. As such, this simple musical figure contributes to the array of spatial information potentially perceived. It describes movement towards the listener.
Once analysed into encoding modes, this relatively simple example of a recorded trumpet panned in a virtual acoustic space reveals many sources of spatial information. Figure 19 illustrates how the activity of each mode contributes to the final listening experience. The composer’s spatial design might essentially concern the spatial simulation of the church, but the spatial information received by the audience includes information present in three other sources; the acoustic of the performance hall, the recording room of the trumpet, and the musical figure performed by the trumpet. This example challenges the conception of the spatialisation of sound as the process of imbibing spatial characteristics. A plurality of spatial information may be present within recorded sounds, within musical constructs, and (as mentioned earlier) within the timbral properties of synthesised sounds. It is perhaps better to describe space as something that is constantly collected by sound, rather than bestowed upon it. The analysis of spatial music by spatial encoding modes thus suggests that the composer might better consider spatial information in sound as something that needs clarification, rather than application.

The notion of the ubiquity of space within sound is well captured by Emmerson in a post-script to a journal article entitled “Acoustic/electroacoustic: The relationship with instruments” (1998b). Emmerson uses the term ‘spatial surrogacy’ to label qualities of space that arise as a secondary effect of other processes involved in creating and presenting music:
But here we are referring to our apprehension of the 'space' around the sound, implied in its recording, processing and projection; it is this which may carry anecdotal or narrative reference rather than the instrumental sound itself. (pp.161-162)

Emmerson is highlighting how spatial qualities inadvertently captured or created in music can carry significant import to the composition. Considered within spatial music, and the nature of our perceptual limitations already identified, these spatial qualities present a challenge to the composer. One of the primary aims of the proposed framework presented here is the identification of what Emmerson has labelled spatial surrogacy. Emmerson’s projection, in the previous quotation, is analogous to the *intrinsic* mode; recording and processing create perceptions of space in the *referential* mode. Emmerson (p.162) closes his discussion by mentioning techniques that are analogous to the *simulated* mode: he states that “the increasing availability of realistic spatial effects and simulated movement and positioning capabilities in sound processing facilities […] may bear fruit”. Of course, this last statement is particularly potent within the concerns of this thesis since the research question is essentially motivated by the desire to understand how spatial verisimilitude might ‘bear fruit’ within a compositional context. The framework presented here takes a step in that direction by introducing a way of understanding the relationship between music and space that clarifies a range of issues introduced by the simulation of spatial sound.

### 3.5 Spatial encoding modes illustrated by examples

The validity of the framework is now illustrated by examining some historical examples of spatial music. Since the modes categorise spatialisation *techniques* rather than *technologies*, their application to pre-electronic music remains valid. In so far as technologies encapsulate specific techniques, the encoding modes also resonate with technologically dependent genres of spatial music. Whilst spatial works tend to predominantly engage in one mode, the other modes will often be present. Examples of the use of multiple modes are highlighted where appropriate.

The *intrinsic* mode, in which the spatiality of sounds is created by the current physical space, is characterised by an implicitly limited palette of spatial gestures.
Sounds are limited in their position and movement, and in their acoustic treatment. The most accessible and controllable spatial attribute is that of location. Whilst limited in scope, the use of the performance venue’s intrinsic spatiality, in combination with musical constructs, has been exploited to realise a large range of spatial gestures.

In outlining the antecedents of the use of location in spatial music Worner (1977, p.156) identifies the compositional exploitation of the spatial division of performers as an important technique. The exploration of the musical possibilities of opposed organs and their associated vocal and instrumental choirs, is a technique typically associated with the Venetian school of the late 15th and early 16th centuries. This practice, known as cori spezzati (separated choirs) or polychoral technique, involves the encoding of location through the intrinsic spatiality of the performance space. Despite predating the technological engagement with space in music, cori spezzati is a practice that is inherently concerned with the spatial dimension of composition. Indeed, in his text Cori Spezzati (1988), Carver introduces his subject by noting that the “continuing interest of contemporary composers in the compositional possibilities of space” appropriates the celebration of “one of the glories of the late Renaissance and early Baroque music – the polychoral style” (1988, p.ix). Strictly speaking, polychoral music predates the Venetian school (p.ix) but Carver cites the work of Venetian composer Giovanni Gabrieli as characteristic of polychoral style (1988, p.144). In Omnes Gentes (1587 or 1597) Gabrieli uses four spatially separated choirs, occasionally paired, to create “a particularly splendid effect” (Carver 1988, p.157). The practice of spatial antiphony eventually developed into some innovative acoustic techniques such as the creation of echo effects, achieved through the carefully timed orchestration of opposing choirs (Begault 2000, p.195). Begault identifies echo music as belonging to a broader concern with the musical “suggestion of environments and distant sound sources” (2000, p.194). The creation of echo effects introduces a second dimension to the intrinsic spatiality of the performance space. These effects depend on the careful orchestration of temporal delays implemented within musical expression. They can thus be understood as attempting the simulation of a very large space since the technique used attempts to replicate an aspect of the auditory stimuli experienced when hearing an echo. Of course, whether the spatial
illusion is sufficiently convincing to be accepted as real is questionable. The resultant spatiality might thus be more accurately considered as belonging to the referential mode, in which a spatial perception arises out of hearing similarity with the sound of other spaces. What this example demonstrates, however, is that multiple modes can both coexist and cooperate to encode a singular spatial attribute: the echo effects depend on both spatial separation and musical orchestration. As expressed by Worner, in the hands of the Venetian school “the deployment of acoustical space became an art in itself” (1977, p.156).

The practice of spatial antiphony is certainly not limited to 16th century music: Mozart’s *Don Giovanni* (1787) includes passages for three orchestras, one in the orchestra pit, another on stage and a third backstage (Brant, 1967, p.235); and Berlioz *Requiem* (1837) employs four brass ensembles placed off-stage. The 20th century saw a renewed interest in the explicit exploration of the spatiality intrinsic to performance spaces (Zvonar 2004b, p.1). In *The Unanswered Question* (1908) the American experimentalist Charles Ives explored the layering of musical forms by strategically positioning strings offstage and trumpet soloist and woodwind ensemble on-stage. Later, the work of fellow American Henry Brant continued Ives’ engagement with intrinsic spatial encoding. Brant’s *Antiphony 1* (1953), introduced in Chapter 1, employs five groups of instrumentalists each playing in different keys performing at distance to each other. Over a decade after composing *Antiphony 1* Brant wrote an article titled “Space as an Essential Aspect of Music Composition” (1967). In this article he describes, in detail, his experiences and observations on the compositional exploration of spatially separated performers. Amongst the many informative insights is a discussion on the impact of the performance hall size in determining the optimal separation between performers (1967, pp.228-230). Another discussion considers “the compelling naturalness of effect” when pitches correlate to the physical height of their source; for example, a performer positioned high on a balcony projecting high pitches (p.232). Here, Brant describes the perceptual correlation of a singular attribute, height, which is encoded in both the intrinsic mode, and the referential mode. The interpretation of pitch as referring to physical height can be considered spatial encoding in the referential mode, and this is
examined in detail in Chapter 4. Broadly speaking Brant’s paper can be considered an important exploration of the compositional possibilities of the *intrinsic* mode.

Other 20th century composers extended Brant’s engagement with the *intrinsic* mode. Iannis Xenakis’ *Eonta* (1963-1964) for piano and brass involves a range of spatial gestures created by the instrumentalists’ movements (Harley 1994b, p.295). The slow axial rotation of the brass instruments is of particular interest here. By slowly turning to the left or to the right whilst performing a chord, the brass players create subtle variations in timbre. Musical instruments emit different qualities of sound in different directions. The rotational movement thus introduces variations in timbre that is further modulated by differences in reflections on walls as the instrument’s principle direction of projection changes. Here Xenakis engages with the *intrinsic* mode by exploring how the *spatiality of instruments* interacts with the *spatiality of the performance space*. *Eonta* expresses Xenakis’ understanding, whether intuitive or otherwise, of the significance of directivity in instrumental projection and how it affects spatial perception. *Eonta* is an example of a work in which the composer’s spatial thinking has become aligned with the opportunities afforded by a particular spatial encoding mode.

The importance of the spatiality of the *intrinsic* mode is not lost to electroacoustic composers who use technology to simulate and reference space. Many electroacoustic composers have specifically discussed the challenge of adapting work, often developed in a studio environment, to suit the specific acoustic of the performance space (Truax 1998, p.145; Smalley in Austin 2000, p.11; Barrett in Otondo 2007, p.10). One of the fundamental challenges with developing works in a studio environment is that spatial design is entirely divorced from the effects of the *intrinsic* mode. When performed, a studio-developed work will almost always be subject to a different acoustic.

An interesting example of the exploration of the *intrinsic* mode within electroacoustic music is Alvin Lucier’s *I am sitting in a Room* (1969). In this work, Lucier captures the acoustic treatment of the current listening space by recording a sound coming out of a loudspeaker. He then iteratively re-plays and re-records the
sound, thus compositing the effect of the *intrinsic* mode many times over. Lucier’s interest is in drawing the listener’s attention to how physical spaces affect sounds (Wishart 1986, p.52). One might say that Lucier’s interest is in highlighting the *intrinsic* mode.

The *referential* mode is discussed in detail in Chapter 4. Since perceptions of space arising from musical figures are considered to be largely referential, the question of the relationship between the *referential* mode and spatial verisimilitude is critical to this thesis and thus is allocated its own chapter.

What is worth noting here, however, is a particular genre of spatial music that straddles both the *referential* mode and the *intrinsic* mode. Diffusion, as it is called, refers to “the distribution of the (usually stereo) sound in a space through the use of a mixer and multiple loudspeakers” (Truax 1998, p.141). Diffusion can be categorised *intrinsic* since it allows “tailoring the spatialisation to work within the space of the concert hall” (Barrett in Otondo 2007, p.12). Indeed, Smalley goes so far as to describe it as “the ‘sonorising’ of the acoustic space” (Austin 2000, p.10). Diffusion can also be categorised as *referential*; firstly because the use of recognised sounds bring associations (*referential index*) but also because manipulations of the sound can affect perceptions of space (*referential icon*). Smalley (cited in Austin 2000, p.14) explains, within the context of diffusion practice, that “quite often composed space is created through artefacts or spatial by-products of the sounds, textures, and processing techniques” used. Smalley draws a specific distinction between the practice of diffusion and the *intrinsic* mode in that he argues that electroacoustic music is never designed for a specific acoustic (Austin 2000, p.11). This is in direct opposition to works of the afore mentioned *cori spezzati*. A significant component of diffusion practice involves adapting works, often composed in the studio, to fit the acoustics of diverse performance spaces. As such, diffusion has an inherent consciousness of spatial encoding native to the *intrinsic* mode.

Francois Bayle, a French acousmatic composer who, in 1974, founded a sound diffusion system known as the Acousmonium, identifies this same distinction between the spatiality present within the musical content and the spatiality of the performance space. He calls them the “internal space” and the “external space”
These two terms broadly correlate with the *referential* mode and the *intrinsic* mode. Bayle identifies the effects of the “external space” as “often undesirable or from time to time exploited” (p.243). Within Bayle’s account, there is evidence of a lesser embrace of the effects of the *intrinsic* mode. Indeed, Bayle considers the act of diffusion as consisting of the faithful adaptation (p.243) of the space already conjured in the musical material, to the projection space. This differs slightly with Smalley’s approach, mentioned above, that emphasises the importance of respecting the acoustic contribution of the *intrinsic* mode.

Bayle’s characterisation of the acoustic characteristics of the performance space as ‘often undesirable’ is an appropriate introduction to the *simulated* mode. The techniques used by reality-equivalent technologies mostly assume the non-existence of the *intrinsic* mode. In the *simulated* mode, the effect of physical space on sound is manufactured; it has no relation to the present listening space. Indeed, as is discussed in Chapter 5, one key attraction to the *simulated* mode is precisely the ability to escape from the present listening space. The extent to which the *simulated* mode can successfully create the impression of acoustic characteristics that are different to the current listening space is a complex matter that concerns its practical implementation. However, what these challenges highlight is that whilst diffusion practice is perhaps far more limited in its spatial repertoire, by acknowledging and incorporating the effects of the *intrinsic* mode it also avoids a struggle with it. The importance of this distinction is highlighted by Dow who says: “Diffusion, even in these heady days of refined multi-channel loudspeaker systems, still has an important role as mediator between artistic inception and reception” (2004, p.26)

Significant engagement with the *simulated* mode is only beginning to receive compositional attention. It is a form of spatialisation technique that has attracted limited critical attention from composers. Two notable contributors are Ambrose Field (cited in Austin 2001) and Natasha Barrett (2002; cited in Otondo 2007), who have each discussed the incorporation of ambisonic techniques within their compositional practice. Barrett’s approach, in which the simulation mode is characterised as increasing the “composer’s spatial options” (2002, p.322) is discussed earlier in this chapter.
Field takes a critical approach to reality-equivalent technologies. He recognises the importance of parsing out any spatial information present in a sound recording prior to its spatialisation (2001, p.2461); this is the recognition of spatial information present in the referential mode. He (cited in Austin 2001, p.24) argues that whilst ambisonics depends on the careful placement of speakers within the concert venue, it allows for the design of spatial gestures more precisely than is possible with diffusion practice. This reflection suggest one way that reality-equivalent technologies might mediate compositional conception: the ability to achieve greater accuracy in spatial gesture might engender more, and subtler, spatial articulation. In an interview with Larry Austin, Field discusses another characteristic; the quality of ‘realism’ engendered by ambisonic techniques:

Austin: That was what was distracting me. I was disconcerted when I first heard ambisonics, because it was too ‘‘real.’’
Field: This is a big issue. ‘‘Too real.’’ I absolutely agree with you. I would hate to take away the act of people’s engaging their imaginations with a sound or a piece or whatever from any form of composition. And you might argue—that’s the basis of what we said earlier about people in the audience—that we have to learn to decode sound diffusion. So we’re left with an ambisonic reality where everything has an accurate physical space, and so on. But that doesn’t mean, as a composer, you stop there. You think of other ways that you can allow the audience to imagine things—how you can transport them to other spaces that might not exist. Now, that’s powerful. If you can make a physical space which, at the same time, encourages a depth of imagination in the listener, then you’re really getting somewhere. There is a problem there with the reality aspect of it. Reality is always a problem, though, isn’t it? (Field in Austin 2001, p.28)

Within this discussion is the acknowledgement of a very different type of spatial attribute exposed by reality-equivalent technologies. ‘Reality’ is a quality that introduces broad and complex dimensions to compositional thinking. These dimensions and their implications for spatial composition form one of the most important components of the research question, and are explored in depth in Chapter 5.
Field goes on to draw an interesting parallel with ‘reality equivalence’ in other art forms. Contemporary UK artist Damien Hirst, he argues, engages audiences through such works as *Away from the Flock (a sheep in a tank of formaldehyde)* (1994) by making ‘reality more real’. The question of how ‘reality’ can contribute to musical meaning, which is often abstract and ephemeral, is perhaps more difficult to consider. Field alludes to the importance of employing realism as a way to draw and direct the listener’s imagination to other things. In so saying, he is alluding to the semiotic dimension of sound, in which what is heard acts as a *sign* to other meanings. In the next chapter, I consider music’s capacity to create perceptions of space through signs. As the discussion develops, reality’s capacity to act as signs that contribute to musical meaning more broadly, is also examined.
Air-bound, above and stretching out beyond the plane tree to the right, is the swifts’ zone, their high-frequency calls delineating the canopy of the soundscape’s space. […] For a moment I wonder what would happen to the swifts’ space were I to record it and bring it indoors for loudspeaker listening. The recorded image could not reproduce the spatial elevation, but I would nevertheless deduce it: in acousmatic music actual spatial localisation is not essential to create elevation. (Smalley 2007, p.36)

Smalley, a prominent composer and commentator on electroacoustic composition, is making the point that sound has the ability to imply spatial qualities by association. Using qualities of space as a compositional device does not require the physical simulation of sound in space. Nor does one need to position a speaker at height to create the impression of elevation. Instead, simply referencing a sounding object that is typically heard in elevation can imply height.

This example illustrates the signification of spatial qualities by reference. It is an example of the referential mode, introduced in the previous chapter. Through the identification of the object creating the sound, in this case a bird, associations can be drawn to certain spatial qualities, such as height. However, the referential mode is not limited to this particular mechanism. Indeed, one of the key points introduced in this chapter concerns how spatial qualities can be referenced without the intermediary of an identified sounding object.

A single channel recording of a child running in a reverberant church will reveal something about the size of the physical space and may even reveal the type of space itself, a church. It is not associations with the identified sounding object, a running child, that suggests a large reverberant indoor space. Instead, there is something recognisable about how the sound has presented itself. Listening to a recording of a church is far from the experience of being in a church, yet the spatial qualities perceived are consistent with that experience. Here, space is referenced through recognisable artefacts present within the sounds heard. What is important is that this
mechanism remains active even when the sounding object is not identifiable, as is the case in much electroacoustic music.

The notion of referenced space plays an essential role in the exploration of the research question. The use of reality-equivalent technologies does not preclude the activity of spatial references. As such, spatial design must consider the presence of referenced space, which may coexist and perhaps conflict with or contradict simulated space. This chapter is primarily concerned with the identification of spatial references present in sound and music. As introduced, there are different mechanisms by which space can be referenced and these have musical implications. The identification of spatial references thus requires a structured approach, in which the referential mechanism can be understood. Here, a framework need not be devised since this subject is precisely the concern of semiotics; the study of the interpretation of signs and symbols. This chapter exposes an understanding of sound’s capacity to reference space through the application of a semiotic framework developed by 19th century logician Charles Peirce.

The application of Peirce’s framework to spatial audio underlies some important insights. It is shown that sound’s capacity to reference space is dominated by a referential mechanism whose activity is extensive: it occurs easily and often. The iconic mechanism, defined and described below, is largely responsible for the presence of perceptions of space in sounds that host no explicit spatial design by the composer, and may have no associations to identifiable sounding objects. It is shown to be present in sounds, processing techniques and perhaps most importantly, in common musical constructs. Examples of the ways that spatial qualities can be perceived in ‘non-spatial’ music are provided.

Within this line of enquiry the early writings of Pierre Schaeffer, pioneer and inventor of musique concrète, are considered. Schaeffer’s published research is deeply concerned with sound’s referential capacities as they manifest within a musical context. His text Traité des objets musicaux: essai interdisciplines (1966) is examined for evidence of the distinction between references caused by the recognition of sounding objects, and references caused by the recognition of spatial
effect artefacts. The distinction is not found, but an examination of Schaeffer’s understanding of the semiotics of sound raises a key and critical question for spatial music. As is elaborated later in this chapter, Schaeffer argues that musical meaning is derived from sound’s capacity to reference abstract form over and above the referencing of real-world physicality. As such, for Schaeffer, whether real-world physicality is referenced by the intermediary of a sounding object or otherwise is immaterial; what is of importance is that the perception of this real-world physicality is suppressed to allow the emergence of abstract form. Such an understanding of music highlights the following question: what is the relationship between spatial verisimilitude and abstract form? Thus, an understanding of the semiotic dimension of sound begins to project insights onto the broader question of how spatial verisimilitude might contribute to musical meaning.

