ANGLIA RUSKIN UNIVERSITY

INTERACTIVE SONIFICATION OF A PHYSICS ENGINE

RHYS JOHN PERKINS

A Thesis in partial fulfilment of the requirements of Anglia Ruskin University for the degree of Doctor of Philosophy

Submitted: September 2013

ACKNOWLEDGEMENTS

There are a number people that I would like to thank for helping me during the course of my studies.

The Faculty of Arts, Law and Social Sciences for providing me with a scholarship, without which this project would have never been possible.

The Department of Music and Performing Arts for supplementary financial assistance.

My supervisor Dr. Richard Hoadley for his excellent guidance and advice.

Supporting supervisors Dr. Tom Hall and Dr. Paul Rhys for being available when required.

Prof. Robert Home and Dr. Douglas Nunn for helping me to satisfy bureaucratic administrative requirements.

Patrick Hemmerle for having the patience to teach me the piano and reaffirming my love of music.

My family and friends. You have always been the best motivators and have provided me with much needed time away from work while lending a sympathetic ear.

And last, but by no means least, my parents, Mike and Linda, for their financial support, kindness, and showing continued interest when reading and commenting on my work. It is to them that I dedicate this thesis.

ANGLIA RUSKIN UNIVERSITY

ABSTRACT

FACULTY OF ARTS, LAW AND SOCIAL SCIENCES

DOCTOR OF PHILOSOPHY

INTERACTIVE SONIFICATION OF A PHYSICS ENGINE

RHYS JOHN PERKINS

August 2013

Physics engines have become increasingly prevalent in everyday technology. In the context of this thesis they are regarded as a readily available data set that has the potential to intuitively present the process of sonification to a wide audience. Unfortunately, this process is not the focus of attention when formative decisions are made concerning the continued development of these engines. This may reveal a missed opportunity when considering that the field of interactive sonification upholds the importance of physical causalities for the analysis of data through sound. The following investigation deliberates the contextual framework of this field to argue that the physics engine, as part of typical game engine architecture, is an appropriate foundation on which to design and implement a dynamic toolset for interactive sonification. The basis for this design is supported by a number of significant theories which suggest that the underlying data of a rigid body dynamics physics system can sustain an inherent audiovisual metaphor for interaction, interpretation and analysis. Furthermore, it is determined that this metaphor can be enhanced by the extraordinary potential of the computer in order to construct unique abstractions which build upon the many pertinent ideas and practices within the surrounding literature. These abstractions result in a mental model for the transformation of data to sound that has a number of advantages in contrast to a physical modelling approach while maintaining its same creative potential for instrument building, composition and live performance. Ambitions for both sonification and its creative potential are realised by several components which present the user with a range of options for interacting with this model. The implementation of these components effectuates a design that can be demonstrated to offer a unique interpretation of existing strategies as well as overcoming certain limitations of comparable work.

Keywords: auditory display, data representation, human computer interaction, physics engine, rigid body dynamics.

TABLE OF CONTENTS

COPYRIGHT

CHAPTER 1 - INTRODUCTION

1.1 Overview

The introduction begins with a brief background before noting several pertinent developments which function as the motivation behind a more detailed investigation. This is followed by outlining the methodology and aims pertaining to the creation of a novel tool for the interaction sonification of a physics engine, known as Mhyusics Auditory Display, which is defined within the scope of the thesis. A potential audience for such a tool is then discussed along with the contributions it can provide. The second chapter familiarises the reader with the fields of sonification and auditory display serving as a rationale behind their employment and determining much of the corresponding vocabulary. It also determines the main body of references for the reader should a more in depth account of the two fields be required. The third chapter considers many of the key areas comprising an AD along with the decisions made in the implementation process. A detailed narrative, backed by the literature, supports the many arguments underpinning each aspect of the design defining the current version of Mhyusics. The fourth chapter details the requirements for Mhyusics to be implemented along with the rationale for the software chosen during its development process. The current implementation of this AD is then compared with other relevant work. The concluding stage of the thesis summarises the salient points within the preceding chapters, identifying several common themes in order to support the relevance of such a tool. Propositions are then made for possible future research directions upheld by notable developments in the appropriate domains.

1.2 Background

Interactive sonification bases much of its contextual framework on how we interact with and perceive the physical world. Two of its pioneers (Hermann & Ritter, 2004) demonstrated that sound is a crucial source of information about our physical environment and looked at how sound could encode meaning from the perspective of physics. They refer to interaction sounds, which are produced when physical objects touch or have an impact on one another. Furthermore, they suggest that by employing analytical listening in our perception of this event we can learn about properties behind the sound-producing process. For instance, we exercise this particular type of listening when shaking a box to try and guess its contents from the sound generated in response to our actions. Much of their work stems from Gaver (1993) who distinguishes the act of everyday listening as one where we attend to events rather than sounds. He proposes that we can recognise events based on the relationship between sound structures and their underlying physical cause.