4.1 Peirce’s theory of signs

Different semiotic frameworks use different terms to broadly indicate similar concepts. In this thesis the word ‘reference’ is used to describe a sound that results in the perception of spatial information outside of the veridical experience of sound in space. In his discussions on semiotics Charles Peirce uses the word ‘sign’:

I define a Sign as anything which on the one hand is so determined by an Object and on the other hand so determines an idea in a person's mind, that this latter determination, which I term the Interpretant of the sign, is thereby mediately determined by that Object. (1998, p.482)

This triadic structure identifies not only the relationship between the sign and the thing it represents; the object, but also the mind’s interpretation of this object; the interpretant. Within the context of this thesis, the sign will typically be a sound but the spatial quality perceived may be either the object, or the interpretant. For example, the sound of a bird in flight is a sign that represents a bird and can be interpreted as, or in Peirce’s words ‘determines an idea’ of, height. The sound of reverberation, however, is a sign whose object is a spatial quality: that of a large space. In other words, the spatial quality of height is a secondary interpretation of the sound of a bird, whilst the spatial quality of a large space is a primary interpretation of the sound of reverberation. This subtle distinction, between the object and the
interpretant, of a sign, is a characteristic of Peirce’s approach that distinguishes it from other important semiotic frameworks such as that of Swiss linguist Ferdinand de Saussure. In his text *Music and Discourse: Toward a Semiology of Music* (1990), Nattiez argues that the concept of the interpretant is particularly relevant to music. He (1990, p.5) explains that Saussure’s definition of the sign depends on a “static” relationship between what Saussure calls the ‘signifier’ and the ‘signified’. This duality is perhaps more relevant to language but Nattiez (p.5) describes it as uncharacteristic of music. Peirce’s definition differs in that it caters for the role of the interpretant that holds an iterative dimension: the interpretant can itself become a new sign pointing to further interpretants. For example: the sound of a bird has the possible interpretant of height, which can itself become a sign that holds the interpretant of falling. This iterative dimension, in which interpretants are concatenated, is important for composers generally and spatial composers specifically. The spatial composer needs to consider not just how space is signified, but also how that space signifies other things that might contribute to the expression of the composition. A reverberant space may, for example, signify a public space, which may in turn signify a loss of privacy and so on.

Peirce’s writings explore different ways to categorise signs. The categorisation, or typology, mainly used here is that which Peirce describes as the ‘most frequently useful’:

[I] had observed that the most frequently useful division of signs is by trichotomy into firstly Likenesses, or, as I prefer to say, *Icons*, which serve to represent their objects only in so far as they resemble them in themselves; secondly, *Indices*, which represent their objects independently of any resemblance to them, only by virtue of real connections with them, and thirdly *Symbols*, which represent their objects, independently alike of any resemblance or any real connection, because dispositions or factitious habits of their interpreters insure their being so understood. (1998, p.460)

This is a typology that is concerned with the mechanism by which a sign points to its object. The use of a bird to sign the spatial quality of height, for example, employs just one of these mechanisms. The sound of a bird in flight can be described as an
index of height; it logically follows that the bird is at height when this sound is heard. There is nothing within the sound of a bird that is similar to the sound of height. The sign operates by virtue of a logical connection between birds in flight and height.

<table>
<thead>
<tr>
<th>Playback of bird recording</th>
<th>is an icon of</th>
<th>which is an index of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The sound of a real bird in flight</td>
<td>Height (a spatial quality)</td>
</tr>
<tr>
<td>Playback of reverberation recording</td>
<td>The sound of a real large physical space (a spatial quality)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 20. The difference in the referential mechanisms of: a recording of a bird and a recording of reverberation, in signifying spatial characteristics.

Conversely, a recording of a child running in a church contains an auditory signature of the physical space concerned. That auditory signature can be considered an icon of a large physical space. When a listener listens to this recording they hear similarity with the sound of large spaces. There is a likeness between the recorded sound and the sound of a large space. There is no logical association between the sound of a child running and large spaces. It does not follow that where there is a child running, there will be a large reverberant space. As an icon, the recorded reverberant sound is similar to sound remembered from the experience of large spaces.

The most important distinction between these two examples is that in the first example, a spatial quality is the interpretant of a bird in flight, and in the second example, a spatial quality is the object of reverberation. In other words, there is a more direct relationship between the sound of reverberation and spatial attributes than there is between the sound of a bird in flight and spatial attributes. Of course, both the sound of a bird and the sound of reverberation will also act as signs to other referents. For example, the sound of reverberation may act as a sign of cold temperatures, or as a sign of religious institutions. What is of concern here, however, are only signs whose referents are spatial attributes.
Peirce’s third type, the symbol, describes a connection between the sign and the object or interpretant that involves neither similarity nor logic, but rather is merely understood by ‘dispositions or factitious habit’ (p.460). One example of a symbol of space is the use of voice suggestions that can sometimes be found in demonstrations of binaural recordings, which are concerned with the realistic spatial representation of sounds auditioned on headphones. For example, in a demonstration of a binaural spatialisation technology (A Virtual Barber Shop, n.d.) a voice says “I’d like to start the demonstration by moving over to your right hand side and picking up this bag” (1’00”). The spatial interpretation of the speaker’s location is thus informed by a verbal description, which acts as a symbol of space. It is worth noting that the sound of hair clippers cutting hair, in this same demonstration, acts as an index of space. The identification of the specific sound of proximate hair clippers is logically associated to short distances from the head. Begault (2000, p.29) notes that binaural demonstrations often employ sounds that are typically heard close to the head, such as a cigarette lighter, drinking a glass of water or hair clippers. The success of these sounds, to create perceptions of space, is more a function of the cognitive dimension of perception, than the accurate spatial rendering of auditory stimuli (p.29).

Another example of the use of symbols in spatial audio involves non-auditory symbols pointing to spatial qualities. For example, in his summary of the historical use of space in music, Begault describes the use of written text to aid in the spatial interpretation of sound:

[Orchestral] music created aural virtual realities not only through distant, imitative sounds, but also by writing program music. For example, one doesn’t realize until they’ve read the program that one is hearing a “March to the Scaffold” in Hector Berlioz’s Symphonie Fantastique. (2000, p.195)

Here it is the text within the program notes that acts as a symbol whose object concerns spatial qualities that will be projected onto the music heard. The listener’s interpretation of the music is thus guided by what has been read. This is analogous to the French philosopher Roland Barthes’ notion of the role of text within advertising imagery. Barthes describes text as having one of two functions with respect to the
iconic messages in visual imagery: *anchorage* and *relay* (Barthe's essay translated to English in Trachtenberg 1980, p.274). The use of text to guide the listener’s interpretation is an example of Barthes’ notion of *anchorage*:

>[The] text *directs* the reader through the signifieds of the image, causing him to avoid some and receive others; by means of an often subtle *dispatching*, it remote-controls him towards a meaning chosen in advance. (Trachtenberg 1980, p.275)

Program notes allow the composer to ‘remote-control’ the listener’s spatial interpretation of the sounds heard.

Conversely, spatial audio signs can also act as symbols of non-spatial characteristics. For example, in filmic media the use of reverberation on voices can suggest the memory or recollection of past events. There is nothing about reverberation that necessarily equates to thinking about the past, nor does it follow that when one is thinking about the past there will be reverberation. Reverberation thus has the capacity to act as a symbol of the remembered past.

However, the way in which a spatial property might itself act as a sign to further meaning is not of direct concern here. What is of interest is the understanding of the different mechanisms by which perceptions of space may arise independently of the simulation of space. This understanding helps clarify the relationship between music and realistic spatial representation because it allows the identification of spatial properties potentially present in sound and music that have no simulated dimension.

Of course, Peirce’s understanding of the chaining of interpretants indicates that perceived spatial attributes will, in turn, act as signs to further meaning. How a specific perceived space might *qualitatively* contribute to the expression of a composition is not considered here. Such issues are left to the mind of the composer. However, what is of concern to the thesis is the more general question of the relationship between auditory illusion and abstract musical form. This question is later considered in the context of an examination of Pierre Schaeffer’s conception of musical meaning.
4.2 The Spatial Icon

The application of a semiotic framework to sound articulates that there are two principal mechanisms by which sound can reference space: as an index and as an icon. The first way involves associations drawn to the sounding objects identified; the second involves a likeness with the experience of sounds projected in space. The second might also be called the identification of a spatial modulation, where the manner in which space has affected the sounds heard is recognised. Symbols are excluded from this discussion because they are less relevant to musical contexts than they are, perhaps, to mediated environments more generally.

Symmetry exists between these two sign types and the physical nature of sounds. Since all energy, which includes sound, ‘has to be deployed in space’ (Levebvre cited in Smalley 2007, p.38), the modulation of space must exist in all sounds. In other words, a sounding object is always accompanied by a spatial modulation. As expressed by Blauert (1997, p.3) in his scientific text Spatial Hearing: ‘The concept of “spatial hearing” … is in fact tautological; there is no “nonspatial hearing”’. The presence of space in sound is perhaps most evident in sound recordings, in which the ear is metaphorically replaced by the microphone’s diaphragm:

> Usually, any sort of live recording will carry with it information about the overall acoustic properties of the environment in which it is recorded. (Wishart 1986, p.45)

Sounding object and spatial modulation can thus be seen as two sides of a coin. Where one exists, so must the other. Their capacity to reference space, however, can operate independently. The identification of a sounding object, irrespective of the spatial modulation it carries, can act as an index of spatial characteristics. A recording of a bird is such an example. Conversely the identification of a spatial modulation, irrespective of what the sounding object might be, can act as an icon to the experience of space.

Whilst this analogy expresses the implicit presence of spatial modulation in all sounds, the perception of the source cause of sounds holds significant importance in the literature of electroacoustic music. In the words of Pierre Schaeffer (1966, pp.93-
we are “instinctively, almost irresistibly inclined” to identify the source cause of sounds. However, in his text *Traité des Objets Musicaux* (1966) Schaeffer does not explicitly address our ability to identify spatial characteristics independently of the identification of a sounding object. Schaeffer’s writing discusses the apprehension of the sounding object, and this focus retains an importance even in spatial composition. As stated by Smalley (2007, p.37) “Identified sources carry their space with them”. What Schaeffer does not specify is that non-identified sources also have the capacity to carry their space with them. By hosting a recognisable spatial modulation that acts as an icon of the sound of particular spaces even non-identifiable sounds can lead to the identification of physical spaces.

Our ability to identify the spatial modulation present in sound, with some exceptions, has attracted little critical attention. Smalley (1997, p.122) uses the term *spatiomorphology* to describe the subset of ‘spectromorphologies’ which contribute to spatial perception. Smalley argues that “Space, heard through spectromorphology, becomes a new type of ‘source’ bonding” (1997, p.122). Barrett (2002, p.314) uses the term *spatial-illusion* to describe a similar concept. Doornbusch (2004) again makes a similar analogy and labels it *space-identification*, but also distinguishes *setting-identification* in which a place is recognised independently of the recognition of its physical spatial attributes.

Here, a distinction needs to be drawn between referencing the spatial modulation of sounds and simulating the spatial modulation of sounds. The distinction is identical to the difference between the referential mode and the simulated mode articulated in the previous chapter. Applying an artificial reverberation algorithm on a sound will result in the creation of a referential icon of a large space. Applying an artificial

---

13 Schaeffer is aware that recordings capture the spatial modulation of sounds. He discusses how the use of microphones results in the collapse of the multi-dimensional acoustic space into one or two channels (1966, p.77); observing that this amplifies the perceived volume of the reverberant field, and also highlights the perception of background ambient sounds (pp.78-79). He touches on the spatial aspects of sounds throughout the book. For example, he briefly discusses the perceptions of space made possible by binaural hearing as opposed to the monaurality imposed by microphones (pp.213-214).

14 Spectromorphology is a term coined by Smalley concerned with how a sound’s spectral profile changes over time.
reverberation algorithm on sound that is projected through multiple speakers in such a way so as to attempt to convince the listener that they might actually be within that large space, is the creation of a simulated spatial quality. One seeks to create a recognisable likeness of the sound of a large space, the other seeks to create a suspension of disbelief.

A spatial icon can be seen as an approximation of simulated space. The distinction is well expressed by Barrett (2002, p.315) who talks about the robustness of what she calls ‘spatial illusion’; a term which is analogous to ‘spatial icon’ as discussed here (see Chapter 3 for a clarification of these terms):

A spatial illusion does not need to satisfy all of the real spatial laws. In fact it is interesting how much one can get away with before the perception of a spatial illusion breaks down. A primitive reverberation effect can provide the illusion of a spatial enclosure [...] We know the illusion does not exactly resemble a real-world image, but nevertheless accept the information as a good enough approximation. (p.315)

‘A good enough approximation’ of spatial modulation is an apt description of the spatial icon. Conversely, simulated space needs to ‘mimic reality such that departure from reality is below our perceptual threshold’ (p.317). As expressed by Barrett, it is easy to create credible approximations of spatial modulations. It is much more difficult to create realistic spatial illusions.

The distinction between simulated space and spatial icons is often not clearly distinguished in the discourse on spatial music. This lack of distinction obfuscates the ease with which spatial icons come to exist in sounds. In Wishart’s discussion on the nature of the perceived acoustic space, for example, he expresses that “Such real, or apparently real, acoustic spaces may be re-created in the studio” (1986, pp.45-46). Wishart does not articulate what might constitute the difference between a ‘real’, an ‘apparently real’, or even the mere suggestion of an acoustic space. He cites, as an example, the creation of spatial depth “by using signals of smaller amplitude, with their high frequencies rolled off” (p.45). Whilst this technique involves the separation of a sound and its spatialisation, which is characteristic of the simulated
mode, it would here be classified as a spatial icon. The two audio cues modelled represent only a fraction of the five or more (Malham 2001a, pp.33-34) required to mimic reality, but would be largely sufficient to create an impression of distance; a spatial icon.

Spatial icon and simulated space can be seen as two ends of a discrete spectrum, divided by a line which marks the beginning of the suspension of disbelief. Simulated space aims to achieve illusion, whereas spatial icons only aim to produce a sufficient likeness to the real experience of sound in space.

For Peirce the distinction between the different sign types is a matter of definition. For the electroacoustic composer working with space, the differences between icons and indices of space hold some fundamental implications. Whilst the recognition of spatial modulation has attracted less critical attention than the recognition of sounding object, it has some important properties that necessitate careful consideration within spatial composition. As is elaborated below, spatial icons are robust, ubiquitous, are easy to create, will occur accidently, and they can be applied to abstract sounds.

Since the iconic mechanism for referencing space does not require the identification of the sounding object, abstract sounds are not precluded from referencing spatial characteristics (Barrett 2002, pp.314-315). But perhaps a more important attribute of spatial modulation concerns the facility with which spatial icons come to exist. Single channel reverberation is one well-known example of a spatial icon, but sound’s capacity to act as an icon of space goes much further than reverberation. As expressed by Smalley, spatial icons lie in the ‘sounds, textures, and processing techniques’ used by electroacoustic composers:

[Quite] often composed space is created through artifacts or spatial byproducts of the sounds, textures, and processing techniques you are using. For example, delays, phase changes, pitch offsets, or accumulation processes that give the sound more space also give it more depth. You can’t necessarily simply say “Well, I want it to have less depth or more depth,” because you’ll be affecting the actual music, the spectromorphologies, not just the spatial setting. So, one has a
limited power to control space as such, because the perception of space is the sum of many interrelated features. (Smalley in Austin 2000, p.14)

For Smalley electroacoustic music and space are fundamentally inseparable. They are inseparable because spatial icons are to be found in both the sound material and how it is processed. This view, in which space is ubiquitous in sound, challenges the creation of spatial music via the simulated mode. The simulated mode depends on a logical separation between a sound and its spatialisation. If, as Smalley suggests, spatial iconography is implicit and unreserved in sound then its subsequent spatialisation will be open to conflict and contradiction with those pre-existent spatial icons.

Spatial icons are thus worthy of significant consideration by the composer working with spatial verisimilitude. Whilst identified sounding objects also have the capacity to bring about perceptions of space through the indexical mechanism, their presence is more overt. Spatial icons, however, can be inadvertently created through the manipulation of sounds prior to spatialisation. Their presence confirms that the pursuit of verisimilitude cannot be considered independently of the sound material that is spatially projected.

4.3 Schaeffer’s semiotics and musical meaning
Within the discourse of electroacoustic music, the importance afforded to the recognition of the sounding object overshadows the recognition of the mutual and implicit spatial modulation. However, the examination of some of the early writings and ideas of Pierre Schaeffer indicate that this overshadowing may be incidental more than intentional. The below discussion identifies that Schaeffer’s concern with the identification of the source-cause was secondary to his interest in suppressing it. In Traité des Objets Musicaux (1966) Schaeffer argues for the disavowal of external references so that sound’s “intrinsic acoustic properties” (Wishart 1986, p.41) may be exposed. He advocates the denial of musical “instruments and cultural conditioning” (1966, p.98, my translation) thereby allowing sounds to present themselves as pure musical possibility. Whilst Schaeffer’s language indicates a concern with the identification of the sounding object, his real agenda is broader:
I listen to the event, I seek to identify the sound source: “What is it? What has just happened?” I do not stop at what I have just perceived, instead I use it. I treat the sound as an *index* that signals something to me. (1966, p.114, my translation)

Schaeffer specifically mentions the ‘sound source’ but his precise concern is that sound will act as an ‘index’ that ‘signals something’. Whether what is signalled is the recognition of a sounding object, or a spatial characteristic born of a spatial icon, becomes less important within his wider argument.

Schaeffer’s use of the word index, the French ‘indice’, also sheds light on his understanding of sound’s referential capabilities. Here, the word ‘index’ is more consistent with the broader meaning of the Peircian ‘sign’, since there is no specification of the referencing mechanism. However, Schaeffer also uses the French word ‘signe’ which translates literally as the English ‘sign’ but takes on an entirely different meaning within his writing. In the above passage he expresses that the listener can treat a sound as an ‘index’ (a Peircian sign) that references certain things, and in the below passage he describes how a sound can also act as a ‘sign’:

[1] can treat the sound as a *sign* that introduces me to a certain realm of qualities, and I can open myself to its meaning. A typical example is, of course, speech. This concerns semantic listening, aligned with semantic signs. Among the diverse possible “signing” listening modes we are, of course, specifically concerned with musical listening, which leans on musical qualities and provides access to musical meaning. (1966, pp.115-116, my translation)

Schaeffer thus distinguishes between sound’s capacity to reference everyday things, and sound’s capacity to reference things that contribute to musical meaning. He further clarifies this distinction by naming the latter ‘abstract signs’ where the ‘abstract is in opposition to the concrete materiality’ (1966, p.117, my translation) signed in the former. Thus, for Schaeffer, musical meaning is born of signs that reference the abstract, as opposed to real-world physicality. Underlying the apparent predominance of source-cause identification, in Schaeffer’s writing, is not the suggestion that sounding object is more important than spatial modulation, but rather
that the referencing of concrete and real-world material things sits apart from the 'abstract signs’ that provide musical meaning.

The notion that it is abstract signs that bring musical meaning is challenging to the pursuit of spatial verisimilitude. Indeed, abstract signs are effectively the inverse of realistic representation. Schaeffer seeks to rid sound of any semiotic association with real-world physicality such that sound's intrinsic characteristics can stand alone within abstract musical form. By seeking out realistic representation within a musical context, the composer seeks to re-instate real-world physicality, thus moving in the opposite direction to Schaeffer. In the pursuit of spatial verisimilitude the perception of sound’s intrinsic characteristics, which Schaeffer argues contributes to musical meaning, is thus demoted in favour of spatial illusion. In other words, by Schaeffer’s understanding of music, seeking out realistic representation may result in the obstruction of musical meaning. Of course, Schaeffer’s ideas on musical meaning are not incontestable, and further insights into the relationship between abstract signs and real-world physicality in forming musical meaning are discussed again shortly.

A proper investigation of how Schaeffer’s semiotics might be informed by Peirce’s semiotics is beyond the scope of this thesis. However, Schaeffer’s suggestion that the referencing of external physical things is distinct from the abstract signs that contribute to musical meaning is worthy of a little more consideration. ‘Acousmatic’ sound, which is at the source of Schaeffer’s ideas, derives its meaning from ancient Greek tradition. Schaeffer provides the Larousse Dictionary’s definition of the noun form of ‘acousmatic’

Name given to Pythagoras’ disciples who, for a period of five years, listened to his teachings hidden behind a curtain, unable to see him, and engaged in the strictest silence. (1966, p.91, my translation)

For Schaeffer, the ‘acousmatic’ listening situation illustrates that through isolation the ‘perceptual reality’ of sounds is exposed (1966, p.91). The Pythagorean veil is much cited and discussed in electroacoustic writing but it has been argued, by Kane (2008, section 1), that it also serves as a myth subscribed to by the electroacoustic community to appropriate its own origin within ancient heritage. Kane proposes an
“alternative reading to displace and challenge our thinking about the acousmatic” (section 4) based on the writings of Christian theologian Clement of Alexandria. This reading exposes a very different interpretation of the acousmatics, in which the veil is understood figuratively instead of literally. Kane’s interpretation is particularly potent to the consideration of the role of spatial verisimilitude in music because it denies that the isolation of sounds from concrete world associations leads to a perceptual clarity. Thus, if Schaeffer’s notion that musical meaning comes out of abstract signs holds true, sounds that reference real-world physicality cannot be excluded from contributing to the production of that meaning.

The Clementine reading suggests that the veil is used to distinguish between those students who understand Pythagoras and those who are merely curious about his ideas. In other words, the veil is a metaphor for a lack of clarity in understanding, which is far, indeed almost opposite from Schaeffer’s interpretation that it represents perceptual clarity.

Kane (2008, section 3) proposes several consequences of the Clementine interpretation. One concerns the questioning of the “over-determined division of the eye and the ear” (section 3, point 1): “The Clementine reading makes no explicit mention of any visual concealment, of any hiding or masking of sources” (section 3, point 1). Another consequence (section 3, point 3) concerns the necessary rejection of the ‘acousmatic’ as a tradition grounded in ancient heritage. A third consequence (section 3, point 2) questions the value of considering sounds in perceptual isolation. Schaeffer’s interpretation of the Pythagorean veil expresses that through greater isolation comes greater perceptual clarity of the intrinsic characteristics of sound, which in turn allows the ‘abstract signing’ of musical meaning to shine. In Kane’s Clementine reading the veil represents an opposite; the isolation created by the veil represents a loss of clarity. Kane subsequently explores the implications of this interpretation:

The old emphasis on residual listening, or other modes of listening that reduce the presence of semiotic aspects of sounds would be greatly diminished. The new emphasis would be placed on understanding, not apprehension. Concomitant with this move away from reduced listening would be a move away from
morphologies, typologies and epistemological approaches to the theory of electroacoustic music based on eidetically reduced sound objects. The ontology built on the sound object underemphasizes questions of figurality, tropology or symbolic meaning and purpose, and cannot account for the intensities and possibilities of rhetorical unfolding. (section 3, point 2)

If the importance of perceptual clarity achieved through isolation is rejected, then sound’s capacity to reference external ‘concrete material’ things can no longer be precluded from contributing to musical meaning. Kane’s description of the “intensities and possibilities of rhetorical unfolding” holds a resonance with the iterative dimension of Peirce’s understanding of semiotics, in which interpretants can themselves become signs to further things.

The reality of the relationship between semiotics and music may lie somewhere in the middle of a continuum which is bound on one side by the pure abstract sounds and abstract musical form pursued by Schaeffer and on the other side by the realistic illusion of real-world materiality. Emmerson provides an example of a comparison of two works that sit at different points on this continuum. In “The relation of language to materials” (1986), Emmerson references the argument that Berlioz’s *Symphonie Fantastique* “is a better work” than Beethoven’s ‘Battle’ Symphony because “the Berlioz has more ‘abstract musical’ substance”, which is in finer balance with its “mimetic content” (p.19). The criticism levelled at the ‘Battle’ Symphony is that it is “mere sound effects” (p.19). For Emmerson, the two opposing aspects, abstract form and mimetic content, combine to make “the totality of ‘musical discourse’” (p.19). Emmerson does not comment on the realistic representation of sound; the “mimetic content” he refers to concerns referential icons, that is, likenesses with real-world materiality. Indeed, having been published in 1986, this discussion was not privy to reality-equivalent technologies. However, some thirty years later, and as is mentioned in Chapter 1, Emmerson (2007, p.143) expresses reservation on the judgement of whether or not realistic spatial illusion might be of any interest to compositional concerns. Perhaps, for Emmerson, the realistic representation of space in sound has progressed beyond the ‘mimetic content’ that he sees as part of ‘musical discourse’. Nevertheless, this discussion suggests that spatial verisimilitude will need to be balanced by the presence of abstract musical form.
4. 4 **Signs of space in music**

The ubiquity of the presence of spatial icons in sound raises the question: To what extent are perceptions of space present in music and musical styles that are not explicitly concerned with space? Just as Blauert (1997, p.3) states that ‘spatial hearing’ is a tautology could it be stated that ‘spatial music’ is also a tautology? Certainly the previously discussed association between changes in loudness and distance supports the suggestion that perceptions of space play a role in ‘non-spatial’ musical expression.

In this section, evidence of spatial references implicit in common musical constructs is provided. The aim of this exploration is not so much to argue that space already exists in ‘non-spatial’ music, but rather to highlight that composing with spatial verisimilitude needs to cater for the spatial references already existent in music. The spatial references found in musical constructs discussed below include: the association of height with pitch; the ability for changes in sound loudness to reference distal movement; perceptions of space as a result of changes in timbre; and the impression of movement as a result of changes in tempi. The examples presented here are by no means exhaustive, but they demonstrate references of space present in a broad selection of musical concerns.

The interpretation of pitch as height, where pitches of greater frequency are associated with greater height, is a well-known association between musical constructs and space:

> Musicians have always implicitly accepted some kind of metaphorical or analogical link between perceived pitch and height. We speak about ‘high’ pitches and ‘low’ pitches and about melodic lines or portamenti going ‘upwards’ or ‘downwards’. (Wishart 1996, p.191)

Considered within Peirce’s framework, if the association between pitch and height arises through mere “dispositions or factitious habit” (Peirce, Kloesel & Houser 1998, p.460) then the pitch can be described as a *symbol* of height. Dahlhaus (1985, p.18) concurs that there is little logical association between pitch and height, describing this analogy as ‘a convention of limited historical validity’. Dahlhaus
highlights that before the words ‘high’ and ‘low’ were employed to describe pitch, the words ‘sharp’ and ‘blunt’ were favoured, and “modern psychological studies of aural perception use ‘light’ and ‘dark’” (p.18). Neither of these later labels conjures associations to spatial characteristics.

Wishart (1996, p.191), however, suggests that the association between pitch and height might originate in an “environmental metaphor” caused by the “small body weight”, and subsequent “small sound-producing organ” of flight-capable animals. This interpretation suggests that the pitch-as-height association could be considered an index; the sign functions by virtue of a logical association to the flying capabilities of different sized animals. Smalley (1997, p.122) proposes a similar reasoning also placing the pitch-as-height association as an index.

Yet another view proposes that pitch-as-height might also be considered an icon. Kendall (2010, p.231) cites several listening tests that confirm that “the higher the spectral energy distribution of the source, the higher the perceived elevation”. Kendall attributes this to the effect of the pinnae on sounds originating from above the head. In other words, the shape of the ear causes sounds above the head to be perceived as having a higher frequency spectrum. Thus, higher pitched sounds will have a ‘likeness’ to sounds heard above the head. Within the context of this understanding, pitch can be described as an icon.

If anything, these different interpretations indicate that the delineations between the different sign types may be blurred. However, whichever the correct mechanism for pitch’s capacity to reference height, it is a sign that cannot be ignored. Certain spatial manipulations in combination with pitch-based musical constructs may be difficult to present without perceptual confusion. For example, simulating the downward spatial movement of a melodic line that rises in pitch may create a sense of spatial contradiction. Similarly, the use of increasing pitch on a static source may cause the perception of the upward movement of that source.

In his summary of the historical use of space in music, Begault (2000, p.195) identifies auditory distance as a characteristic of Romantic-era orchestral music. He
(p.195) explains that the “Manipulation of auditory distance in instrumental music can be produced by varying dynamic indications or by reducing the number of instruments in an orchestral desk”. In other words, changes in loudness, however implemented, can act as a reference of distance in space.

Indeed, applying a change in sound loudness is perhaps the single most commonly used sound processing technique used in electroacoustic music. Wishart labels it the most obvious example of the formalisation of acoustic space in contemporary music:

> The formalisation of acoustic space is found in all kinds of contemporary music production […] The most obvious is the rebalancing of instruments by means of differential amplification. (1986, p.46)

Changes in volume, whether applied over time to a single sound or applied differentially to the same two sounds, is a very common musical figure. Within a context where sounds are being spatialised using reality-equivalent technologies, the application of this musical figure may introduce perceptible inconsistencies with the simulated space. The musically realised reduction of loudness may reference a sound moving farther away whilst that same sound’s spatialisation may model a static position. This is a very clear example of how the act of spatialisation cannot be considered independently of the musical material being spatialised, since this last may implicitly contain its own significations of space.

Another common musical parameter, which is intimately related to perceptions of space, is the manipulation of timbre. Timbre is described by Begault (2000, p.27) as the perception of “the spectral content (spectrum) of a sound source, along with the manner that the content changes over time.” Spatial modulations, or the way in which physical space modifies sound, can also be described as changes in a sound’s spectral content over time. Indeed the effect of physical space on sound has been identified as an aspect of timbre (Dannenberg cited in Worrall 1998, p.94). Worrall explains that:

> [Modifications] of timbre (including all its sub-dependencies such as fundamental frequencies, amplitudes, spectra, etc.) affect our perception of different aspects of
Indeed, the ability of timbral manipulations to affect perceptions of space is known to electroacoustic composers who modulate spatial qualities precisely by using processing techniques that change timbre (Emmerson 1998b, p.150; Smalley cited in Austin 2000, p.14). The musical use of timbral modulations must thus be carefully considered when used in spatialised music. The application of a low-pass filter, for example, in which high frequencies are filtered out of a sound signal, can reference a range of spatial characteristics such as: distance; obstruction by a physical object present between the sound-source and the listener; or containment within a closed room. Humidity in the air absorbs high frequencies so distal sounds appear to have a lower frequency spectrum (Blauert 1997, pp.126-127). Lower sound frequencies refract around physical objects more easily than higher frequencies so the obstruction of a sound source by a physical object causes that sound to appear to have a lower frequency spectrum. Hard physical materials tend to absorb high frequencies and transmit lower ones, so sounds heard through a wall, such as outside a nightclub, will appear to have a lower frequency spectrum. All of these spatial scenarios can be referenced by the attenuation of a sound’s higher frequencies, whether that is achieved through digital filters, instrumental manipulations, or orchestration.

Musical constructs that focus on changes in timbre and affect perceptions of space are also prevalent in orchestral music. As Smalley explains “One can […] consider all sorts of orchestral textures as spatial, even in tonal music” (Austin 2000, p.20). Smalley highlights that, in some orchestral music, the identification of sounding objects, in this case musical instruments, can be subsumed to the spatial effect created by timbral modulations:

The best examples of really spatial orchestral music – where one forgets about the instruments, forgets that the sound is coming from people’s blowing and scraping—occur in some of Xenakis’s and Ligeti’s orchestral pieces (Austin 2000, p.20)
In Chapter 3, Xenakis’ composition *Eonta* (1963-1964) is cited as an example of the spatialisation of sounds through the intrinsic mode. In the above passage, Smalley is referencing a different aspect of Xenakis’ work, in which the careful orchestration of multiple instruments is used to create complex textures and timbral modulations. One example of such a work is Xenakis’ *Metastaseis* (1953–4). Examples of György Ligeti’s work that make extensive use of spatial textures are *Apparitions* (1958–59), *Atmosphères* (1961) and later *Lontano* (1967). These compositions, and the techniques they use, form a part of the musical grammar established outside of the explicit spatial concerns of electroacoustic music. This makes their use within spatial composition deserving of careful consideration. Again, the composer cannot assume that existing musical constructs do not contain references of space. The act of spatialising music, using reality-equivalent technologies, must therefore involve the conscious reconciliation of spatial information present within musical constructs, and the spatial information encoded in the *simulated* mode.

Spatial movement is another quality often referenced within musical contexts. It has at least two possible sources; one involves changes in pitch to reference vertical movement as already discussed, and the second involves changes in tempi. Movement is a physical phenomenon that involves both time and space, and can thus be referenced through either changes in time, or changes in space. Indeed Dahlhaus (1985, p.18) identifies both of these in the referencing of movement found in 19th century orchestral music. This highlights that spatial movement is not just another example of the presence of spatial references in music; it is also one that is steeped in musical tradition.

The use of increasing tempi on musical material that is realistically spatially simulated as stationary may result in a musical expression in which the tempo does not reference spatial movement. In this example, where the tempo references increasing speed but the simulation describes a static source, there is not necessarily a conflict of spatial information. However, the realistic spatial representation of the musical material as static may lead to a significantly different interpretation of the changes in tempo. Rather than referencing movement, the sounds heard might rather reference a musical performer. Here, spatial verisimilitude has not necessarily been
contradicted by the spatial reference found in the music, but it has altered its expression.

A range of other spatial references exists in music. As discussed in Chapter 4 the simulation of echoes, which reference large spaces, can be found in the Renaissance, Baroque, and Classical eras (Laurie cited in Begault 2000, p.195). Tools that approximate the simulation of echoes can also be found within contemporary music production environments. ‘Digital delay’ effects are commonly found within Digital Audio Workstations (DAWs), and essentially replicate how an echo might be scored for an instrumental group. Of course, the echo effects found in different classical music traditions are also often accompanied with a spatial antiphony, and DAW’s also cater for a limited spatial separation of sound sources.

Indeed, the most common sound processing functions bundled with popular DAWs such as Protools, LogicPro and Cubase effectively reference space. Panning, reverberation, digital delay, volume control, and low pass filters can each act as referential icons of space; they imbue sounds with qualities that have a likeness to sounds projected in real physical space. The incorporation of these processes into mainstream digital music production environments has broadened the presence of spatial references in common contemporary musical gestures. As stated by Wishart:

The formalisation of acoustic space is found in all kinds of contemporary music production […] However, as these techniques are usually applied on an all-or-nothing basis in a particular song, and are used frequently, they cease to have meaning as features of an imagined acoustic space and become conventions of a formalized musical landscape. (1986, p.46)

Here, Wishart draws a distinction between spatial icons that contribute to the design of a coherent ‘imagined acoustic space’, and those that do not. It is worth mentioning that the use of spatial icons outside of a singular coherent spatial scene does have meaningful historical precedent. As introduced in Chapter 1, there is a strong musical tradition of the use of space to affect music such that certain musical elements are easier to perceive. From the work of orchestral composers such as Charles Ives and Henry Brant, to the work of more contemporary composers such as John Cage and
Iannis Xenakis, space’s effect on sound has been used to support the act of listening to the music, rather than manufacturing new and consistent acoustic scenes. The use of sound processing effects that create spatial icons, applied to music by way of DAWs, can be seen as a continuation of this musical tradition. These ‘effects’ are prevalent in the production of contemporary electronic music, and are used explicitly to facilitate the perceptual separation of different parts within the music. This technique, in which space is recognised to be an effective mechanism for the perceptual segregation of auditory streams, is discussed in Bregman’s text *Auditory Scene Analysis: the perceptual organization of sound* (1990, pp.293-294).

In conclusion, the presence of references of space in music demonstrates that the term ‘music spatialisation’ is problematic. Within the word ‘music’, space already exists. Reality-equivalent technologies can create perceptions of space, but so can music. On the one hand, this is a technical or procedural problem for composers working with spatial verisimilitude; they must reconcile the space present in music with the space simulated by technology. On the other hand, if music already has a capacity to create perceptions of space, then what does the simulation of space bring? In the following chapter, the notion of illusion in sound is examined. In this examination it is suggested that the simulation of space is more concerned with bringing illusion to music, than with synthesising spatial characteristics. As such, the question of the relationship between music and spatial verisimilitude is transferred to a consideration of the relationship between music and illusion.
5 VERISIMILITUDE IN SOUND

By persistently underpinning the realist element in his work with the help of those everyday intonations which will convince and persuade all strata of his audience, and have without doubt become a completely real expression of ideas and emotions in the view of most people, the composer attains indisputably to realism and to general recognition of his music over the course of many generations (Boris Asafyev cited in Dahlhaus 1985, p.101)

The notion of realism in music, as expressed by the Russian and Soviet composer Boris Asafyev, concerns the convincing depiction of ideas and emotions. Through musical form the composer can achieve the ‘completely real expression’ of human concerns. This interpretation of realism is congruent with the theories of art prevalent in the 16th, 17th and 18th centuries, which adopt the classical Aristotelian idea of ‘Imitatio naturae’, that is: the imitation of nature (Dahlhaus 1985, p.17). Where ideas and emotions are concerned, however, the nature that is being imitated is humanly centred (p.19). In his text Realism in Nineteenth-Century Music, Carl Dahlhaus seeks to identify evidence of the late 19th century notion of realism, characteristic of the visual and literary arts, within music. He begins by identifying six different interpretations of ‘imitatio naturae’ in the music leading up the 19th century (pp.16-29). Examples also include the imitation of ‘outside’ nature such as the naturalistic copying of non-musical sounds, pitch-based metaphors for vertical spatial movement and the musical representation of speech intonations. Some of these exist in 19th century music but Dahlhaus (p.121) does identify new aesthetic, dramaturgical and compositional phenomena related to notions of realism. He concludes, however, that any evidence of realism in music is intimately tied with the concerns of the time and, for the 19th century, these involve epistemological doubt about the nature of reality. Ultimately Dahlhaus argues that the relationship between realism and music is generally problematic, and ‘ever since the aesthetics of the beautiful was displaced by the aesthetics of the true, the problem of what ‘true’ reality ‘really’ is has plagued compositional practice, as well as theories about music’ (p.115).
The type of realism discussed within this thesis is concerned with representing external reality. It is an interpretation of realism concerned with the realism in the representation, rather than the subject of the representation, as was the predominant inclination in 19th century music. However, far from being incongruent with Dahlhaus’ account of realism in 19th century music, this chapter elucidates that the convincing representation of external reality is a contemporary musical concern. It is a concern born of technological advances, beginning with recording technology, which effectively transfer the understanding of reality away from philosophical thinking, towards the science of perception. In so doing, the engagement with realism is drawn away from the interpretation of ideas and becomes heavily technologically mediated. In this technological mediation, the electroacoustic composer is drawn away from the question of reality, and becomes subsumed to its scientific interpretation, its technological embodiment, and its focus on appearance.

In this chapter, two distinctly different sources for the technological mediation of realism in sound are identified and discussed. The first is the early history of recording and playback technology in which realism is achieved through the mechanical capture and reproduction of sounds. Here, the notion of fidelity plays a central role. The second is the appropriation of electronic technology, and later digital technology, to music making. The technological manipulation of sounds exposes the composer to the limits of auditory perception; thus making perception a central concern, and crystallising the psychology of perception as the principal interpretation of reality.

5.1 Fidelity as the technological interpretation of verisimilitude.

The end of the 19th century saw a technological development that was to dramatically alter sound’s capacity for realism. The invention of the phonograph allowed the capture and playback of sounds through the precision of a mechanical device. With this technology, notions of what could be considered ‘real’ in music take on an entirely new dimension.