From this standpoint he determines two categories that we are capable of estimating via the perception of ecological information, that is, obtained by direct sensitivity to invariant structures in our world. The first type, referred to as structural invariants, can be described as the attributes of the objects involved such as size, shape, mass, elasticity and material. The second, described as transformational invariants, comprise velocities, forces and the position of contact points. By examining acoustic patterns that correspond to a number of physical events including scraping, bouncing and spilling Gaver derives a variety of algorithms to create sounds that would capture relevant structural features for each event. The synthesised versions accompanying these interactions were not necessarily perceptually identical to the original sound but instead conveyed recognisable information that convinced the listener. Since then, with the continued increase in computer processing power we find ourselves in a position where we can produce evermore convincing sounds for these physical events. This includes various states such as solids (van den Doel et al, 2001; O'Brien et al, 2002; Lloyd et al, 2011), liquids (Moss et al, 2010) and gases (Chadwick and James, 2011). In all these cases the data incites a physically informed sonic model where events translate into a sound structure that resembles its real world equivalent. However, we can also consider these events as the driving force behind a more musical purpose, presenting the opportunity for a wider range of sonic interpretation while still being governed by the fundamental structure of the data.

According to Chion (1994) our brain is capable of providing a link between raw physical events and a more aesthetic meaning. He illustrates the concept of synchresis as demonstrating how we possess an innate ability in our 'forging of an immediate and necessary relationship between something one sees and something one hears.' (p. 224). This psychological phenomenon allows viewers to accept cinematic events as logically connected where sound and image can be completely unrelated, presenting us with many ways to utilise sounds in conjunction with visual events. In the broadest sense, a musical soundtrack can be used to punctuate the events of physical action; a process generally referred to as "Mickey-Mousing". From a more discrete point of view, individual instruments can be used to imitate physical interaction, focusing attention on their own set of unique characteristics. For instance, in cartoons we can often hear the crash of cymbals accompanying a collision, or the fall in pitch of a slide whistle accentuating a falling coyote in an episode of "The Road Runner Show". Video games will produce sound to signify character events, such as jumping or shooting, where audio is considered as an essential component that often improves the player's involvement (Coleman et al, 2005). The idea of involvement is also the drive behind the work of Foley artists. In this manner, their creative use of properties (theatrical props) attempts to convince the audience of the sonic counterpart to an onscreen action while producing a more natural experience in the process.

It should be noted that all of these examples share similarities with Gaver's synthesis routines in the fact that they capture the relevant features of their sources. Nevertheless, in this case their relationship can be seen as more abstract where there is considerably less emphasis on scientific simulation. One reason for this approach could be viewed as a solution for a situation where it would be impractical to try and recreate the same sounds as found in the original event. For instance, in the context of a cartoon, the actual sound of an anvil falling on a coyote's head would most likely be met with feelings of horror as opposed to the humour it is intended to evoke. Likewise, a fantasy situation may involve objects that do not exist in reality and therefore have no tangible sounds associated with them. In both of these cases the audio must try and persuade the listener that an event has taken place thus providing a greater experience over the visual modality alone. As such, the sound designer must find a convincing link between event and sound in order to reach a solution, a process that will often find inspiration in the form of a metaphor.

Lakoff and Johnson (1980) revealed that metaphors often have their basis in the physical world around us and derive from our cultural experiences. They portray them as fundamental in the shaping our perception of reality and argue that they form the justification for science where 'the intuitive appeal of a scientific theory has to do with how well its metaphors fit one's experience.' (p. 19). Indeed, the language of physics employs many metaphors to describe its phenomena (Stahl, 1987) and the sonification of such phenomena has proved to be successful when based on metaphorical principles (Vogt, 2010; Vogt & Höldrich, 2010). Perhaps this should come as no surprise given that sonification follows a close analogy of metaphor when viewed as the understanding of a target domain (data) by placing it in the source (sound). The simulation of our physical environment as a metaphor can also be found in the main areas of interest pertaining to interactive sonification. With appropriate use it can enhance human computer interaction (HCI) (Jacob et al, 2008), help elucidate meaning in data transformation methods (Fels et al, 2002) and bring greater immersion to ADs (Grimshaw, 2007). Central to all of these areas is the idea that metaphors allow us to make sense of the gap between physical and virtual environments where the laws of physics provide a comprehensive set of principles for interaction and understanding. In this regard, the modern physics engine could be seen as a general means of accessing and presenting universal truths within the virtual world. Simply put, it serves as a source of inherent physical analogies that can close the gap between virtual and physical. Furthermore, given that the information provided lies within the framework of computer software, it affords the opportunity to render possibilities beyond that of the physical world. Chowning (2006) stated that 'the integration of programming concepts into the musical imagination can extend the boundaries of the imagination itself.' (p. xii). By establishing these concepts in the context of this investigation it will become apparent that we can envisage new, and perhaps more importantly, universal methods of musical control, representation and interpretation.