In Mahler’s First Symphony, realism exists through the quotation of folk-like melodic content, collaged into a grandiose and abstract symphonic setting (Dahlhaus 1985, pp.109-110). Here, the appearance of being real is achieved within the subject
matter of the music. In a music recording, the appearance of being real is not associated with the subject matter of the music at all; it is only associated with the reproduction of the sounds. The verisimilitude does not support the expression of the music; it supports the act of listening to the music as though one was ‘there’. In other words verisimilitude, as it is considered in early sound recording technology, is divorced from any musical intent. Music reproduction instead primarily serves the goal of accessibility to listening experiences. Early recording technology needs to be considered here because of two reasons; firstly, it sets an expectation for what can be considered real in technologically mediated music, and secondly, it highlights an important distinction between the appearance of being real, and actually being real.

Recording and playback technology has given rise to a quality intimately tied to verisimilitude: that is, fidelity. Fidelity is defined by The Oxford English Dictionary (Simpson & Weiner 1989, p.877, vol 5) as ‘Strict conformity to truth or fact […] correspondence with the original; exactness’. As is discussed below, correspondence with the original is often used as a measure of quality of reproduced sound. Yet, to achieve the impression that a recording is faithful, it is found that recordings need to satisfy preconceptions of the appearance of being real. Within this perspective, the moral and ethical dimensions of the word ‘fidelity’ are unceremoniously discarded. In order to achieve the impression of being faithful to the original, verisimilitude eclipses veridical accuracy. The early history of the phonograph provides a context for the exploration of this dichotomy, which remains potent in contemporary sound production.

American inventor and businessman Thomas Edison is credited as being the inventor of the first practical sound recording and reproduction device. Edison’s phonograph, invented in 1877, was developed in an era in which sounds had never before been captured and reproduced. As such there was no precedence for the comparison of the quality of reproduced sounds. Thomson (1995) gives a detailed historical account of how the idea of fidelity was used to support the introduction of the phonograph into different markets. The first commercial incarnation of Edison’s phonograph was directed at business users as a means of creating aural letters. Within this proposed use, fidelity was defined as the “retrievable truth of the message” (p.137). The
verisimilitude in the voice recording served to confirm the authenticity of an oral contract or agreement “rendered permanent and therefore indisputable” (p.137). The phonograph eventually failed as a business machine but was successfully appropriated as a coin-operated entertainment device, installed in saloons and hotel lobbies, playing recorded musical works. As a provider of entertainment, the phonograph’s measure of fidelity moved away from the intelligibility of the spoken word to the faithful rendition of musical works. As early as 1890, prior to its general public availability, reviews of the phonograph emphasised the realism of its reproduction; recordings “rendered with so startling and realistic effect that it seems almost impossible that the human voice can issue from wax and iron” (New York Journal 1890 cited in Thompson 1995, p.138). Such enthusiastic reviews begin to indicate the cultural and contextual dependence of the appearance of being real. It is now difficult to believe how the sound of a voice mechanically reproduced over a hundred years ago could be described as startling and realistic.

As the commercial potential of recording and playback machines began to be realised, competitors entered the market and Edison was eventually forced to discard cylinders, as the capture medium, in favour of discs. In a 1914 concert, Edison’s new Diamond Disc was directly compared to his old Edison Standard cylinder machine. The audience reaction consisted of bursting into laughter upon hearing the sound of the older device (p.137). Thus, through exposure to a better appearance of being real, previous expectations were revealed as inadequate. In other words, experience plays a role in the perception of what is real. Having received a level of ‘education’ about the possible quality of reproduced sounds, the audience’s perception of what could be perceived as real had significantly evolved.

A year later, in 1915, Edison devised a public demonstration that reportedly successfully convinced audience members that the sound of the Edison Diamond Disc was ‘absolutely identical with and indistinguishable from the original’ (William Maxwell, Edison Retail Sales Laboratory cited in Thompson 1995, p.148). This startling claim, of the equality of reproduction and original, was supported by the audacious comparison of a live performer and a recording played back on an Edison Diamond Disc. This demonstration, called a tone test, was quickly adopted as a
marketing campaign purportedly witnessed by a total of 2 million people between the years of 1915 and 1920 (p.137). Over those five years, the art of creating tone test concerts evolved. Theatrical elements were included and any techniques that better supported the illusion of similarity between real and recorded sounds were adopted; female voices had better success so women performers were favoured (p.137); the live singer would never sing without the accompaniment of the phonograph, thus maintaining a level of invariance in the comparison (Thompson 1995, p.152; Milner 2009, p.6); and the singer adjusted their volume to match that of the quieter machine (Milner 2009, p.7). In analysing reports of the success of these tone tests, Thomson (1995, p.156) describes accounts that invert the role of imitation and authenticity. In one review of Christine Miller’s 1915 tone test concert, the reviewer notes that the singer ‘adjusted the power of her voice to that of the ‘record’ with skill and the reproduction was closely imitative’ (Boston Evening Transcript cited in Thompson 1995, p.156). Within this review is a subtle but telling reversal of roles. The recording is cast as the authentic work, and the live performer takes the role of the imitator of the recording.

Such reports begin to cast Edison’s tone tests as exercises in creating convincing illusions of reality, rather than truly attempting to compare a reproduced sound’s equality with reality. Reality was manipulated to support its own illusion. Whether motivated by the need to drive commercial success or by an understanding of human perception, Edison’s tone tests demonstrate that the power of the illusion of reality is not diminished by falseness. In this sense, the use of the word ‘fidelity’ to describe the impression that a recording is similar to the original contradicts the word’s moral dimension. Fidelity is achieved through the illusion of faithfulness, not the faithful representation of an original. Indeed, the creation of that illusion may involve blatantly inaccurate representation.

Thus, reports on the beginnings of sound recording technology suggest two points about verisimilitude in sound. The first is that what is perceived as real is subject to change; the second is that the illusion of being real need not be veridical. Both points are supported by certain approaches to illusion in the visual arts. Notwithstanding the differences between sound and vision, the depth of the tradition of illusion in the
visual arts allows a perspective that pre dates technology. To this end an important and influential text examining the psychology of representation in the visual arts is now consulted.

In *Art and Illusion* (1977) Gombrich examines how visual artists create illusion. Initially published in 1960, many commentators have since scrutinised and extended Gombrich’s ideas but, as one of the first art theorists to take a cognitive approach to understanding illusion, his basic premise translates well to audio illusions. Indeed, Gombrich introduces this approach by recounting a personal experience, from his time working in the Monitoring Service of the British Broadcasting Service, in which audio transmissions required decoding.

Some of the transmissions which interested us most were often barely audible, and it became quite an art, or even a sport, to interpret the few whiffs of speech sound that were all we really had on the wax cylinders on which these broadcasts had been recorded. It was then we learned to what an extent our knowledge and expectations influence our hearing. You had to know what might be said in order to hear what was said. [...] For this was the most striking experience of all: once your expectation was firmly set and your conviction settled, you ceased to be aware of your own activity, the noises appeared to fall into place and to be transformed into the expected words. So strong was this effect of suggestion that we made it a practice never to tell a colleague our own interpretation if we wanted him to test it. Expectation created illusion. (1977, p.171)

This short anecdote alludes to the backbone of Gombrich’s explorations of pictorial representation, in which he argues that perception is governed by the presence of ideas or concepts called schemata. Both the artist and the audience perceive by comparing stimuli against these schemata. Visual artists do not begin a work by capturing what they see, but rather with a conception of what they see, which is then tested and adjusted. Gombrich (1977, p.24) describes this as “making and matching”. The schemata inform what is made, and this is matched against the objects in nature, the ‘motif’. The presence of schemata in perception means that there is no “neutral naturalism” (p.75). This is a far-reaching hypothesis with many implications in the pursuit of verisimilitude. For Gombrich, the appearance of being real is not created
by copying nature; it is created by the work’s resonance with the audience’s conception of how reality might appear. The creation of verisimilitude does not require coincidence with reality but rather coincidence with the audience’s expectation of the appearance of reality.

Thus the ability of an early 19th century technology, the phonograph, to create the illusion that a reproduced voice has the appearance of being real is explained. Simply put; the listener’s expectations, or schemata, were met. Newer and better recording technologies resulted in the adjustment of listener’s schemata. And Edison’s strategy of sacrificing accuracy in his tone tests in the interest of supporting the illusion of being real now appears valid, rather than cynical. Fast forward to the present, and the brute computational simulation of nature, such as the modelling of the acoustics of physical spaces common to reality-equivalent technologies, can be questioned for its ability to create the appearance of being real. For Gombrich, creating verisimilitude is not a question of mimetic accuracy achieved through such techniques as simulation. The imitation of nature may play a role in the artist’s process but, ultimately, verisimilitude is created through the artist’s capacity to tap into the action of the schemata shared between creator and perceiver. Thus, in pursuing spatial verisimilitude, composers might better ask themselves: what is the audience’s expectation of the appearance of reality in space?

Both Gombrich and his contemporary James Gibson, a perceptual psychologist responsible for what is known as the ecological approach to perception, have leant on the concept of fidelity in sound to illustrate related concepts in visual perception. The difference in their appropriations, however, illustrates the difference in their understandings of perception. Gombrich (1973) cites the use of fidelity within the recording industry to impress the importance of being faithful to listener’s expectations rather than being faithful to the original performance. He notes that recording artists do not hesitate to edit a recording to improve it (pp.196-197). The improvement brings the recording closer to the artist’s expectations of how it should sound, and farther from the recorded reality. Matching the expectation of how a recording should sound, at the expense of its fidelity, creates a perfected idea of reality.
Gibson, however, adapts the concept of fidelity in sound to illustrate an entirely different point. Like Gombrich (1973), Gibson is also concerned with how we perceive mediated images. He uses the concept of a scale of fidelity, from low faithfulness to high, to illustrate what he argues is a falsity about how we perceive mediated information. Gibson (1971, pp.28-33) rejects the idea that high fidelity can ever approach or achieve identity with reality. For Gibson, there is a fundamental distinction between perceptions of reality and perceptions of mediated reality; “And so, despite all the stories of paintings that are said to deceive observers into trying to lift the curtain, or eat the grapes, or walk into the scene, I am sceptical” (p.33). Such a view also finds resonance in Thomson’s account of Edison’s tone tests. Whilst there were many positive reports about the realistic reproduction of the music, Thomson concludes: “the question remains did most, or even many, people conclude that the living performance and its re-creation were actually acoustically indistinguishable?” (1995, p.159). Such a question might engage the scientist more than the artist since, as already discussed; the artist intuitively understands that identity with the original is not a precondition for the perception of verisimilitude. However, in its technological embodiment where Gombrich’s schemata cease to exist, the shared expectations between artist and perceiver are inaccessible, and so the pursuit of verisimilitude is cast as the imitation of nature, supported by the scientific study of perception.

This opposition between true fidelity and verisimilitude in sound also exists in audio-visual media. Emmerson (2007, p.143) cites Chion who argues that aspects of sound in film require exaggeration “in order to appear ‘real’”. Through exaggeration truth is sacrificed to achieve verisimilitude. A similar perspective exists in computer game design. In a personal email discussion with the author, Simon Goodwin; Principle Programmer at Codemasters and responsible for the critically acclaimed sound in some of Codemaster’s games, says:

Like the video people, in AR [augmented reality], cinema and games, synthesis rather than capture is our focus, hyperrealism is our aim, rather than metric accuracy, and that seems both more interesting and more attainable. (Deleflie 2007)
Goodwin’s statement, and his use of the term ‘hyperrealism’, stresses the importance of pursuing the illusion of reality, rather than similarity with reality. Both these accounts highlight Gombrich’s view; that verisimilitude is achieved not by coincidence with reality, but rather by coincidence with the perceiver’s expectations of reality.

5.2 Synthesis of verisimilitude: informed by the perceptual sciences
In Gombrich’s thesis it is the artist, or in our case the composer, who brings about verisimilitude by intuitively and constantly appraising the results of their work. When the composer is excluded from this process, the technological encapsulation of verisimilitude can only be informed by the scientific understanding of perception. There is a broad range of empirically informed perceptual models that can be applied to the simulation of sound in space. The work of R.L. Gregory, for example, holds a certain resonance with Gombrich’s ideas in that perception is understood to involve testing a formulated conception of the world. In his paper ‘Perceptions as Hypotheses’ (1980) Gregory identifies parallels between the scientific method of creating and testing hypotheses and the mechanism of perception. A broad survey of how this and other perceptual models might shed light on the authorship of spatial verisimilitude would be of benefit to spatial audio composers. In the interests of brevity, however, this section presents a critically selected cross-section of research that is specifically concerned with interactive mediated environments and virtual reality. Such research involves the application of perceptual models to better understand the illusion of being real, which is here understood by the notion of presence.

Presence is central to research concerned with mediated environments. It is a concept that has received a great deal of attention, yet exactly what it is and how to define it “has been the focus of considerable research and theoretical attention across several academic disciplines” (Preston 2007, p.278). Within the context of the reproduction of captured music, the notion of presence is intimately tied to fidelity. Rumsey, an engineer and researcher concerned with the spatiality of the reproduction of recorded music, observes that for certain musical genres “it is often said that the aim of high-quality recording and reproduction should be to create as believable an illusion of
‘being there’ as possible” (2002, p.652). In *Spatial Hearing*, a text introduced above, Blauert’s definition of virtual-reality (VR) also implies presence: “Ideally, the subjects feel ‘present’ in the virtual environments and actually accept them as being real: virtual-reality!” (1997, pp.383-384). Within this second definition, which is not directly concerned with music capture, there is no appreciable concept of fidelity.

It is more difficult to pin down a notion of presence that can be broadly applied across different types of mediated environments. The last ten years of presence research has introduced new ways to conceptualise mediated realities. The relationship between mediated realities and consciousness, for example, is one area of research that has attracted attention (Sanchez-Vives & Slater 2005; Preston 2007). In the proceeding discussion several different encapsulations of presence will be identified, and how they might shed light on the pursuit of verisimilitude in sound will be examined.

5.2.1 Presence as the illusion of nonmediation

A concise, elegant, and much cited interpretation, which has already been introduced in Chapter 1, is proposed by Lombard and Ditton (1997) in a paper entitled ‘At the Heart of It All: The Concept of Presence’. Here, presence is defined as “the perceptual illusion of nonmediation” (section "Presence Explicated").

The illusion of nonmediation occurs when the subject “fails to perceive or acknowledge the existence of a medium in his/her communication environment and responds as he/she would if the medium were not there” (section “Presence Explicated”). This explanation indicates how the successful illusion of presence might be measured: by the perceiver’s response to the mediated environment instead of the immediate physical environment. If the perceiver’s response is caused by the mediated reality, then some degree of presence has been achieved. Such a definition perhaps favours mediated environments that allow for a degree of interaction. In certain music listening modes, such as concert music, evidence of response will be limited. Although, as is discussed later, even in the case of listening to concert music there is still some level of response possible in the form of small head movements.
Lombard and Ditton’s (1997) definition is included here because of their discussion on different conceptualisations of presence identified in the literature. They outline six different encapsulations some of which help illuminate notions of presence in sound. IJsselsteijn et al. (2000) group Lombard and Ditton’s outline into two categories; physical and social presence. Social presence “refers to the feeling of being together (and communicating) with someone” (section 1.2) such as in teleconferencing applications. Physical presence refers to the “sense of being physically located somewhere” (section 1.2). In this latter category of presence, Lombard and Ditton identify three types: presence as realism, presence as transportation and presence as immersion. All three have meaning in the light of presence in music.

Within presence as realism, Lombard and Ditton (1997) distinguish between what they call social realism and perceptual realism. Social realism refers to the kind of realism that has the potential to actually be real; it is plausible. Perceptual realism, however, exploits an understanding of perception to create the impression that a fictitious or improbable scene is real. In the faithful reproduction of recorded music, social-realism is targeted since what is heard already has a high level of plausibility since it was, originally, real. In the composition of abstract music, where what is heard has a low level of plausibility, perceptual realism will be targeted. This is illustrated later, in the exploration of Stockhausen’s composition Kontakte, in which the creation of the perception of distance of an abstract sound brings significant difficulties. Distance is more easily perceived in recognised sounding objects, where the listener already has an understanding of how something sounds when it is near or far. The synthesis of distance in an abstract sound requires a greater understanding of the perceptual mechanisms involved in judging distance. With a low level of plausibility, realism in abstract sounds must target perceptual realism.

Within presence as transportation Lombard and Ditton (1997) identify three subtypes; “you are there”, “it is here” and “we are together” (section "Presence Explicated"). The notion of “we are together” most aptly applies to mediums such as video conferencing. “You are there” is a common contemporary encapsulation of presence in sound: both Rumsey’s conception of spatiality in reproduced sound and
Blauert’s definition of VR concerns the illusion that the listener is elsewhere. This has not always been the case. Edison’s tone tests aimed to convince the audience that the recorded performer was actually present in the room: “it is here”. It is interesting to note that the difference between “it is here” and “you are there”, only involves a distinction in spatial acoustics. In both cases, the listener and the heard subject are together, but in “it is here” the spatial acoustic remains that of the listener’s present physical space, whereas in “you are there” the spatial acoustic belongs to the physical space of the heard subject. The difference between these two forms of ‘presence as transportation’ exists only in that the “you are there” modality attempts to create a spatial illusion. Thus, the predominance of the transportation mode of “you are there”, in music concerns, highlights the importance of the illusion of space in contemporary notions of presence.

Presence as immersion is defined as "the degree to which a virtual environment submerges the perceptual system of the user" (Biocca & Delaney cited in Lombard & Ditton 1997, section "Presence as immersion"). Two simple measures of this form of presence are: the count of the number of senses supplied with stimulus, and the degree to which inputs from the real physical environment are limited. Targeting only one sense, music’s capacity for creating immersive presence will be limited. However, certain performers such as Francisco López have employed blindfolds distributed to the audience (Lopez 2004) to block contradictory visual stimuli thus increasing immersive presence.

Lombard and Ditton (1997, section "Presence Explicated") stress that aural stimulus is undervalued in generating presence. Their review of the literature identifies that both sound quality and spatial dimensionality contribute to presence. They cite a study examining soundtracks in action films that found that “the presentations with high fidelity sound were judged more ‘realistic,’ but it was the low fidelity sounds that made subjects feel more ‘a part of the action.’ ” (Reeves, Detender, and Steuer cited in Lombard & Ditton 1997, section "Causes and Effects of Presence"). In other words, sound quality affects presence, but it is the context the sounds are presented in that dictates whether high or low quality will create a greater sense of presence.
Such an observation lends credence to Milner’s observation, cited earlier, where the high quality of reproduced sound is perceived as uncharacteristic of ‘live’ music.

5.2.2 Presence as perceptually-induced

In *Auditory-Induced Presence* Larsson et al. (2005) examine the notion of presence in situations where the aural sense dominates, such as in the act of listening to music. Their interpretation of the term presence focuses on the perceptually-induced sense of “being there”. Their account of auditory-induced presence is cited here for two reasons: they identify how the aural sense is different to the visual sense in contributing to presence, and they propose recommendations for presence-oriented sound design.

Larsson et al.’s (2005) review of the literature identifies several characteristics unique to the aural sense that contribute to perceptually induced presence (p.3). The first is that aural awareness is not restricted to a fraction of the environment, as it is with visual stimulus. Auditory awareness covers the entire spatial environment that contributes to sense of being within a physical space. The second is that temporal resolution in the aural sense is much higher than the visual medium. This can contribute to the sustenance of movement and activity where the visual medium would otherwise be still. The third considers sound as a medium of suggestion that has “affective and mood induction properties” (p.3). In these last two characteristics, the musical dimension of sound is implied. Larsson et al. do not comment on how the presence of music might itself affect auditory-induced presence.

Larsson et al. (2005) conclude their paper by suggesting recommendations for presence-oriented sound design. These are summarised below.

- Spaciousness

For Larsson et al. spaciousness is produced by creating a minimum level of reverberant sound in spatial sound. They reference Begault (2001) to underscore the role of the reverberant field in the externalisation of sounds when rendered binaurally (Larsson et al. 2005, section 4). They (section 4) note that their understanding of
‘spaciousness’ is similar to Rumsey’s definition of the word ‘presence’, already discussed above, as “the sense of being inside an (enclosed) space or scene” (Rumsey 2002, p.663).

- Interaction – near field

Near field interaction refers to the sound of interactions that occur within the listener’s peripersonal space; that is, very close to the listener.

- Interaction – far field

Far field interaction refers to how the acoustics of physical spaces change as a result of movement. This suggests the importance of including transitions between acoustically different spaces, of modelling the effect of occlusion by obstructing objects, and of modelling exclusion: the sounds heard through openings such as doors and walls.

- Scene Consistency

The lack of scene consistency in auditory scenes has a considerable impact on presence. This inconsistency can be caused by incorrectly modelled spatial stimuli or by ecological inconsistency. In ecological inconsistency, sounds are presented that have a low degree of real world plausibility. For example, “strange combinations of naturalistic stimuli representing concrete sound objects (e.g. dog, bus) and artificial sound (modulated tone)” (p.11). This later type of inconsistency creates a low level of what Lombard and Ditton call ‘social realism’ (Biocca & Delaney cited in Lombard & Ditton 1997, section "Presence as immersion") as discussed above. In other words, Larsson et al. suggest that a low level of ‘social realism’ affects presence.

Larsson et al. (2005, section 4) also cite evidence suggesting that “virtual sound environment complexity” can negatively affect spatial perception.
5.2.3 Presence as successful action in the environment

In summarising different approaches to presence, Sanchez-Vives & Slater (2005) describe one fundamentally different view of presence that focuses on interaction with the environment. In this approach what is important is “action (how things are done) and the affordances offered in the virtual environment, rather than just appearances, and that the sense of ‘being there’ is grounded on the ability to ‘do’ there” (p.333). This view of presence references some of the work of American psychologist James Gibson whose theories of perception have been widely adopted in the fields of computer games and virtual reality because of the centrality of the notion of interacting with the environment (Preston 2007, p.278).

Applied to the experience of real-world environments the characterisation of the environment as consisting of affordances for action allows the identification of the truthfulness of perceptions that escapes the problematic comparison between the perceived truth and the notionally objective one. Applied to mediated environments, this characterisation allows identifying whether the perceiver is responding to the mediated environment or the real physical environment. Thus, by focusing on interaction with the environment, this approach allows a measure of presence without needing to consider the verisimilitude of synthesised stimuli.

Within the context of music listening situations such as concert halls, interaction with the mediated environment will be limited, but not inexistent. Small head movements have been recognised as contributing to the perception of location (Blauert 1997, pp.43-44, 383). Here, successful action in the environment can be regarded as the listener perceiving subtle changes in sound as a result of head movements. Listening to sound on headphones thus reduces presence unless, of course, some form of head tracking is included.

5.2.4 Gibson’s ecological approach to perception

Gibson (1986, p.238) denies that perception involves the construction of mental representations from sensory inputs. Instead, his ecological approach to perception places importance on the constantly changing relationship between the environment and the self. For Gibson, information is “picked up” (p.238) not through the
interpretation of sensory inputs informed by experiential memory, but rather through how aspects of the perceived environment change or don’t change (p.239).

The ecological approach to visual perception […] begins with the flowing array of the observer who walks from one vista to another, moves around an object of interest, and can approach it for scrutiny, thus extracting the invariants that underlie the changing perspective structure and seeing the connections between hidden and unseen surfaces. (Gibson 1986, p.303)

The observer’s relationship to the environment not only supports the ‘pick up’ of information, but also confirms their presence within that environment. For Gibson, “self awareness accompanies perceptual awareness” (1986, p.263).

Reed identifies three fundamental concepts presented within Gibson’s *The Ecological Approach to Visual Perception* (1986):

The new concept of persistence and change as reciprocals, which replaces the old distinction between space and time, lies at the heart of ecological psychology. Next is the concept of information specifying its source in the environment. Finally, there is a new concept of the environment as consisting of affordances for action, offerings that can be perceived and used by observers. (1988, p.280)

Whilst Gibson’s ideas concern perception generally, his application focuses predominantly on the visual medium. The appropriation of an ecological approach to sound has received limited attention (Worrall 1998, p.96). One of the first efforts in this vein is from Gaver (1993, pp.1-2) who draws a distinction between *musical listening* which focuses on such things as pitch and timbre, and *everyday listening*. For Gaver, “the distinction between everyday and musical listening is between experiences, not sounds” (p.2). In other words, the distinction is made by the context or environment within which those sounds are heard, not by the sounds themselves.
Worrall (1998) takes some further steps in examining how Gibson’s ideas can be appropriated to sound. He draws parallels between Gibson’s concept of ground\textsuperscript{15} with ambient sound and Gibson’s concept of texture\textsuperscript{16} with the reverberant field, spectral changes, and depth cues such as loudness (pp.96-97). He also examines how Gibson’s notion of invariants\textsuperscript{17} might be applied to sound.

Lennox, Vaughan and Myatt (2001, p.1) define the term “ambience labelling information” to describe perceptually significant aspects of the relationship between sounding objects and their environment. Whilst carefully acknowledging the qualitative distinction between sound and vision, they note that ambience labelling information can be considered a parallel to Gibson’s notion of texture. The comparison is similar to that suggested by Worrall, as discussed above.

5.3 Presence and music
In Edison’s days, the music captured by the phonograph had no explicit interest in the mechanisms of audio perception. It was the pursuit of fidelity, not the music itself that brought interest in audio perception to musical endeavours. This changed with the arrival of the electronic medium. As is discussed below, the electronic medium forces a concern for audio perception and, by extension, introduces a newfound concern for space. In this section the research presented above is applied to comments made by Stockhausen concerning the challenge of authoring perceptions of distance in the composition Kontakte. It is argued that Stockhausen’s techniques differ significantly from techniques informed by the scientific understanding of perception, which would result in a lower level of presence. However, it is also

---
\textsuperscript{15} Gibson sees ground as a continuous background surface against which places, objects and other features are laid out (1986, p.148). Information is perceived through how these relate to the ground. For example, our ability to perceive visual object distance has a dependency on how the size of objects changes relative to the ground (pp.160-161).

\textsuperscript{16} Gibson’s concept of texture concerns the “structure of a surface, as distinguished from the structure of the substance underlying the surface” (1986, p.25). He states, “mud, clay, sand, ice and snow have different textures” (p.28). It is our ability to perceive subtle changes in texture, for example, that Gibson argues provides information.

\textsuperscript{17} Gibson’s invariants refer to aspects of the visual environment, such as textures, that do not change. It is how other aspects change, relative to the invariants, that result in the reception of information.
shown that by expanding the understanding of presence to include notions of consciousness, musical form can itself contribute to the feeling of presence thus acting in parallel to, or perhaps in competition with, presence achieved through spatial verisimilitude.

The arrival of electronic technology in the late 19\textsuperscript{th} and early 20\textsuperscript{th} centuries had a profound effect on music. Milton Babbitt, an early American pioneer in electronic music, describes this effect as a compositional revolution. He takes care to qualify the nature of this revolution, warning against the assumption that the properties of the electronic medium should be interpreted as a new set of properties for composition:

For this revolution has effected, summarily and almost completely, a transfer of the limits of musical composition from the limits of the nonelectronic medium and the human performer, not to the limits of this most extensive and flexible of media but to those more restrictive, more intricate, far less well understood limits: the perceptual and conceptual capacities of the human auditor. (1962, p.49)

Babbitt’s point is that the electronic medium has a breadth of possibilities that extends far beyond both what we can perceive and what we can easily conceive. Composing thus becomes an exercise in discovering how to navigate broad possibilities whilst remaining within the perceptual limits of human listeners. In other words, alongside the challenging of composer’s conceptual capabilities, the introduction of the electronic medium has placed auditory perception as a central concern within the compositional act. Di Scipio concurs, stating that the perception-centric nature of electroacoustic music has been emphasised by many composers, including Stockhausen, Risset and Smalley (1995a, p.381).

This recognised emphasis on the importance of perceptual concerns, within electroacoustic composition, necessarily extends to the spatial representation of sound. This is because space is an implicit part of auditory perception. In Spatial Hearing, an important text summarising the scientific literature concerned with audio perception, Blauert (1997, p.3) states that his text’s title “is in fact tautological; there is no ‘nonspatial hearing’”. In other words, 	extit{all perceived sound is spatial}. As such,
spatial concerns are implicit in any understanding of the perception of sound, and the spatial characteristics of sound must therefore assume a prominent role in the electronic medium. Some composers have expressed awareness of this. Stockhausen (1989, p.192), for example, has expressed the view that the spatial attribute of movement is as important a musical parameter as any other.

As a pioneer of the electronic medium, Stockhausen’s thoughts are particularly interesting because they represent an early conscious engagement with the scientific understanding of perception, and highlight what might be called a compositional engagement with perception. In 1971 Stockhausen (1989) presented a lecture in London, in which he spoke about four criteria of electronic music. The criteria labelled ‘The Multi-Layered Spatial Composition’ concerns the use of spatial depth to layer sounds in a similar way to how it might be done with melody or harmony. Kontakte featured this technique but Stockhausen expresses disappointment at the audience’s perception of the illusion of depth:

Now I come to my point: when they hear the layers revealed, one behind the other, in this new music, most listeners cannot even perceive it because they say, well, the walls have not moved, so it is an illusion. I say to them, the fact that you say the walls have not moved is an illusion, because you have clearly heard that the sounds went away, very far, and that is the truth. (1989, pp.107-108)

Stockhausen laments what he believes is the dominance of visual perception over auditory perception: one denies the verisimilitude of the other. Indeed, as discussed above, the importance of the consistency of illusion across different senses is recognised within the discourse of mediated environments. The use of blindfolds may have helped support the perception of distance by immersing listeners exclusively in auditory stimuli. However, Stockhausen is not so concerned with Kontakte’s encapsulation of spatial sound stimuli. He suggests that the success of the ‘new music’ and its spatial dimensions depends on the receptivity of the listeners:

What makes it so difficult for new music to be really appreciated is this mental block ‘as if’, or that they can’t even perceive what they hear. (1989, p.108)
This approach is very much aligned with Gombrich’s understanding of the creation of illusion. For Stockhausen, the audience has not yet developed the schemata that allow them to perceive the sounds as he has. It is not a question of providing the necessary auditory stimuli; it is a question of the alignment of expectations between composer and listener. Of course, Gombrich might argue that if the audience has failed to perceive the illusion as designed, then it is a reflection on the artist’s failure to understand and match the expectations of his audience. For Stockhausen however, it is the listeners who must adapt so that they can “perceive this music in its real terms” (1989, p.108).

The examination of how Stockhausen attempts to create a sense of depth reveals a technique that is fundamentally different to the synthesis of distance informed by a scientific understanding of auditory stimuli. In *Kontakte*, a sense of depth is created through the sequential revelation of sounds. Stockhausen superimposes six layers of sound, and then proceeds to strip them back one by one:

I stop the sound and you hear a second layer of sound behind it. You realise it was already there but you couldn’t hear it. I cut it again, like with a knife, and you hear another layer behind that one, then again. (1989, pp.105-106)

Each sound is revealed by stopping another, louder, one that originally masked it. Through this technique Stockhausen aims to create layers of depth where farther sounds are revealed by stopping nearer sounds. It is a technique that has a significant *temporal* dimension, which necessarily contributes to the structural order of the composition. He later describes another technique for creating a sense of distance: for an unrecognised sound to suggest distance Stockhausen explains that it must be heard “several times before in the context of the music, in order to know how it sounds when closer and further away” (1989, pp.106-107). Again, here, a temporal dimension is implied. The composer first attempts to educate the listener on how an abstract sound appears at different distances. Exactly how Stockhausen establishes the differences in sound as they are heard at different distances is not clear. As is described in the next chapter, some rich spatial characteristics are present in *Kontakte*, but these are the result of sound capture, not synthesis. Accounts suggest that Stockhausen’s synthesis of distance depend primarily on musical structures such
as changing dynamics and changing speed of repetition (Dack 1998, pp.113-114). In this case, the word ‘synthesis’ might best be replaced with the word ‘reference’. Indeed, this distinction points to earlier discussions, elaborated in Chapter 4, in which musical structures are shown to have the capacity to act as ‘references’ of space.

From the perspective of a scientific understanding of perception, as explored in the field of mediated environments, Stockhausen’s techniques would be subject to much criticism. Firstly, as informed by Lombard and Ditton (1997), the use of predominantly abstract sounds means that successful illusion requires the pursuit of *perceptual realism*, which involves the accurate modelling of audio stimuli. Abstract sounds have low *social realism*, that is: they are typically implausible, and so necessitate a greater adherence to perceptual accuracy. Thus *Kontakte*’s use of abstract electronic sounds, in combination with a limited synthesis of distance means there will be a low level of presence. The use of piano and percussion, performed live alongside electronic sounds, means that there will also be a degree of scene inconsistency created by the audible differences between naturally produced sounds and electronically produced sounds. This will further diminish the composition’s capacity for believable illusion. The use of loudspeakers surrounding the audience, however, supports the illusion of presence by creating a sense of immersion. With respect to Lombard and Ditton’s (1997) encapsulation of presence as transportation, it is not entirely clear within Stockhausen’s discussion of his attempts at creating depth whether he aims to create a sense of “it is here” or “you are there”. Perhaps the composer’s clarification may have helped the verisimilitude. Other factors that have the capacity to affect the success of the illusion include the complexity of the sound scene. Stockhausen’s technique of the layering of six different abstract sounds, which are then sequentially stripped to reveal layers, may inhibit spatial illusion through sonic complexity.

The above comments do not address any specifics concerning the synthesis of the spatial attribute of distance. The simulation of distance informed by a scientific understanding of perception results in an approach that has very little in common with Stockhausen’s techniques. Not only is the approach very different, but it
necessarily results in simulating other spatial attributes. In other words, the scientifically informed simulation of a singular spatial attribute results in the simulation of others. The perceptually centred effort to simulate distance has been known to musical efforts at least since the work of John Chowning (1977). Chowning not only simulated changes in intensity, as sounding objects move closer to or farther from the listener; he also simulated changes in the relative loudness of the direct sound coming from the sounding object, to its reverberant field caused by the environment (p.48). The reverberant field creates perceptions of the environment that indicate its size, shape and the kinds of materials used within the environment (p.48). A more empirical summary of the cues involved in creating the perception of distance, again within musical contexts, is given by Malham (2001a, pp.33-34). Malham adds that first reflections contribute to the perception of distance. The difference between the onset direction of the direct sound and onset direction of its reflections on walls helps determine how far away it is. However, first reflections also contribute information about the positions of walls within the environment. Again, the simulation of one spatial characteristic leads to the simulation of others. Another cue that contributes to the perception of distance is the loss of high frequencies caused by the humidity in the air (p.34). The simulation of this spectral change describes the levels of humidity in the environment climate.

What this account wishes to emphasise is that by describing more and more information concerning the general environment, the scientifically informed simulation of a singular attribute has a tendency to move towards the creation of presence. This is reflected in the characterisation of presence as immersion in stimuli (Lombard & Ditton 1997, section "Presence as immersion"). This tendency also finds resonance with Gibson’s ecological approach to perception. For Gibson the idea that perception is based on the processing of input stimulus is fundamentally flawed (1986, pp.251-253). He proposes that information is ‘picked up’ through its relation to the environment. This suggestion appears consistent with the observation that a better synthesis of an isolated attribute results in a concern for modelling more information describing various aspects of its environment. These different approaches, informed by the science of perception, corroborate that there is a relationship between the pursuit of realistic synthesis of isolated attributes, and the
more encompassing quality of presence. Considered within spatial music: by embracing a scientific understanding of perception the composer necessarily moves towards a concern with illusion more generally. This development suggests that one possible reason for the contemporary interest in spatial verisimilitude is the focus on a scientific understanding of perception that is captured within electronic and digital technology.

However, perhaps the most interesting insight that arises from the juxtaposition of Stockhausen’s techniques with the scientific understanding of presence, concerns an unexpected parallel between music and spatial verisimilitude. Whilst Kontakte can be criticised for poor adherence to perceptual realism, other conceptions of presence indicate that it may still have a high level of presence. In their development of the concept of presence Lombard & Ditton identify other approaches that represent “one or more aspects of what we define here formally as presence: the perceptual illusion of nonmediation” (1997, Section “Presence Explicated”). One such approach concerns the psychological component of immersion. As opposed to perceptual immersion, psychological immersion creates a sense of being “involved, absorbed, engaged, engrossed” (Section “Presence as immersion”). This aspect of presence is not created via a technological reality engine but rather through techniques that help the perceivers suspend disbelief and invest themselves in the work experienced. In a similar vein Biocca and Levy (1995, p.135) draw a parallel between “compelling VR experiences” and “reading a book in a quiet corner”. They suggest that authoring presence may involve helping audiences reject external stimuli, rather than be convinced by realistic and convincing stimuli: “In successful moments of reverie, audience members ignore physical reality and project themselves into the story space” (p.135). Such accounts advocate that music may itself, and independently of technologically realised illusion, have a capacity to create presence.

If both music and spatial verisimilitude are understood to have a capacity to create presence then one possible source for the tension between them lies in the opposition of their approaches. Spatial verisimilitude creates presence through perceptual realism, whilst music creates presence through immersion in musical experience. In such a conception, the source of the tension between them lies not in having
fundamentally different concerns, but rather in having very **different approaches to similar concerns**. This conception also potentially sheds light on the composer’s intuitive interest in spatial verisimilitude: there is a sensed similarity between the effects of spatial illusion and engagement with musical listening.

The suggestion that all music is concerned with presence may seem disingenuous but such a statement is contingent upon what is understood by ‘presence’. Within the context of this thesis, presence is understood as the illusion of non-mediation. This definition is designed to serve enquiry concerned with mediated environments, which includes the pursuit of spatial verisimilitude, but its underlying sense approaches notions appropriate to enquiry on art. One recent attempt to consolidate various understandings of ‘presence’ originating from different academic disciplines is made by Preston (2007, pp.278-279). It is Preston’s reach into the “consciousness literature” (p.278) that reveals an approach to presence, understood as **self-referential awareness**, that allows further insight into how Stockhausen’s *Kontakte* may be understood as creating presence. Preston cites Hunt as a key reference in this conception. In his text *On The Nature Of Consciousness* (1995) Hunt discusses two different forms of symbolic cognition that lead to conscious awareness (p.41). He calls these *presentational* and *representational* symbolisms. He cites language as an example of representational symbolism (p.42), but it is his notion of presentational symbolism that is of significance here.

In the presentational symbolism [...] meaning emerges as a result of an experiential immersion in the expressive patterns of the symbolic medium. It appears as spontaneous, preemptory imagery and is fully developed in the expressive media of the arts. Here, felt meaning emerges from the medium in the form of potential semblances that are “sensed,” polysemic and open-ended, and so unpredictable and novel. (p.42)

In other words art, and thus music, has a capacity to create self-referential awareness by immersing the listener in symbolic meaning. Considered within music, this is a form of presence that is not directly associated to the realistic projection of auditory stimuli, but rather to the ‘felt meaning’ that emerges from musical form. Hunt describes this meaning as ‘polysemic and open-ended’: it consists of many different
possible interpretations. This characterisation approaches the description of the aesthetic experience of music, but understands it as a form of conscious awareness.

Preston states that a “receptive, observing attitude allows such meaning to emerge” (2007, p.278). Within this conception of presence, Stockhausen’s lament about the audience’s reluctance to “perceive this music in its real terms” starts to seem much more reasonable.