1.3 Motivation

1.3.1 Technological advances

Ongoing advances in information technology have made it possible to render a vast quantity of information at high speed with low computational cost. Visual replication of the real world in the virtual has been one of the driving forces behind this evolution where we seek to create ever more realistic and immersive computer imagery benefiting industries such as film, television and video games. Nevertheless, the rendering of real world entities originally called for a shift in focus from two-dimensional sprite-based graphics to polygonal representations. This saw the emergence of dedicated graphics processing units (GPUs) designed to accelerate three-dimensional (3D) visualisation by providing common functions, such as transformation matrices, required to display polygonal vertices. Since their invention GPUs have evolved to allow for full 3D real-time interactive data display satisfying the demand for faster, high-fidelity graphics (Nickolls & Dally, 2010). In addition, we have now reached the point where they are ubiquitous in modern technology and whereas graphics processors used to be an expensive hardware expansion, they are becoming more closely integrated within the central processing unit (CPU) (Vaughn-Nichols, 2009). This move away from discrete hardware has allowed graphics acceleration to be available in a more compact form and ensured its prevalence in everyday technology such as mobile phones and tablet devices.

The increase in the level of realism in the virtual environment not only depends on the enhancement of modelling and rendering effects, but also on the continual improvement of other elements such as animation, artificial intelligence and physics simulation. Most notable in this regard is the physics engine which often operates tasks for other components. For instance, engines can perform collision detection and resolution routines, thus decreasing the prospect of a character perceptibly intersecting its surroundings. They are also are instrumental in the production of increasingly realistic movement for character animation (Geijtenbeek & Pronost, 2012). Physics engines continue to remain relevant to modern software and are the focus of a great deal of attention. As such they are constantly improving, partly in response to gamers who constantly demand a more immersive experience, but also due to need for serious applications where greater stability and fidelity is required (Choi et al, 2010; de Paolis, 2012). Further inspiration behind their ongoing development comes from the many open source examples such as Bullet (Coumans, 2013) and Open Dynamics Engine (Smith, 2006) which bring with them a culture of dedicated contributors who provide insight into a black box for game developers where physics simulation has become a commodity (Millington, 2007). Together with their specific capabilities, they are able to operate on all major operating systems and are adaptable to various configurations involving the CPU and the GPU (Joselli et al, 2008; Harada, 2011) thereby increasing their bearing on technology.

The spread of technology into everyday devices has precipitated the integration of physics engines into popular culture. Whilst already pervasive in the world of video games and interactive media they are now proving to underpin an increasingly popular interaction paradigm in the field of HCI (Jacob et al, 2008). The education sector has also capitalised on simulation training where game engines can be adapted to the many modalities of learning while supplying a secure, augmented environment at a lower cost (Arango et al, 2008; Clark et al, 2010). However, it is important to recognise that the development process for physics engines is incomplete and that current implementations are generally skewed towards video games. In this regard there is still the need to adapt them for more musical aspirations. Software associated with interactivity and music (Puckette, 2002; Liebe, 2013) is now at a turning point where the integration of physics simulations will bring several benefits. By building on the work presented in these fields we can create a new framework for how to understand, teach, generate and manipulate sound within the confines of a natural and immersive virtual space.

1.3.2 Drive from the related fields

On the topic of the advancement of sonification design Worrall (2009a) suggests:

> 'generally available data sets and interesting mapping templates in software environments that combine sonic complexity with experimentation flexibility will encourage the much needed communitywide sharing of examples toward a catalogue of best practice.' (p.328)

As previously discussed we can now consider the freely distributed physics engine as an example of a generally available data set which has become commonplace in modern computing. Their cross-platform approach has seen widespread use that not only utilises long established platforms such as PC and Macintosh but also follows the current trend of technology with software such as Unity (Unity Technologies, 2013) providing mobile based solutions for iOS and Android. Due to the increase in computer processing power in recent years we have reached the point where it is possible to reliably render sound in real-time alongside interactive visuals. This presents us with an opportunity to revisit previous sonification methods pertaining to physics (Sturm, 2001a; Henry, 2004; d'Inverno & Olofsson, 2006; Hamilton, 2008) in order to build upon their ideas with current technology.

Two of the founders of the field of interactive sonification (Hunt & Hermann, 2004) have argued that we should contemplate our corporeal interaction methods with everyday objects in our approach to connecting with virtual data. Much of what they describe is embodied in the first-person shooter (FPS) genre of video games that use the metaphor of an avatar, essentially a controllable camera, to facilitate real-time interaction within the game world. Here, self-simulation grants the ability to generate impromptu, true to life gestures in order to interact with, manipulate and generate information within a virtual space. This has prompted others to indicate that game engines can therefore grant a comprehensive framework for ADs where data can be sonified in real-time (Le Groux et al, 2007; Hamilton, 2008; Dolphin, 2009a; Nguyen, 2012). Indeed, with the integration of the physics engine increasingly seen as mandatory, many modern game engines now facilitate advanced rendering of simulated objects with real-time interaction and audio response. In this regard they can be seen as having all the necessary components required to construct an ideal sandbox for interactive sonification.