An important question arises out of this characterisation: can these two approaches to presence support each other, or do they act independently of each other? Stockhausen’s expressed disappointment with listener’s perceptions of depth suggests that Kontakte’s musical dimension did not contribute to the perception of spatial illusion. Here at least, it does not seem that immersion in musical symbolism might support the effect of perceptual realism. It is the consideration of the inverse of this relationship, however, that strikes at the heart of the research question: would a more sophisticated level of perceptual realism have supported Kontakte’s musical dimension? This question seeks to understand precisely to what extent perceptual realism might have the capacity to support musical affect.

Hunt states that the two forms of symbolic meaning; presentational and representational, are necessarily intertwined (p.42) but this does not speak for the ability of one to serve the interests of the other. If both music and spatial verisimilitude have a concern with presence, is it the same presence? Perhaps perceptual realism has a capacity to create self-referential awareness that has a different quality to that created by a specific musical expression. Given different qualities of presence, would the successful development of one necessarily result in the weakening of the other? Perhaps the success of composers’ engagement with spatial verisimilitude is a function of their ability to meaningfully align the conscious awareness cultivated by each one.

Regardless, what this discussion highlights is that by characterising both music and spatial verisimilitude as having different approaches to similar concerns the underlying tension between them starts to develop a conceptual grounding. A
plethora of specific questions and research paths immediately emerge. The continuing exploration of the dimension of ‘presence’, in its breadth of meanings, is likely to shed further insights.

In conclusion, the notion of fidelity, within music, has always been intimately tied to auditory illusions. In the time of Edison, illusions consisted of creating the impression that the performer is present within the current listening space. Contemporary expectations of fidelity place importance on the illusion that the listener is present in the performer’s space. The difference between the two is that the contemporary concern seeks the experience of the illusion of space. A common concept for describing the success of such illusion is the notion of presence.

In pursuing ever-more realistic representations of spatial attributes, composers have moved towards an interest in spatial illusion. Compositional concerns have thus somewhat aligned with contemporary expectations of fidelity in music. Spatial illusion, however, introduces a set of qualities that are far from the isolated spatial attributes, such as direction, hereto explored compositionally. Composers must now consider the compositional potential of qualities such as presence and immersion. Herein lies an unexpected correspondence. Within the research concerned with mediated environments, presence is also understood to be a result of immersion within the expressive potential of the arts. This conception, in which music and spatial verisimilitude are understood to have similar concerns, offers one avenue for the exploration of the tension that exists between the two.
6 IMPACT OF THE TECHNOLOGICAL DIMENSION

Karlheinz Stockhausen’s composition *Kontakte* (1958-1960) has been described (Chagas 2008, p.191) as “a historical work that became a reference for the organic relationship between sound synthesis and sound space”. It was realised at the West German Radio and Television Broadcasting’s (WDR) Studio for Electronic Music and made use of a purpose-built rotating table, called the *Rotationstisch*. This apparatus employed four radially arranged microphones that captured sounds from a central rotating speaker (p.191). “As the table is turned, the loudspeaker on it projects a sound in different directions” (Chadabe 1997, p.41). The signals from the four microphones were recorded to four-track tape and later routed to speakers similarly aligned to the original microphones. The *Rotationstisch* is thus a device for “achieving sound distribution in a concert hall” (p.41). Zvonar describes it as a device used “to create the effect of sounds orbiting the audience” (2004b).

![Figure 21. Karlheinz Stockhausen operating the *Rotationstisch* in the Studio for Electronic Music of the WDR, Cologne, in 1959 (Chadabe 1997, p.41)](image)

As identified in the discourse of electroacoustic composition, discussed below, the *Rotationstisch* holds significance not just in its ability to distribute sounds, but also in
its mediation of the act of composition. The physical design and material reality of the device has not just left traces within the composition, it has also exposed and arbitrated compositional possibilities. These possibilities are highly gestural, allowing the composer to articulate spatial gestures.

Manning (2006, p.87) highlights that the *Rotationstisch* did not produce a smooth sense of rotation. He attributes this to variation in sound concentration caused by the rotating speaker as it straddles a pair of speakers then moves in line with each speaker thereby creating ‘bursts of sound’. This effect may have been emphasised by the directional characteristics of the microphones used, which may have provided little off-axis response. Manning (p.87) also identifies the presence of a Doppler shift; a change in pitch caused by the movement of sounds towards or away from the listener. This effect is created as the wall of the rotating speaker cone moves closer to, then away from, each of the four microphones. Manning argues that the Doppler shift contributes to the sense of movement. However, while a Doppler shift is undeniably a sonic artefact of movement, it should not exist when a sounding object is orbiting around a central listener since there is no change in the distance between the sound and the listener. The design of the *Rotationstisch* was therefore limited in its ability to represent orbiting sounds, but it was also responsible for the creation of a range of “‘interesting’ effects that are embodied in the aesthetics of the composition” (Chagas 2008, p.191).

Jonty Harrison, a British electroacoustic composer, argues that the *Rotationstisch* creates an illusion of space that is superior to, despite predating, the technique of panned mono sources (1998, p.126). He attributes this to the use of synchronised four-track tape recorders that effectively allows the capture and playback of “real phase information” of a “real environment” (p.126). In other words, the *Rotationstisch* did not just capture an orbiting sound; it also captured information relating to the acoustic of the room it was housed in. The faithful reproduction of that phase information is attributable to the synchronous operation of four-track recorders. Again, whilst the *Rotationstisch* was limited in its ability to capture orbiting sounds, the technique used resulted in the reproduction of the spatial quality of the room it occupied.
Manning (2006, p.87) also notes that the Rotationstisch “was a physical performance tool, operated directly by the hand actions of the composer, rotating the table faster or slower, and subject in addition to the frictional forces and the inherent inertia of the loudspeaker table itself”. The weight of the table, the smoothness of its operation and the physicality of the operator each contribute in determining the resultant sounds, thus helping to define the spatial gestures captured.

These accounts illustrate that the techniques and technologies used to realise a composition are, in the words of Agostino Di Scipio, “more than mere means” (1995a, p.371); they play a significant role in determining the aesthetics and form of the final work. As a technology, the Rotationstisch both limited how sounds could be projected into space, and also introduced new spatial effects. It might be clearer to say that the Rotationstisch framed possibilities for the explorations of space; within the rotating table’s mediation there is both restriction and the creation of possibilities. As is elucidated in this chapter, the resultant composition can be understood as a product of the composer’s relationship with this technological framing. The success of the Rotationstisch’s use in Kontakte can thus be seen as a testament to Stockhausen’s ability to engage with the skew of the technology, allowing it to take meaning within his expression.

This chapter presents three main points: firstly, it outlines one important approach to the technological interpretation of electroacoustic music; secondly, reality-equivalent technologies are examined for how they might frame the composer’s thinking and shape a composition’s aesthetic concerns and musical form; thirdly, the relationship between art and technology is considered at a more abstract level, in an attempt to understand how the composer might approach and negotiate reality-equivalent technologies.

Within this enquiry is an examination into how the technical pursuit of spatial verisimilitude affects the compositional act. It is argued that reality-equivalent technologies introduce some radical mediations of the compositional act that challenge traditional musical form, and orient music in a very different direction.
6.1 Discourse on the technological interpretation of electroacoustic music

There are many accounts of the importance of technology on electroacoustic music. Some examine the compositional legacy of chosen composers. For example: Manning (2006) and Clarke (1998) look at the work of Stockhausen; Georgaki (2005) and Hamman (2004) examine the work of Xenakis; and Palombini (1998) considers Schaeffer. Other examples include Emmerson (1998b) and Teruggi (2007). These accounts vary in their approaches, but broadly subscribe to the thesis, as worded by Manning (2006, p.81), that “the functional characteristics of the equipment available […] materially influenced the ways in which composers developed their compositional aesthetic”. Other accounts take a more abstract approach, aiming instead to qualify the relationship between music and technology more generally. Examples include (Boulez 1978; Becker & Eckel 1995; Hamman 2002). Of this later approach the work of Agostino Di Scipio, Italian composer and writer, stands out in its depth of consideration and solid theoretical foundation. Di Scipio’s approach, and its grounding in a philosophical understanding of technology, forms the springboard for this chapter.

Di Scipio (1995b) argues that electroacoustic music challenges the form/material dualism inherent to a classical interpretation of music; where composition is seen as the temporal organisation of sound material. His discussion is supported by the presence of techniques that treat both material and form homogenously. He cites techniques such as Xenakis’ stochastic synthesis, FOF synthesis and granular representations of sound (pp.39-40). Di Scipio also identifies the impact that such techniques have on the aesthetics of composition thus articulating the importance of technē: “the realm of techniques and technology captured in the creative process of music composition” (1995a, p.371). His thesis is that the study of electroacoustic music “needs to undertake a critique of the technical processes of composing and a characterisation of the technological working environment” (p.380).

Di Scipio outlines two core lines of inquiry which “require from us a profound critical insight into the technology of music” (1997, pp.63-64). The first echoes the above stated thesis and consists of the relationship between the composer’s work and
the techniques of its production. The second is a broader inquiry questioning to what extent electroacoustic composers have taken consciousness of the pervasive role of technology within their practice. This inquiry challenges the composer to consider how technology, rather than serve a musical intent, might partly determine it. It is an enquiry that reverses the standard interpretation of technology where it is seen as something that serves, and instead casts it as something that shapes. Di Scipio (1997, pp.72-76) thus advocates a perspective on music technology that he calls “subversive rationalisation” in which the composer takes consciousness of the cognitive, aesthetic and social values embedded in technology. He illustrates the importance of the subversive approach by pointing out that both early movements of early electroacoustic music; musique concrète and elektronische Musik arose out of the re-appropriation of technology designed to serve other aims (pp.73-74). Within this re-appropriation lies a subversive agenda in which the original aims of the technologies used were understood and disregarded.

Contemporary music producers and sound artists, whether experimental or otherwise, continue this tradition of re-appropriating technology to aesthetic effect. Artists such as Chris Watson, Jana Winderen, Carsten Nicolai, Mark Fell, Taylor Deupree, Ryoji Ikeda, Russell Haswell and many others explore technology as a form of black box that harbours unknown musical potential. Engaging with these technologies then becomes an exercise in discovering what can be revealed from within them. This approach is not to be confused with the contemporary composer’s engagement with the pursuit of verisimilitude.

When reality-equivalent technologies are used to serve a composer’s interest in the realistic illusion of space, the technology is not being re-appropriated. Rather, it is simply, or even ‘blindly’, being employed to pursue what it purports to deliver: realistic illusion. There is therefore a lack of a subversive agenda, to use Di Scipio’s terminology, in the spatial composer’s engagement with reality-equivalent technologies. The motivator for using them is the pursuit of verisimilitude, rather than the promise of new musical expression born of their inherent contingencies and biases.
This chapter precisely seeks to uncover what lies hidden within these technologies. In line with Di Scipio’s expression the following question is explored: What are the cognitive, aesthetic and social values embedded in reality-equivalent technologies? It is a question that rejects the characterisation of these technologies as neutral tools, and seeks to understand how their use might affect, influence or indeed partly determine the composition. This understanding then allows the composer to interact with and transform the technology’s ‘shaping’ rather than be subsumed by it. Expressed in the words of Di Scipio:

Composing always includes a personal view and an interpretation of the technological environment. It includes an interaction with and a transformation of the knowledge embodied in the techne by which we compose. (1997, p.78)

In his enquiry, Di Scipio consults arguments found within the philosophy of technology. His concept of ‘subversive rationalisation’ (p.72) is adapted from the writings of Andrew Feenburg, but the core of his approach stems from an interpretation of 20th century philosopher Martin Heidegger’s essay, written in 1949, titled The Question Concerning Technology (Heidegger 1977). Other researchers writing on the technological dimension of electroacoustic music have also referenced this essay (Palombini 1998; Hamman 2002; Manning 2006).

Of course, Heidegger is by no means the only philosopher to have considered technology and its effects on humanity. Other prominent 20th century philosophers who have contributed research on technology include John Dewey, Jean Baudrillard and Lewis Mumford. Contemporary philosophers continue to address the expanding role of technology and these include Bruno Latour and Bernard Stiegler. Aspects of each of these philosophers’ work could be applied to research concerned with the role of technology within music. One promising critical framework is known as Actor Network Theory (ANT). ANT, largely developed by Latour (1992) and his colleague Michel Callon (1991), posits that non-human actors must be taken into account alongside human actors to understand assemblages of aligned interests known as ‘actor networks’. If one considers contemporary music production as an actor network, then music technology can be cast as one of the contributing non-human actors. There is already some work taking this approach. Prior (2008), for
example, suggests that ANT can help cater for some of the failings of Bourdieu’s field theory in understanding the contemporary music known as ‘glitch’. Again, by casting technology as a ‘non-human actor’, the development of ‘glitch’ as a music aesthetic is better understood.

In the interests of brevity, philosophical enquiry into the significance of technology is here largely limited to Heidegger’s writing. This allows focusing Di Scipio’s existing work, which considers music technology broadly, onto one type of music technology: the type that purports to deliver realistic illusion in spatial audio. In the following section, some consideration is first given to interpreting Heidegger’s essay. This analysis then forms the ground for teasing out the values embedded in reality-equivalent technologies and understanding their impact on composition.

6.2 Reality-equivalent technologies as Bestand

In The Question Concerning Technology (1977) Heidegger seeks to expose the essence of technology. He begins by observing that the common conception of technology is that it is a simple means to an end that is executed as a human activity. Whilst this characterisation, encapsulated in the word equipment, is correct, he argues that it does not reveal technology’s essence. For this he turns to Classical Greek philosophical thought and cites Plato: “Every occasion for whatever passes beyond the nonpresent\(^\text{18}\) and goes forward into presencing\(^\text{19}\) is poiesis, bringing-forth” (Plato cited in Heidegger 1977, p.10). Heidegger then examines the meaning of the word ‘bringing-forth’ and finally arrives at the essence of technology (p.11):

\[
\text{Bringing-forth brings hither out of concealment forth into unconcealment.} \\
\text{Bringing-forth comes to pass only insofar as something concealed comes into unconcealment. This coming rests and moves freely into what we call revealing.}
\]

For Heidegger, technology is a way of revealing; that is, it operates in the realm of truth (p.12). Nothing new is made. What has come forth has only moved from being hidden to being shown. This view denies that the essence of technology lies in creating new objects or finding new ways of manufacturing or delivering new ways

---

\(^{18}\text{nonpresent can be understood as ‘what is not here’}\)

\(^{19}\text{presencing can be understood as ‘bringing forth into being’}\)
of communicating. In essence, technology is something that shows things previously hidden. In Heidegger’s words, it ‘unconceals’.

Di Scipio is concerned with identifying that which is unconcealed by the technologies used in composition. What is unconcealed by technology will be present within the expression of the composition, and so it is imperative for the composer to take consciousness of that unconcealment, and respond to it. Failure to do so may result in the expression of the technology, at the expense of the expression of the composer.

The question that is difficult to answer, however, is what exactly is unconcealed by a technology? Another line of inquiry in Heidegger’s essay throws light on this question. He cites the four causes, which have been known to philosophy for centuries (p.6), responsible for bringing something about. The _causa materialis_ is the material out of which something is made; the _causa formalis_ is the form which the material takes; the _causa finalis_ is the intended final object and the _causa efficiens_ is the agent that brings these together. Heidegger gives, as an example, a silver chalice. Silver is the _causa materialis_; it is the material out of which the cup is made and brings with it certain properties, both physical and social. The cup, its handle and base are the _causa formalis_; the chalice needs to hold liquid and be handled. The ritual being served is the _causa finalis_; it is the need that the cup satisfies. The silversmith is the _causa efficiens_; responsible for bringing together the causes and producing the final cup. Heidegger explains that the chalice is indebted to the silver, and also indebted to the form it has taken. Whilst the chalice is also indebted to the silversmith, it is ultimately not the silversmith that has brought about the chalice, but rather the _causa finalis_; the ritual that has provided the need for such a chalice. It is the _causa finalis_ that “sets the standard for all causality”, not the silversmith (p.7).

At this point, one can ask: what are the four causes of reality-equivalent technologies? The _causa materialis_ could be interpreted as the digital medium, the interpretation of signals as binary digits, and perhaps include some electronics. The _causa formalis_ might be a software package, with a user interface, and perhaps include the algorithms developed by engineers. The _causa efficiens_, the agents who
bring the other causes together, might be a mix of audio engineers and software developers. The *causa finalis*, which represents the original need for the resultant technology, is the most difficult question to answer. What need do reality-equivalent technologies satisfy? Why do they exist? Do they exist as a result of a need expressed by composers? Do they exist to serve the needs of audiophiles interested in the quality reproduction of recorded sounds? Do they exist to satisfy the engineer’s mind to find elegant solutions to difficult problems? Do they exist to satisfy an intrigue with the sonic illusion of reality?

One answer to these questions might be: all of them, and maybe more. The difficulty in answering this question with any precision suggests that reality-equivalent technologies belong to the group of technologies that Heidegger labels ‘modern technologies’. The chalice is very different to reality-equivalent technologies. It has a clear and very specific need as its *causa finalis*. However, there is a key and critical distinction between the chalice and these technologies, which is that the *chalice is not a technology*; it is the result of technology.

There are other framings of the four causes in which reality-equivalent technologies hold a much clearer role. For example, one might understand the *causa finalis* as the resultant musical work produced with reality-equivalent technologies. In such a formulation reality-equivalent technologies have acted as the *causa formalis*, and the *causa efficiens* could be understood as being the cultural purpose of the musical work.

However, it is precisely the consideration of technology as a thing in itself, divorced from any specific end result, which Heidegger is concerned with in his essay *The Question Concerning Technology*. Reality-equivalent technologies exist independently of any clear and specific *causa finalis*. They serve a realm of possibilities that aim to satisfy a range of very different needs from compositional endeavours to interactive installations to high quality sound capture and playback. Heidegger calls this realm of possibilities ‘standing-reserve’ (1977, p.17) and describes it as an attribute of ‘modern technology’. Standing-reserve’s principal characteristic is its disassociation from the thing it will ultimately bring forth:
Causality now displays neither the character of the occasioning that brings forth nor the nature of the causa efficiens, let alone that of the causa formalis. It seems as though causality is shrinking into a reporting [...] of standing-reserves [...] (p.23)

This can be illustrated by considering how reality-equivalent technologies are indebted to, as in ‘caused by’, the intent (or the causa finalis) of a composition. Simply put, they are not. Their development exists largely outside of compositional concerns. Instead it is more likely that it is the composition that is indebted to the intent of the technology. Compare this to the silversmith’s hammer, which, as a technology, is indebted to the causa formalis of the chalice. It is not the design of the hammer that has given rise to the shape of the chalice; it is the shape of the chalice that has given rise to the design of the hammer. The danger of standing-reserve is that technology assumes an importance disassociated from any causa finalis. Heidegger explains that instead of being indebted to the causa finalis, standing-reserve ‘enframes’ it; that is, standing-reserve challenges the causa finalis to fit within its own bounds (p.19). Thus the composer’s compositional intent is challenged to fit within the enframing of reality-equivalent technologies. Heidegger finally explains why this is dangerous:

The rule of enframing threatens man with the possibility that it could be denied to him to enter into a more original revealing and hence to experience the call of a more primal truth. (p.28)

Hubert Dreyfus, a scholar on Heidegger, interprets Heidegger’s warning about the danger of modern technology in simpler terms:

The danger, then, is not the destruction of nature or culture but a restriction in our way of thinking – a leveling of our understanding of being. (1997, p.99)

Di Scipio’s call to understand the values embedded in technologies can thus be interpreted as asking: how do reality-equivalent technologies result in the restriction of the composer’s thinking? The use of the word restriction within the same sentence
as the term reality-equivalence begins to hint that perhaps these technologies restrict the composer to using material that is equivalent to reality. Indeed, this line of enquiry is confirmed and elaborated below.