1.3.3 Going beyond physical limitations

Computers have changed the manner in which we interact with the world and we are continuing to see the percolation of computing devices into ordinary ubiquitous objects and everyday activities. As we increasingly rely on technology we look to imbue many of the qualities of the physical world into the virtual. However, one of the advantages of the virtual is the greater freedom of control where we can impose our own personal ideas. Although we should be aware that our physiology dictates the limitations in what we can perceive and understand, we are dealing with an environment that harbours considerable artistic merits. In sonic terms it now becomes possible to create timbres that exist outside the realms of the physical constraints of traditional acoustic instruments. From the point of view of a physics simulation, Sturm (2001b) identifies that we can bend scientific laws to suit a composer's taste. For instance one could define an object that is unaffected by gravity and instead generate its own vector field, influencing other neighbouring objects with a defined area. In fact, we have the power to accurately configure any number of modular abstractions of simulated objects where subsequent instances can be created and deleted at will. Additionally, by combining these elementary components in various configurations we can achieve something greater and more complex within a comparatively short time frame. This is of particular importance in the framework of composition where it can be exercised in prototyping musical ideas.

Cadoz (2002) used the etymological basis for the word compose ("put together") to provide an analogy for the process of assembling modular entities, or building blocks, for the sake of music creation. He portrays a series of simulated masses that when connected form a rudimentary orchestra comprising vibrating structures. In the context of rigid body dynamics this would suggest that it is possible to construct a vast number of contraptions in the style of Heath Robinson or Rube Goldberg. To some extent this has been manifest in popular culture with software like Garry's Mod (Face Punch Studios LTD, 2012), described as a physics sandbox, providing extant examples within its user-generated content. However, it is through the sonification of these systems that we allow these potential users to introduce their creations to a new distinctive quality. Here, both man and machine can determine dynamic gestures, generative rules and object configurations, where a subsequent simulation will unfold into a unique tapestry of sound.

1.4 Approach and aims

This thesis is an investigation into a real-time interactive sonification of a 3D physics engine where the focus will be primarily concerned with a particular subset of classical mechanics, rigid body dynamics. This interaction type was chosen because it can be regarded as an industry standard where there are many well established, cross-platform examples (Boeing & Bräunl, 2007). Common solutions are capable of responsive, real-time simulations with headroom for more computationally demanding scenarios alongside separate tasks including sound synthesis and networking. By comparison, other simulation types, such as fluid dynamics or soft-body dynamics, often come with higher computational demands and are not as widely supported.

What follows outlines the theory behind the design of an AD which is built within the framework of a games engine comprising real-time user interaction with a complex data set. Its main aim is to demonstrate that sonification can appeal to a wider audience by leveraging a well-established, accessible source of information with a broad cultural understanding. The inquiry takes an inductive approach to the literature where many aspects of the final design are built upon extant work in order to support a new tool that can play an active role in the sonic interpretation of a dynamic, multivariate data set. This involves combining distinct approaches from the fields that define interactive sonification along with a direct comparison to what can be considered the closest comparable examples in this area.

The topics presented here articulate the argument that the creative process of interpreting the gap between a physical event and its corresponding sound provides an opportunity for a fresh insight into the underlying data while forging the foundation for new ventures in music composition. In the most elementary sense this body of work supports a tool that encourages the user to interact with a customisable system comprising virtual objects that can emulate, or go beyond, the limitations imposed by their physical counterparts. Visually, the didactic nature of the polygonal models not only afford a common, logical approach to comprehending rigid body behaviour but also yield to a number of well-established methods for gestural interaction. However, it is through the sonification of object behaviour, or more specifically their mapping configurations, that the software can be shown to support a wide variety of uses. In this regard, several notable ideas will become manifest throughout the thesis, from which compositional examples can be derived. These examples should be seen as a starting point where one can explore data to sound mappings and gain ideas for musical scenarios. Taking the form of both a video demonstration and saved configurations files, they are supported by the relevant literature and serve to reflect on their associated area of discussion while highlighting prominent features included in the design [Appendix 1]. As only a limited number of instances can be demonstrated within the scope of the thesis it is left up to the reader to form inspiration from what is accommodated by the design of Mhyusics Auditory Display. Further propositions regarding potential implementations are also available in the concluding chapter.