Whilst Heidegger does not specifically consider the use of technology for the purpose of art, the concept of art does assume a very important role towards the end of his discussion:

Because the essence of technology is nothing technological, essential reflection upon technology and decisive confrontation with it must happen in a realm that is, on the one hand, akin to the essence of technology and, on the other, fundamentally different from it.

Such a realm is art. But certainly only if reflection upon art, for its part, does not shut its eyes to the constellation of truth, concerning which we are questioning.

(1977, p.35)

And so whilst Di Scipio proposes to question technology in the interest of art, Heidegger proposes that art might help question technology. Herein lies a resonance that is explored in the last section of this chapter in which the relationship between art and technology is considered at a more abstract level. Having elaborated Heidegger’s characterisation of modern technology, we are now in a position to uncover the values embedded in reality-equivalent technologies.

6.3 Embedded values of reality-equivalent technologies

Reality-equivalent technologies un-conceal the pursuit of realistic spatial illusion as a primary concern. Before the composer has even attempted to engage with the tools of reality-equivalent systems, spatial verisimilitude has assumed a central role. Reality, which has historically been the subject of diverse musical interpretation (Dahlhaus 1985), is now implicitly interpreted as realistic representation. It does not matter what is represented, as long as the representation moves towards illusion through its realism. This theoretically leaves the subject of the representation as the concern of the composer.
However, research into the manufacture of illusions suggests that the subject of the representation does have an impact on its potential to be perceived as real. This research, introduced below, indicates that spatial verisimilitude does have a concern with the sound material that is spatialised and thus, the use of reality-equivalent technologies has an impact on the composer’s choice of sounds. In line with Heidegger’s characterisation that modern technologies restrict or limit our understanding of being, reality-equivalent technologies can be understood to limit the composer’s musical conception to sounds and spatio-temporal gestures that have the capacity to be perceived as real. This point is succinctly expressed by Malham, in this passage already quoted in Chapter 1:

For a system to produce reality-mimicking sound, it is necessary for all departures from reality equivalence to be below the relevant thresholds of perception. (2001a, p.32)

This understanding raises the question of what might constitute a ‘departure from reality equivalence’. Of specific interest to the composer is whether or not music might in itself represent a departure from realism in sound. Chapter 3 has already uncovered that musical grammars have the capacity to reference space, which can interfere with spatial information presented by reality-equivalent technologies, but the question being asked here is different. The question is not: what space exists in music, but rather: how ‘realistic’ is music? For example, a melody might remain reality-equivalent if it is perceptually tied to an instrumentalist. But can a melody voiced by an abstract sonic texture be considered a departure from reality-equivalence? Expressed thus, and as is elicited below, reality-equivalent technologies don’t just limit the composer by restricting sounds to those that have the capacity to be perceived as real, they also ask the composer to question whether music might itself present a challenge to realistic representation.

As discussed in Chapter 5, by engaging with reality-equivalent technologies, the composer partly transfers responsibility for the illusion of reality over to the science of perception and its technological embodiment. However, these systems still remain limited in their ability to author convincing auditory illusion. Malham (2001a, p.37) stated a decade ago that no systems yet have the capacity to project sounds that are
indistinguishable from reality, and there has been little evidence since that this mark
has been met. In practice, whilst reality-equivalent technologies inspire with the
promise of the illusion of reality, the composer’s accessibility to this illusion is
limited. Reality-equivalent technologies thus mediate the act of composition in two
significantly different ways; the conceptual centres on the potentiality of convincing
realistic illusion; but the execution is mediated by limitations in the available
systems.

These limitations, in the available systems, have two sources. The first concerns the
design that underlies reality-equivalent technologies, taking a specific perspective on
what constitutes sound in space. One such limitation is the logical separation of a
sound and its spatialisation, discussed below. Another example is the tendency to
model sounds in space as point sources which, whilst non-existent in reality, are
simple to model mathematically. Here, the ease of implementation, which is a
function of the physical model of sound that underlies different reality-equivalent
technologies, influences the aspects of space that software developers choose to
expose to the composer.

The second source of limitations concerns the practicalities of hardware and software
solutions. Due to implementation complexity, and limited computational resources, the
simulation of the spatiality of sounds favours spatial qualities that can be easily
produced, such as movement. In other words, the kinds of spatial characteristics and
gestures available to the composer are partly determined by the implementation
facility and computational accessibility of certain spatial audio cues.

A selection of these technological mediations are considered and explored in greater
detail below. Other considerations consist of such things as control over parameters,
such as location, and whether they have a physical embodiment or are limited to
software modelling.

---

20 One might argue that modern computers are less and less limited. However, as
described in Chapter 2, the portfolio of works involved software tools and techniques
that could take up to twelve hours to render ten minutes of spatial sound. The real-
time realistic simulation of the behaviour of sound in space still requires a great deal
of computational resources.
6.3.1 Logical separation of sound and its spatialisation

One of the most significant impacts in the conceptual processes of the composer using reality-equivalent technologies involves the logical separation of a sound and its spatialisation. As already expressed in Chapter 3, audible sounds can only exist within the context of a physical space, so the premise that a sound can exist before being spatialised is fundamentally erroneous, and brings significant implications. The sound being spatialised may be synthesised, processed or recorded. Recorded sounds will always hold a spatial imprint. This imprint concerns not only the acoustic of the room of the recording but also the orientation of the sounding object relative to the microphone. The spatialisation of recorded sounds presents a potential conflict of spatial information between the spatiality captured in the recording, and the spatiality simulated in the act of spatialisation. Indeed, as expressed by Worrall:

> Sound is not an abstract ideal which is projected into 3-space. The space is in the sound. The sound is of the space. It is a space of sound, but there is no sound in space. (1998, p.97)

The spatialisation of synthetic or processed sounds holds its own implications, considered next.

6.3.2 Timbral modulations and other sound processing techniques

Timbre is an important component of electroacoustic composition that is intimately tied to the perception of space. The relation between the musical use of timbre and perceptions of space, in which timbral modulations have the capacity to act as references of space, is introduced in Chapter 4. What is of concern here is how the musical use of timbre might contradict or conflict with spatial verisimilitude, and how this might affect its musical use. Worrall explains:

> Both timbre and spatial cues depend on the morphology of the spectral modification (emphasising some frequency bands whilst attenuating others) of the sound source… Clearly the two interact; modifications of timbre (including all its sub-dependencies such as fundamental frequencies, amplitudes, spectra, etc.) affect our perception of different aspects of location: proximity (distance), angle
Thus, to maintain realistic representation in sound the composer will need to be careful and conservative with the use of timbral modulations. This challenges the composer to be conscious of how processes, which affect the timbre of sounds, might change perceptions of space. British composer Denis Smalley highlights that typical sound processing techniques can affect perceptions of space:

But, quite often composed space is created through artifacts or spatial byproducts of the sounds, textures, and processing techniques you are using [...] So, one has limited power to control space as such, because the perception of space is the sum of many interrelated features. (Austin 2000, p.14)

Smalley’s comments concern the practice of diffusion, introduced in Chapter 3, in which it is implicitly understood that sounds will have a spatial character before being projected spatially. Considered in the context of reality-equivalent technologies, Smalley’s characterisation of the impact of ‘sound textures and processing techniques’ on the perception of spatial information highlights the problem of the logical separation between a sound and its spatialisation. Processing techniques, common to electroacoustic composition, and applied to a sound prior to its spatialisation may interfere with the success of that subsequent spatialisation.

Thus to maintain realistic spatial representation the composer needs to be concerned with and attentive to timbral modifications specifically, and to the processing of sounds generally. This attention, if respected, will have a significant impact on: firstly, the composer’s choice of sounds that are subsequently spatialised; secondly, how those sounds might be processed; and thirdly, how successful the resultant verisimilitude will be.

6.3.3 Direction as the principle attribute

The invention of ambisonics, a reality-equivalent technology briefly introduced in Chapter 1, is primarily attributed to British audio engineer and mathematician Michael Gerzon. Gerzon’s understanding of spatial audio can shed light on the
values embedded in the technology he pioneered. In a paper titled *Criteria for evaluating surround-sound systems*, Gerzon argues that:

>[The] ultimate aim of surround systems is to provide a good illusion of an intended encoded directional effect (1977, p.400)

Direction thus takes on a key importance in the value system of ambisonics. Indeed, the mathematics at the core of ambisonics (and WFS) is essentially concerned with the encoding of direction. Other spatial attributes such as distance, early room reflections and reverberation need to be modelled *outside* of the encoding of direction. The direction of sound thus assumes an importance, in reality-equivalent tools, that insists composers consider direction above all other spatial parameters.

The bias that reality-equivalent technologies hold towards direction becomes significant for the composer, given that there have been questions concerning the importance of direction for listener preference. Rumsey (2002, p.653) reports that “recent experiments seem to suggest that precise source position rendering in the spatial reproduction of music and other natural signals is not the most important spatial factor governing listener preference”. These accounts reference spatial attributes that listeners would *prefer* to hear, not those that most contribute to the creation of verisimilitude. That said, the kinds of spatial attributes that are preferred by listeners are those that increase the sensation that they are “present” at a musical event, and these include such qualities as surround immersion in sound (Berg, J & Rumsey 2000, p.12). In exposing direction as the primary spatial attribute, reality-equivalent technologies place greater importance on a spatial quality that evidence suggests is not favoured by listeners.

6.3.4 Sounds primarily modelled as points

In (1998, p.167) Malham asserts that “real sound sources” have “extended sound-emitting surfaces” that have a “more complicated behaviour” than the “hypothetical point source”. Similarly to direction, the mathematics of ambisonics and WFS make it very easy to spatialise sounds modelled as point sources. Composers are thus typically limited to considering sounding objects as point sources that project the
same sound in all directions. In this modelling, sounds have neither dimension, nor orientation.

There have been a few documented efforts to move past the limitation of point sources (Malham 2001b; Menzies 2002; Potard & Burnett 2003, 2004), including my own work (Deleflie & Schiemer 2009, 2010) but the complexity involved in modelling non point sources currently makes it only a remote option for composers generally.

Xenakis’ *Eonta*, discussed in Chapter 3, can be seen as an example of spatial music that could not have eventuated if realised in today’s reality-equivalent technologies. This work explores changes in timbre caused by rotating brass instrumentalists. It is the instruments’ lack of homogeneity in directional sound projection that allows complex timbres to emerge as their orientation changes. The rotation of point sources of sound, as typically modelled in reality-equivalent technologies, would create no net effect.

6.3.5 Ease of simulation of certain attributes

Some spatial information will be simpler to simulate, and less computationally intensive. The simulation of the Doppler effect, for example, is computationally cheap and contributes to the perception of movement, particularly those that involve moving past the listener. This only requires slight modulations of pitch in accordance with changes of direction. The relative simplicity of creating the perception of movement through simulations of the Doppler effect means that movement will be one of the most accessible spatial attributes for composers. This has a significant aesthetic impact, in that composers will favour movements around the listener, and small modulations of pitch will be prevalent.

This has a secondary impact in that changes in pitch will potentially interfere with musical devices, such as melody, that predominantly assume a static position in space. The composer is thus challenged to either: accept and embrace the corruption of harmonic and melodic content; refrain from employing spatial movement with this
content; or sacrifice the realism in the representation of moving sounds by excluding the Doppler effect in their modelling.

6.3.6 Interpretation of reality is denied to the composer

It has already been argued that reality-equivalent systems imply an interpretation of reality that is concerned with realistic representation. Theories of the illusion of reality from the visual arts have been presented in Chapter 5. In one prominent understanding of illusion in the visual arts, historian and theoretician Gombrich (1977) argues that successful illusions arise when the artist applies the same expectations, of what appears real, as the perceiver of the work. In other words, it is the schemata shared between artist and beholder that allow the artist to successfully create illusions. When the illusion of reality is handed over to a scientific understanding of perception, and a computational simulation of the behaviour of sound in space, these schemata are partly disengaged because the composer is largely excluded from the details of the spatial modelling.

It is interesting to consider that within the practice of diffusion, control of the illusion of space is not handed over to the science of perception, but remains firmly within the schemata of the diffusionist, who manipulates the spatial projection in real time thus being able to respond to their hearing of the space. Differences between diffusionism and the use of reality equivalent systems, to project sounds in space, are discussed in Chapter 3.

There is considerably less scope, within reality-equivalent systems, for the composer to manipulate the representations of space in such a way to tweak the perception of reality. In other words, the mechanism described by Gombrich for the creation of illusion in art, is limited in reality-equivalent technologies.

6.3.7 Limitations of the scientific interpretation of reality

The design and implementation of reality-equivalent technologies is significantly informed by the scientific understanding of perception and its supporting empirical evidence. This understanding, however, brings its own set of limitations. Zielinski, Rumsey and Bech (2008), for example, identify and discuss biases in listening tests designed to measure quality in sound. Lennox, Myatt and Vaughan (2001, p.3)
question whether research into the perception of the direction of sounds remains valid once applied within real environments. Lennox and Myatt identify and cite further evidence:

Jarvilehto [10] considers that environment and percipient comprise a holistic system and that sensible conclusions about perception cannot be drawn from the study of perceivers in isolation. (2007, p.2)

The handing over of realistic representation in sound to the science of perception means that the composer’s access to this realism is limited by the practicalities of a scientific understanding. Malham provides one example of a possible perceptual mechanism that he believes has been disregarded by science:

Unfortunately, because of the difficulty of working experimentally on, for instance, chest cavity pickup or bone conduction mechanisms, little work has been published on these means of perception and their directional discrimination capabilities. Instead, because of the relative ease with which headphone-based measurements can be made, almost all the major studies of directional hearing have concentrated on headphone-presented information. (2001a, p.34)

Consider these purported perceptual mechanisms, not yet identified and explored by science, within light of Gombrich’s understanding of the artist’s process in authoring illusions. As described in Chapter 5, for Gombrich, the artist’s process is one of ‘making and matching’: what the artist makes is then perceptually matched against the artist’s expectations of the illusion. In such a process the role of the chest cavity, in contributing to the perception of direction, is inherently included. When informed solely by a scientific understanding of perception, only the mechanisms identified by science will be included.

6.3.8 Interaction with environment

This last example, of the technological mediation of composing with spatial verisimilitude, only indirectly concerns reality-equivalent technologies. By focusing the pursuit of verisimilitude on the scientific understanding of perception, reality-equivalent technologies project the composer towards new spatialisation techniques
that further leverage this understanding. One example of this is the compositional consideration of the possibilities offered by allowing listeners to interact with the environment. It is originally Gibson’s ecological approach to perception, in which the environment can be understood “as consisting of affordances for action” (Reed 1988, p.280), that suggests presence can be created by allowing the perceiver “to successfully act or behave in the environment” (Preston 2007, p.281). Catering for an interactive component, within a spatial composition, can therefore advance the pursuit of spatial verisimilitude. It is to be expected, then, that compositional efforts interested in the pursuit of spatial verisimilitude should begin to consider how interactive elements might be integrated into spatial composition. In this sense, compositional thinking can be seen as mediated by the development of techniques that create greater levels of spatial verisimilitude.

Here, two ways that spatial composition might introduce an interactive element are identified. Each has very different implications for spatial composition. The first consists of using head-tracking technologies to change the spatial orientation heard by a listener when they move their head. The second concerns designing geographically explorable soundfields in which the listener position is used to manipulate what is heard.

Head movements provide a subtle but perceptually significant action in the environment. Changes in heard sound as a result of head movements have been identified as a significant contributor to the perception of location (Blauert 1997, pp.43-44, 383). Thus, movement of the head can act as a ‘successful action in the environment’ when it is corroborated by changes in the perceived positions of heard sounds. This effect can be simulated when sounds are heard through headphones. By using head tracking techniques changes in head orientation can be used to correspondingly alter the orientation of the spatial simulation. It is a technique that, when well implemented, can be very successful in supporting perceptions of space (Blauert, 1997, p.383). In such a technique, the interactivity is subtle and the compositional impact is relatively low; the listener’s interaction does not affect the composition’s sound material nor its temporal dimension, but does affect its orientation. The resultant indeterminacy of the heard soundfield’s orientation will
need to be considered compositionally; there is potentially significant symbolic
difference in such things as hearing a sound behind the head as opposed to in front.

The second technique is described by Lennox and Myatt (2007) in their paper
*Concepts of perceptual significance for composition and reproduction of explorable
surround sound fields*. They propose to create explorable sound fields, in which the
listener may wander at will. Lennox and Myatt recognise the sizeable implications of
this approach:

> Radical changes are necessary with respect to both content generation and
> conception of sound designs, and in engineering techniques to implement
> explorable fields. An incidental benefit lies in the creative possibilities. For
> instance, a composer could create a piece with no ‘listening centre’ that can only
> be fully appreciated by ambulant listeners. We are currently exploring the
> aesthetic potential of this approach. (2007, p.211)

Such an approach introduces a significant change in compositional concerns. Here,
the fixed temporal structure of a composition is handed over to the listener’s freedom
to explore. The pursuit of the realistic representation of sounds, as informed by at
least one predominant theory within the perceptual sciences, compels the composer
to radically reinterpret the temporal dimension of their work. In composing for
‘ambulant listeners’ predetermined sequential organisation of sounds is no longer
possible, and alternate forms of temporal structuring need to be considered. The
composer must cater for a multitude of possible sequences, within which the listener
will draw their own linear timeline.

These two examples demonstrate that the pursuit of spatial verisimilitude can give
rise to new techniques and technologies that demand a radically different approach to
composition. In such examples, the pursuit of spatial verisimilitude’s mediation of
compositional concerns can be considered more of an opening up of possibilities,
than a restriction. That said, by demanding that the composer discard control over the
temporal dimension of a work, one might say that the act of composition is itself
challenged and the composer subsequently assumes a very different role. As
expressed by Paul Berg, a composer from the Netherlands:
I have no idea what the future of composition may be. Perhaps it will radically change due to the non-linearity of current media. Perhaps the shifting cultural paradigm will redefine the role of a composer. (1996, p.26)

The exploration of compositional possibilities exposed by the technological pursuit of spatial verisimilitude certainly seems to give credence to Berg’s speculation.

6.4 Composing with an understanding of technology

Given an appreciation for the important role of technology in composition, and an understanding of the embedded values, the restrictions, the enframing of reality-equivalent technologies, how is the composer to proceed?

Di Scipio (1997) touches on this line of inquiry again by leaning on Heidegger. In a lecture delivered in 1950, and published as the essay Language (1971), Heidegger explores a relationship that Di Scipio interprets as a parallel to the relationship between art and technology. This is the relationship between poetry and language. Poetry belongs to the arts and while language is not a technology, for Di Scipio, it is analogous to something very similar, technical process:

The difference between poiesis as revealed in the process of art and as revealed in the technical process is illustrated by Heidegger himself (1950) with an example regarding poetry: to the poet the language tells while to others it serves; the poet qualifies its medium, by putting it into question and showing its limitations; doing so the poet eventually enriches language by transforming it; the non-poet uses and exploits the medium within the given, existing boundaries, thus being spoken by language more than speaking it. (1997, p.70)

Di Scipio offers this interpretation in passing, but the analogy is enlightening and is worthy of a more detailed exploration. In his lecture, Heidegger denies a prevalent idea concerning language:

According to this idea language is the expression, produced by men, of their feelings and the world view that guides them. Can the spell this idea has cast over
language be broken? Why should it be broken? In its essence, language is neither expression nor an activity of man. Language speaks. (1971, p.194)

The notion that it is language that speaks, and not man, is challenging. The notion that language is not man’s expression is also challenging. The adaptation of these comments to this thesis’ concerns might express that it is reality-equivalent technologies that engage in musical expression, not the composer. Another way to phrase this would be to state that reality-equivalent systems use the composer to engage in musical expression. Another way again might suggest that reality-equivalent systems simply express their intent(s), or their *causa finalis*, without any concern for musical expression. Heidegger clarifies this cryptic statement by explaining the nature of man’s relationship to language:

Man speaks in that he responds to language. This responding is a hearing. It hears because it listens to the command of stillness (1971, p.194)

Language holds meaning that is not controlled by the person speaking. For the person to speak they must listen to the meaning of the words they have put together and subsequently respond by putting together different words. Projected onto the act of composition using technology, Heidegger’s elucidation might be worded thus: composers create music in that they respond by listening to the music-making of the technology. Here the nature of the relationship between the poet and language is slightly different to how Di Scipio has translated it to the composer and technology. Whereas Di Scipio (1997, p.78) sees composing as involving an ‘interpretation of the technological environment’, Heidegger sees poetry as the *responding* to language. Interpretation requires listening and re-presenting, whilst responding requires listening and answering. For Heidegger, the responding ‘hears because it listens to the command of stillness’ (Heidegger 1971, p.207). In other words, the listening does not judge; it does not interpret; it is still.

There is a symmetry between Heidegger’s notion of listening and responding and Gombrich’s (1977, p.24) concept of ‘making and matching’ introduced in Chapter 5. Gombrich argues that the artist does not represent real objects by observing and copying their appearance, but rather by placing strokes on the canvas (making), then
stepping back and observing if those strokes fit the subject being represented (matching). It is an iterative process. Gombrich’s ‘matching’ can be considered symmetrical with Heidegger’s ‘listening’. For Gombrich, the artist ‘responds’, that is; adjusts the painting, by ‘listening’, that is; observing how their strokes fits their conception of the subject’s appearance.

Within both Heidegger’s account of poetry and language, and Gombrich’s account of the painter’s technique of representation, there is no tension between the artist’s aim and the technique. Instead, the poet and the artist resign to the speaking and the representing of what they have put forward, and they listen and observe it. It is how they respond to that listening and observing that allows the progression towards their intended expression. In other words, there is no manipulation, control or mastering of materials, techniques and technology. There is only response to it. By iteratively responding to the speaking of language, or the representations of strokes on the canvas, there is an incremental movement towards the intended expression.

Evidence of this characterisation also exists in compositional critique. Theodor Adorno, 20th century philosopher and musicologist, had a specific interest in the Schoenberg’s twelve-tone technique, introduced in Chapter 2. Adorno published a number of essays in a new Vienna-based musical journal titled Anbruch (Müller-Doohm 2005, p.112) in which he sought to defend Schoenberg’s radical new technique from its detractors. Within the formulation of that defence, Adorno develops a concept of progress “that would be valid for every branch of aesthetics that was based on the assumption of ‘historical appropriateness’” (Müller-Doohm 2005, p.112). The key characteristic of this concept involves the importance of engaging with the dialectics of contemporary materials and techniques. For Adorno “the freedom available to composers was no mere figment of the imagination, but was embedded in the dialectic of the material” (2005, p.112). In other words, the realm of possibilities does not lie with the composer’s ideas, but rather within the material served as a result of the historical process. For Adorno, the development of twelve-tone technique is such a material. Only the composer:
who submits to the work and seemingly does not undertake anything active except to follow where it leads, will be able to add something new to the historical constitution of the work, to its questions and challenges, something that does not simply follow from the way it happens to have been handed down historically. And the power to resolve the strict question posed by the work, by giving a strict response to it, is the true freedom of the composer (Müller-Döohm 2005, p.112) ²¹

Here, the composer’s freedom to express is characterised as the result of a response to the materials at hand. For Adorno, the composer must engage with the dialectic of those materials, but may respond to them however they see fit. This characterisation resonates with both Heidegger’s understanding of expression through language and Gombrich’s understanding of artists’ creation of illusions. The common ground exists in the notion that human expression results as the response to the expression of materials and techniques. In other words, human expression cannot exist independently of the dialectic of materials and techniques.

Only time will tell if reality-equivalent technologies, and their encapsulation of the pursuit of spatial verisimilitude, will ever be regarded as a significant part of the historical evolution of compositional concerns. In so far as the electronic medium has thrust audio perception as a central concern for electroacoustic music, and perceptions of space are intrinsic to audio perception, then spatial concerns will undoubtedly remain important. Regardless, the concepts developed in this chapter indicate that composers interested in pursuing realistic spatial illusion must respond to the expression of reality-equivalent technologies.

This brings us back to the question of the dialectic of reality-equivalent technologies. One way to examine this question involves returning to Plato’s notion of the causa finalis of reality-equivalent technologies. As already expressed, in abiding to Heidegger’s definition of modern technology, reality-equivalent technologies serve no clear singular aim; they serve a range of different purposes. However, within a musical endeavour, there are three characteristics that have a clear relation to

²¹This quote originates from an essay titled “Reaction and Progress” published in the Vienna journal Anbruch in 1930. Despite being referenced in several prominent texts on Adorno, I have found no complete English translation of it.
compositional concerns: firstly, reality-equivalent technologies express a human interest in creating and experiencing realistic spatial illusions in sound; secondly, the understanding of illusions is encapsulated technologically, and thus scientifically; and thirdly, the technologies that aim to deliver spatial verisimilitude have limited success.

Given the above summary, the dialectic of reality-equivalent technologies can be expressed through the musical exploration of how technology falls short of spatial verisimilitude. Such incidental sonic material could be explored in a similar way to Stockhausen’s exploration of the characteristics of the Rotationstisch. Compositional thinking would then have the opportunity to express the contemporary concern with spatial illusion in music, whilst responding to the underlying inferences of its technological embodiment.
7 CONCLUSION

The sophisticated control of this dimension [the sound landscape] of our sonic experience has only become possible with the development of sound-recording and synthesis and the control of virtual acoustic space via projection from loudspeakers. It would certainly be foolish to dismiss this new world of possibilities on some a priori assumption that they are not of ‘musical’ concern. In fact, any definition of musical activity which does not take them into account must, from here on, be regarded as inadequate (Wishart 1986, p.60)

In the twenty-five years since Wishart wrote these words, spatial design tools for composers have increased in availability and decreased in cost (Otondo 2008, p.77). Whilst this has assured the compositional engagement with space, there remains hesitation about the extent to which spatial synthesis ‘concerns’ music. It is only relatively recently, with scientific and technological advancements, that what Wishart calls ‘virtual acoustic space’ has begun to be qualified. The neutrality of the term ‘virtual acoustic space’ belies a complexity that is born of its technological embodiment and its musical appropriation. A class of contemporary spatialisation technologies here grouped under the label ‘reality-equivalent technologies’, have cast the composer’s concern with virtual space as the pursuit of realistic spatial illusions. Indeed, the performance of these technologies has been evaluated by their ability to attain ‘reality-equivalence’ (Malham 2001a, p.36). With this characterisation, Wishart’s ‘new world of possibilities’ begins to reveal itself as an entanglement for the composer who must now address the question of the relationship between music and the spatial dimension of realism. This thesis seeks to establish arguments and research to help the composer answer that question.

The approach to the research begins by engaging in the act of composing with spatial verisimilitude. The works produced, accompanied by critical reflections on them and the compositional process, serve as empirical support and guidance for the subsequent research. In producing this portfolio of works, bespoke software was
developed to ensure independence from any assumptions concerning the relationship between music and spatial verisimilitude embedded within existing tools.

The principal challenge in executing this research has been in adequately covering the breadth of associated fields. The closest context for the research topic concerns spatial music composition, and this subject is covered in some depth. However, the consideration of the element of verisimilitude, in spatial composition, leads the research into some broad areas of enquiry. For example, the fields of cognitive perception, psychology, and philosophy are all implied in the understanding of illusions of reality. The relationship between music and technology is also a broad field that itself enchains philosophical debate on the nature of technology. All of these areas contribute valuable arguments and insights to the research question. The approach taken has been to avoid engaging in a complete survey of the critical debates in all of these peripheral fields, and instead to pursue insights that specifically unblock the research path by offering alternative perspectives.

The research has produced several key insights. The source of certain tensions in the relationship between music and spatial verisimilitude have been identified and articulated. The importance of the technological mediation of this relationship has been detailed. How the composer might thus engage in negotiating the tensions between music and verisimilitude is clarified. Within my own practice as a composer, future spatial compositions will begin with a solid degree of informed understanding, rather than a naïve utopianism. It is hoped that this thesis will help other spatial music composers do the same.

7.1 Arguments presented
Perhaps the most consequential characterisation of the relationship between space and music is that the concerns of each are not clearly discernible from each other. This is manifested at the perceptual level: music has a broad capacity for creating perceptions of space, and space has a broad capacity for affecting perceptions of music. This interdependence, however, only seems to cause tension when the spatial dimension of a composition seeks to be ‘realistic’. The pursuit of realism imposes significant constraints on the musical dimension. In other words, the issues raised by the compositional engagement with spatial verisimilitude are more appropriately
described as concerned with realistic illusion, than as concerned with space. It is this reframing, in which space is demoted in favour of a focus on auditory illusion, which underscores many of the original insights presented in this thesis.

Historically, the reciprocal relation between music and space has not presented as a significant challenge to musical endeavour. As outlined in Chapter 3, composers have used space in many different ways. The use of existing physical space to spatially present music has been explored for centuries; from before the practice of cori spezzati in the 16th century to the 20th century experimentations of John Cage, Iannis Xenakis and many others. The use of references of space, detailed in Chapter 4, can be found in many common musical figures and constructs, across all music genres, and has an intimate relation to the musical use of timbre particularly in electroacoustic music. It is the goal of creating illusion through realistic representation that introduces a new dynamic between space and music. This new dynamic is characterised by the sense that one is served at the expense of the other. The fragility of auditory illusions places a slew of preconditions, on what sounds are presented and how they are presented. These preconditions encroach on musical thinking, which in turn seeks to free itself from the restriction of presenting sounds realistically.

This tension seems to deny the common assumption that spatial verisimilitude might serve compositional expression. The composer’s intuitive interest in spatial illusion thus seems misguided. However, research into the notion of presence, as it spans different academic disciplines, introduces concepts in which music and spatial verisimilitude can be understood on a common conceptual ground: both involve the cultivation of self-referential awareness. The composer’s intuitive interest in engaging with spatial illusion is thus potentially explained: spatial illusion is perceived as a means to enhance ‘experiential immersion’ in the music. The tension between music and spatial verisimilitude is also potentially explained: both seek to cultivate self-referential awareness, but in fundamentally different ways. In other words, the tension between them can be characterised as a competing between two different approaches to engaging the listener.
The composer is thus left to reconcile the differences between these two approaches. Herein lies an apt summary of the compositional engagement with spatial verisimilitude: it involves a process in which the composer must reconcile two very different approaches to cultivating the listener’s self-referential awareness.

Three such areas that require reconciliation have been presented. Firstly, in embracing the pursuit of verisimilitude the composer must answer to the question of the relationship between music and realism. Music is already known to have a tenuous relationship with realism (Dahlhaus 1985, p.115), but the pursuit of spatial verisimilitude forces its re-consideration. Verisimilitude insists on a specific understanding of realism: it exists in how things are represented, not in what is represented. In other words, the musical engagement with realism is perceptually centred and without immediate concern for the subject of the representation. The composer must address how these characteristics relate to musical concerns.

Secondly, the realisation of spatial illusion introduces new parameters that are foreign to compositional thinking. Presence and immersion, for example, have a largely unknown and unexplored relationship to musical form. Whilst some spatial attributes, such as direction, have been shown to fit into existing compositional methodologies such as serialisation\(^\text{22}\), attributes such as presence, which don’t have a clear discrete scale, require fresh consideration.

Thirdly, composers must negotiate the implications of the technological encapsulation of spatial illusions, which makes assumptions on the nature of illusions that have long been challenged. The technological encapsulation of illusion is informed by the scientific understanding of perception, in which realism is positioned as an objective empirical phenomenon. Reality-equivalent technologies thus necessarily project the limitations of a scientific worldview, in which auditory illusion is primarily understood as the presentation of auditory stimulus that simulates the behaviour of sound in space. In other words, nature is imitated. Both the imitation of nature and the synthesis of stimuli have been criticised as limited models of illusion. For example, Gombrich (1977) suggests that successful illusion

\(^{22}\) Stockhausen used serial techniques to score direction (Harley, 2000, p.155)
results from aligned expectations between artist and perceiver, rather than similarity with physical fact. Through the technological encapsulation of auditory illusions, the composer has a much-reduced scope for exercising their intuitive judgement of what constitutes convincing real space.

Reality-equivalent technologies exercise a very broad mediation of the compositional act that extends far beyond their encapsulation of illusion. This mediation, it is argued, offers unexpected opportunities for the reconciliation of the above-mentioned tension. In failing to deliver verisimilitude (see Chapter 5) reality-equivalent technologies contribute two compositionally significant things: firstly, they provide incidental sound material that is potentially musically interesting; secondly, they express an aspiration to spatial illusion. Within the resultant expression of a composition, the articulation of this aspiration is perhaps more significant than the delivery of spatial verisimilitude itself. Within Adorno’s terms, one might say that the musical material brought by reality-equivalent technologies is not realistic spatial illusion, but rather the aspiration to realistic spatial illusion. Hence, to create compositional expression, the composer must respond to this aspiration, not to spatial illusion.

7.2 Presence: a final insight
There is one common use of reality-equivalent technologies in which there exists no tension between music and realistic spatial illusion. In the spatial capture and playback of live acoustic musical performance the verisimilitude in the recording sits perfectly comfortably with the music heard. Indeed, the spatial recording of classical music performances is the most common usage of ambisonics found amongst the hundreds of works hosted on ambisonia.com (Deleflie 2005).

In such cases there is no direct relationship between the music heard and the spatial verisimilitude. In effect, the spatial verisimilitude does not serve the music: it is not a compositional concern. Instead, it serves the illusion that one is present at the musical performance. The actual compositional expression heard within the music has no importance or bearing in creating this sense of presence. Such a scenario proposes that the successful combination of spatial verisimilitude and music intimates the witnessing of musical performance.
When no musical performance is suggested and music is still heard, the realistic representation of sound in space enters into tension with the music. This tension, as discussed above, can be understood as the result of two different approaches to creating a sense of presence. Spatial verisimilitude seeks to create presence through perceptual realism, whilst music seeks to create presence through immersion in musical symbolism. The fundamental question engendered by this identification is: can the two approaches coexist to serve a singular musical expression, or do they necessarily act apart?

The consideration of the spatial capture and reproduction of music performance offers one answer: they sit apart but can both remain active. There are two possible levels of conscious awareness within this scenario: one concerns the illusion of being present at a musical performance; and the other concerns the conscious immersion in the musical expression of the work being performed. Here two levels exist, but they don’t necessarily conflict with each other, nor support each other. The fundamental question of whether the illusion of being present at the original performance might support immersion in the expression of the music remains unanswered.

7.3 Further research
Initial explorations into the body of literature concerned with the development of consciousness have already contributed significant insights. A more detailed investigation into the difference between Hunt’s (1995) notions of presentational and representational symbolism may offer a deeper understanding on the relation between music and verisimilitude. The identification of other conceptual foundations, in which music and spatial verisimilitude can be understood on similar terms, is also worthy of further research. One such conceptual ground exists in critical discussions concerned with the difference between illusion and illusionism. For example, visual arts theorist Mitchell (1995, pp.329-344) examines different perspectives on this difference, of which one is: “The opposition of illusionism to illusion is that of human to animal, self-conscious being to machine, master to slave” (1995, p.339). In this statement, an analogy exists between Mitchell’s illusion and the simulation of auditory stimuli, and Mitchell’s illusionism and immersion in musical symbolism. He later characterises illusionism as ‘aesthetic illusion’ or ‘self-
referential illusion’ (1995, p.339). The notion of ‘aesthetic illusion’ references Hunt’s notion of presentational symbolism, and the notion of ‘self-referential illusion’ references understandings of presence that focus on the relationship between the self and the environment. Of course, Mitchell’s discussion is concerned primarily with visual media, but their careful appropriation to auditory media may yield further insights.

From the perspective of compositional work, it is an appreciation for the significant technological mediation of the authorship of spatial verisimilitude that harbours exciting opportunities. This appreciation encourages the abandonment of the conscious pursuit of verisimilitude, in favour of letting existing technologies express their own dialectic. There is musical material within the way technologies fall short of spatial illusion. It is the compositional engagement with, and response to, this ‘falling short’ that offers a new way to explore the musical potential of spatial verisimilitude. In such an approach, spatial verisimilitude is not used to support the compositional subject, but rather becomes a part of the compositional subject. Expression is born from the compositional response to that subject.

Another avenue for research concerns the combination of Gibson’s ideas on perception with the above-mentioned concept of levels of illusion. The application of Gibson’s ideas to the concerns of spatial music has already received critical attention (Worrall 1998; Lennox, PP, Myatt & Vaughan 2001; Lennox, P & Myatt 2007). The increased level of spatial illusion possible by incorporating head-tracking into binaural auralisation is widely known. What is suggested here is that head-tracking be used in conjunction with no head-tracking. The carefully articulated and temporal juxtaposition of spatialised sounds that respond to ‘action in the environment’ (head turns) and spatialised sounds that do not respond to head turns would allow for the musical exploration of the meaning of different levels of illusion.
REFERENCES


Babbitt, M 1962, ' Twelve-tone rhythmic structure and the electronic medium', *Perspectives of New Music*, vol. 1, no. 1, pp. 49-79.


Bayle, F 2007, 'Space, and more', *Organised Sound*, vol. 12, no. 3, pp. 241-249.


Hamman, M 2002, 'From technical to technological: The imperative of technology in experimental music composition', Perspectives of New Music, pp. 92-120.


Harley, MA 1994a, 'Space and spatialisation in contemporary music: History and analysis, ideas and implementations', PhD thesis, School of Music, McGill University, Montreal.


Potard, G & Burnett, I 2004, 'Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays', in *Proceedings of the International


Sanchez-Vives, MV & Slater, M 2005, 'From presence to consciousness through virtual reality', *Nature Reviews Neuroscience*, vol. 6, no. 4, pp. 332-339.


Truax, B 2002, 'Genres and techniques of soundscape composition as developed at Simon Fraser University', *Organised Sound*, vol. 7, no. 1, pp. 5-14.


APPENDICES

Appendix 1: Contents of DVDs

DVD 1 contains:

- *Index.html*. This page provides easy access to the complete contents of DVD1, including the capacity to play sound files from within an Internet Browser. It also provides an installation manual for installing and configuring ImageSynth.

- *ImageSynth software*. The complete source code of the spatial compositional environment developed.

- *Sound files*. Sound files are provided for all the works listed in Appendix 2 except for the 3rd order ambisonic encoding of *First Flight Over an Island Interior*.

DVD 2 contains:

- The 3rd order ambisonic encoding of *First Flight Over an Island Interior*. This file is provided on a dedicated DVD because of its large size (3.83 GB).

Appendix 2: List of Works in Portfolio

*Torch video sonification*, video including stereo sound track (0:10)

*Fluoresce*, presented in 5.1 surround, stereo, UA and b-format (0:37)

*Original sound recording used in Fluoresce*, mono (0:10)
Drivetrain, presented in 5.1 surround, stereo, UA and b-format (1:02)

Lost Corps, presented in 5.1 surround, stereo, UA and b-format (7:26)

Hammer, presented in 5.1 surround, stereo, UA and b-format (3:44)

Gorged, presented in 5.1 surround, stereo, UA and b-format (6:33)

First Flight Over an Island Interior, composed in collaboration with Kraig Grady, presented in 5.1 surround, stereo, UA and b-format (45:17)

Test rendering of ‘First Flight Over an Island Interior’ using Xenakis' Gendy algorythm, stereo (4:04)

Test rendering of ‘First Flight Over an Island Interior’ using Sawtooth generator, stereo (3:33)

Test rendering of ‘First Flight Over an Island Interior’ using Sine wave generator, stereo (3:47)