1.5 Audience and contribution

1.5.1 Music

There are several examples in the literature which support the notion that the sonification of physics simulations can be of potential significance. Regarding musical aspirations, the system described in this thesis can be used to foster the talents of composers from a multitude of backgrounds and of varying technical abilities. At its most accessible level it can be considered as a sonic art installation or sound toy, in same sense as Dolphin (2009b) describes one of his compositional applications, MagNular. A setup of this type would require loading two preconfigured files, the first being an AD state that presents the user with a set of preconfigured objects. For instance, a collection of spheres contained inside a box or an arrangement that simulates Newton's Cradle. This file also bypasses the need to generate objects and determine their mapping arrangement. The second file serves as the means of audio generation by taking the form of a programming environment patch or synthesiser preset. Communication between the two environments, denoted by the corresponding files, relies on the integrated MIDI and OSC messaging protocols. The software also supports a more advanced level of user intervention such as a custom mapping configuration, creating and deleting objects, or the changing of simulation parameters. This would encourage the user to engage with Mhyusics Auditory Display as a compositional environment in the same vein as Versum (Barri, 2009) and µ (Bukvic & Ji-Sun, 2009).

At all levels of operation the audiovisual environment allows the composer to quickly prototype a musical idea using a dynamic set of components. Indeed, graphical abstractions of physics simulation systems (Castagne et al, 2009) have previously demonstrated the power of utilising simulated entities for a wide variety of musical purposes. Some notable examples include the ability to control audio engines and synthesis parameters (Henry, 2004), form compositions (Cadoz, 2002) and illustrate complex algorithms (d'Inverno & Olofsson, 2006). Tache and Cadoz (2009) acknowledged this concept when affirming that 'modularity has obvious advantages over the non-modular approach: it brings generality to the environments in which it is used, and gives more freedom to the user' (p. 411). Gresham-Lancaster (2012) corroborated the idea of generality when assessing sonification as a musical craft. Additionally, he added the intention of appealing to a broader user base maintaining that 'sonification has the promise to revolutionize a major area of the human computer interaction environment, but only when it is framed in a generalized cross-cultural way that is fully inclusive of its entire potential.' (p. 207). With the modification of a 3D physics engine for the purposes of sonification we can take the generality of these environments further by placing them in the context of a more deep-rooted, realistic representation. This presents the viewer with a lower entry level for their comprehension while eliciting a common ground for the discussion of musical activities via audiovisual analysis.

1.5.2 Science

Visual simulations have already proven invaluable as an aid in reducing complexity of physics data by presenting it in a context where the viewer can quantitatively predict outcomes (Clark et al, 2009). By adding an auditory channel to these tools the algorithms underpinning the simulation can be described through sound. Both Sturm (2001b) and Vogt (2010) laid out some of the benefits that the sonification of a particle physics simulation would bring to science, including new ways of understanding physical phenomena. Sturm (2001b) discussed pedagogical advantages that Western music composition students might gain from interacting with an audiovisual simulation. In particular he suggests there is an advantage to combining the audio and visual modalities in order to present musical ideas, likening it to listening to a piece whilst reading a score, rather than partaking in one or the other activity separately. Vogt (2010) suggested that the data found in physics was well suited for sonification purposes stating that 'while standard scientific graphs plot time on one axis, this is not necessary in a sonification, where physical time persists as sonification time.' (p. 30). In other words, the dynamic nature of physics granted a very direct mapping of physical time to the temporal evolution of sound in sonification. Sturm (2001b) was also aware of this idea when he proposed that 'by employing an audification of physical principles, most of which are dynamic as well, the physics student may follow along in the notation, and thus gain more exposure to the principle at hand and its interpretations.' (para. 34). This sentiment has now been substantiated by the portrayal of physical models through graphical, gestural representations alongside their underlying equations (Castagne & Cadoz, 2009; Schroeder et al, 2011). In this manner, music and sound become the means of demonstrating mathematical constructions, granting the viewer with a new vantage point for comprehension and creating the possibility of alternative hypotheses.

1.5.3 Popular culture

There are several examples in the field which uphold sonification as a suitable candidate for popular culture. The first-person shooter genre is arguably one of the most popular within the video game community with titles such as the Call Of Duty series regularly ranking as the top selling game (Caoili, 2012). Work by Hamilton (2008) and Berthaut et al. (2011) has successfully modified some notable examples in order to accommodate the sonification process while others have upheld the application of game engines, often skewed towards this particular genre, as sonification systems (Grimshaw, 2007; Nguyen, 2011). Commercial examples have combined interaction with physics for the purposes of composition and play where sonifications of simplified physics simulations have proved to be viable as both games (Nintendo, 2012) and applications (Simian Squared Ltd, 2012). In addition, recent developments have seen the physics engine integrated into more conventional applications designed for complex musical purposes (Cycling '74, 2012).

Wang et al. (2011) demonstrate that we can reach a global audience by placing musical applications on ubiquitous devices such as smart phones. This gave rise to the use of mass human intelligence, or crowdsourcing, to encourage the making of music as a popular activity and allowing the audience to engage both actively and passively in the process. By applying these ideas to the AD designed as part of this thesis we can encourage others to create and share ideas for scenarios associated with a set of dynamic simulated entities. In this regard, rigid body interaction introduces new users to a globally inherent metaphor for sound generation where the flexibility of the object behaviour, along with their sonification, imparts a broad range of interpretation and promotes musical creativity on a grand scale.

1.6 Summary

This chapter provided a brief introduction to the interactive sonification of a physics engine by placing it in context before determining both its significance and potential audience. The next chapter discusses two subjects that are of central importance to the scope of the thesis, sonification and auditory display. Several major approaches to the sonification process are also examined along with some thoughts on why it can be considered as relevant to both art and science.

CHAPTER 2 - SONIFICATION AND AUDITORY DISPLAY

2.1 Overview

The following chapter is intended to provide the reader with a brief introduction to the field of sonification and auditory display. While a full summary of sonification would be beyond the scope of this thesis, this account serves as a guide to the key literature, terminology and topics that predicate the design of Mhyusics Auditory Display. The first section gives a brief historical overview which considers the field as both science and art. The second section reveals some common arguments in favour of employing the ear as a means of understanding complex data structures followed by a segment that discusses several notable methods of data transformation. The final part of this chapter aims to outline the subfield of interactive sonification whose components provide a framework for the topics discussed in later chapters.

2.2 Historical summary

In comparison to visualisation, sonification is considered a young discipline (Hermann, 2008) and auditory display has been referred to as in its stone age (Walker & Nees, 2005). The ideas for both were consolidated at the first International Conference on Auditory Display (ICAD), held in 1992, resulting in a publication that is often regarded colloquially as the Old Testament (Kramer, 1994b). This proved significant as it brought together like-minded researchers from many disciplinary backgrounds and provided a common language. Kramer (1994a) summarises a brief history of the sonification field up to this point including the conception of data sonification, as established by Bly (1982). The collaborative efforts of The Sonification Report (Kramer et al, 1999) described a major summary of events during the following interim period while providing a springboard for discussing underlying theories and drawing attention to many relevant areas of research. In recent years The Sonification Handbook (Hermann et al, 2011b) aimed to give a more thorough and detailed overview. Specific technical fields falling under the rubric of sonification and auditory display were provided with their own chapters so that each topic could be rigorously explored. As such, this tome is now widely regarded as the current authoritative reference, or guidebook, and is informally considered to be the New Testament of the literature.

A common theme throughout all these publications, and amongst the ICAD community, is the drive to legitimise the field as an academic science in its own right (Supper, 2012). Regrettably, science frequently has its focus on vision, where it plays a primary role in research, teaching and presentation of data. However in a considered effort, Hermann (2008) calls to use sonification with the same level of generality as its corresponding term visualisation is used in visual display. He proposes new definitions in order to account for emerging techniques where older variants could be considered too unspecific. His stance seeks to affirm the significant role sonification can play in the scientific domain, but in doing so he clearly distinguishes it from music. Conversely, Vickers and Hogg (2006) described many examples that blur the line between the two subject areas by shifting their focus towards more aesthetical considerations. They concluded that strict segregation was not feasible stating that 'to distinguish between musical and non-musical sonifications is not necessarily helpful (or even meaningful).' (p. 215). Instead, their work associated examples with their closest analogue in the musical domain based on their arbitrariness of mapping or indexicality. From this analogy they describe a representational continuum that ranges from low, a symbolic or interpretative connection, to high, deriving directly from the data source. The lower end sought to depict full artistic freedom, often linked with more musical purposes, while the opposite suggested the tighter constraints of scientific rigour and clearer representation.

There are many composers who might regard their sonifications as substance for musical listening (Childs, 2007). Indeed, Barrass and Vickers (2011) have shown that calls to include composers in the sonification design process have been around since its inception and are still relevant today. The basis for their argument considers how extra-musical sources have long since been the subject of compositional inspiration. External data sources continue to be contextualised both technically and culturally, making them more accessible and harder to ignore. This implies that data sonification is of significant interest from both a cultural and experimental compositional perspective and goes further in strengthening the validity of the process as a form of musical art.

2.3 Benefits of using an auditory display

Regardless of the ongoing debates, there is clear evidence which demonstrates the benefits of providing an auditory channel as the primary or auxiliary means of conveying information. Much of this argument stems from the limitations of relying purely on the visual modality. Kramer et al. (1999) cite Moore (1997) when affirming that auditory perception is particularly sensitive to temporal characteristics. They put forward the notion that fast-changing or transient data, that could be blurred or completely missed by visual displays, may be easily detectible in a well-designed AD capable of a superior temporal resolution. This is of particular pertinence in a physics simulation where determining the exact angular velocity of a rigid body based on visual clues alone can prove impossible due to temporal aliasing such as the stroboscopic effect. When dealing with more complex disciplines of physics, much of the activity is often beyond the perceptual limens of our visual system. In this regard, audio has proved to be a useful alternative for exploring phenomena such as quantum mechanical behaviour (Pereverzev et al, 1997) and computational physics models (Vogt, 2010).

When Bregman (1990) described auditory scene analysis he proposed that our auditory system has the innate ability to segregate audio by making use of our brain's cognitive functions incorporating both attention and learning. He demonstrates that we are not only capable of processing several streams of information in parallel but our ability to perform selective listening also allows us to segregate high-dimensional systems by focusing our attention on single streams of audio. Furthermore, acoustic cues allow for primitive grouping of information where the ear exhibits a high sensitivity for rhythm, spatial location and proximity in both frequency and time. Much of this theory has been integrated into the sonification process (Barrass, 2009) and knowledge of analogous auditory gestalts has proved to be vital when uncovering hidden structures and meaning in collections of data (Kramer, 1994c).

The appropriateness of the ear as a tool for data exploration is evident in our ability to communicate physically with sound. By listening to someone talk we can extract multiple layers of information from their voice such as emotion, gender, age and health. Chion (1994) defines semantic listening as referring to a code of language, the interpretation of a message or gleaning information from a sound. He suggested that the ear was adept for language communication hence we can understand spoken language faster than the eye could read the transcribed equivalent. Ballas (1994) supports arguments in this area when using a framework of linguistic analogies to show how many of the functions of language are found in non-speech sound and are applicable to the process of data sonification in order to effectively communicate information.

Kramer (1994a) summarises a great number of advantages when using ADs alone or in conjunction with visual displays. He notes that by utilising the auditory system's unique strengths we can keep our eyes free to monitor elsewhere thus increasing the bandwidth for a human-computer interface. Corroborating evidence comes from Hermann and Ritter (2004) who demonstrate that audio is complementary to other sensory modalities when performing everyday tasks providing valuable information about the interaction process. Weinberg and Thatcher (2006) proposed that a sonification could be more effective in a spatial application stating that 'the human auditory system is able to perceive synchronous spatial stimuli from every point within a space, while visual perception is limited to the physical range of sight.' (p. 9). Correspondingly, a number of recent examples in this area have supported this argument (White et al, 2008; Neidhardt & Rüppel, 2012; Niinimäki & Tahiroğlu, 2012). In a more general context, evidence for the successful employment of audio in the absence of a visual modality can also be found from the proceedings of the haptic and audio interaction design (HAID) workshops.

2.4 Data mapping techniques

One of the core components in the sonification discipline is the data transformation technique, or as it is more commonly known, data mapping. By definition, in every sonification some form of mapping takes place. This section looks at the most common techniques employed in the field along with a few of the advantages and disadvantages associated with each. It should be noted that many hybrid sonifications will often use mixed methodology, according to the data under investigation, making strict classification unfeasible.

Auditory icons and earcons can be considered the least complex approach for mapping data onto sound. Both of these techniques use audio to signal an event in the same manner that a doorbell will alert a tenant to the presence of someone outside their home. Auditory icons were introduced by Gaver (1994) and serve as the equivalent of the visual icon in the desktop metaphor. They mimic everyday non-speech sounds based on our experience of the real world, accelerating the learning process by drawing upon our common sense for their interpretation. In contrast, earcons use structured musical messages in order to convey information. The key difference with the earcon being that there is no assumption of an existing relationship between the sound and the information that it represents. Examples of both auditory icons and earcons can be readily found in modern technology. For instance, mobile phones allow the user to customise the mapping of ringtones in order to denote a particular caller. The modern computer serves as another common example where many operating systems will use context sensitive sounds to accompany events such as emptying the recycle bin, or displaying an error.

Orders of sonification, first described by Scaletti (1994), refer to levels of abstraction in the data transformation process. At the lowest level there is a direct mapping to sound, a common process that can seen in many devices but not one we would necessarily associate with sonification. For instance we can consider that the engineer of an alarm clock must make a conscious effort to determine a particular brand of component to play the role of the alarm. In this process he establishes the final pitch of the alarm sound which will signal the onset of an event. The successive level involves mapping data to an audible synthesis parameter, for example, pitch and amplitude. A further level of abstraction would concern a parameter that controls another such as modulation depth. Undoubtedly, these levels of abstraction are important for one to take into consideration when seeking to describe data through the medium of sound. Vickers (2005) makes the distinction between first and second-order sonifications by exploring the relationship of functions and graphs. He argues that information is lost when using a visual abstraction of an underlying function. Therefore, the sonification of a graph should be considered as second-order while not utilising the full potential of the data transformation process. Conversely, Nguyen (2011) discusses an approach where he favours a loss in resolution of the data in the mapping process. He promotes greater abstraction in return for an increase in musicality where the use of musical structure has the potential to communicate compound relationships with a greater clarity over high resolution data.

Audification, also referred to as zeroth-order sonification, is a 'direct translation of a data waveform to the audible domain for the purposes of monitoring and comprehension.' (Kramer, 1994c: p. 186). This technique uses the data values directly as a series of sound pressure waves in order to provide a sample specific relationship between data and sound. Intermediary elements such as a mapping function are usually restricted to scaling and frequency shifting. Signal processing techniques such as compression and filtering may be used but no sound generating elements should be introduced. As a result of this transformation the size of the data set becomes the duration and pitch, while the timbre reflects aspects of its temporal evolution. Musical listening is then exercised in order to interpret the sounds and thus the underlying data. While temporal variations can be well emphasised, the length is limited by size of the data set. If the length is too short it can be difficult to make out what is happening in the data, although looping can help alleviate this problem somewhat. High-dimensional data, that of many parameters, can be analysed by mixing audifications together or playing them through separate channels in a multichannel sound system. However, this method, along with earcons and auditory icons, is not suited for generic high-dimensional data sets since they can only reflect a small number of carefully selected parameters. For this reason it would be advisable to adopt a different approach when dealing with such source material.

Parameter mapping sonification (PMS) was first introduced by Scaletti (1994) as first-order sonification and is the current dominating paradigm in the field. This process involves mapping individual data attributes to acoustic attributes making it suitable for data of arbitrary size and dimensionality. In essence, it provides a more flexible approach as both the sound and data parameters can be specified by the designer of the AD, according to the needs of the data analysis task at hand. However, the drawback to this new-found freedom is that the wide variety of mapping possibilities poses a challenge in terms of consistency and comprehensibility. It is often the case that a PMS may be difficult to interpret without explicit knowledge of the employed mapping, hence the specification of a good mapping can turn out to be a nontrivial requirement. This is such a well known issue that it has been dubbed by some as "the mapping problem" (Worrall, 2011) and has prompted several approaches toward finding a solution (Walker & Lane, 2001; Schertenleib & Barrass, 2010; Worrall, 2010). Many of these solutions will be discussed later in the next chapter where they can be put into context in order to support an appropriate resolution.

Model-based sonification (MBS)(Hermann & Ritter, 1999) was designed to overcome data mapping issues associated with PMS by taking 'important dimensions of sound semantics into account by grounding them in physical sound generating processes in a natural and user-transparent way.' (Hermann & Ritter, 2004: p. 730). Drawing inspiration from physics, its inception was based on how acoustic elements were generated in response to our actions. As such, interactions with our environment could provide a general framework (Hermann & Ritter, 2004) to define, design and implement specific task-oriented sonification techniques. A design or instance obtained with MBS is referred to as a sonification model which, under typical circumstances, does not generate sound in the absence of any form of excitation. This could include a virtual object, or a parameterised sound model, combined with a tangible user interface (Figure 1). Interaction with the model follows the real-world analogy where objects can be manipulated to obtain information about their structure and substance.

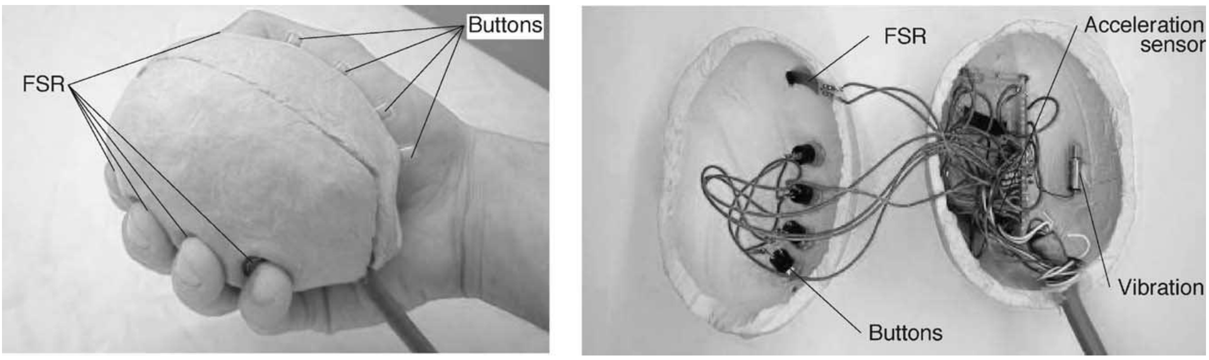


Figure 1: A tangible user interface containing force-sensing resistors, two accelerometers and a motor for interacting with a sonification model

A major disadvantage of adhering to this approach is that the sound design is restricted by the model, where the model itself renders the sound. This led to others (Vogt, 2010) observing that, unlike PMS, MBS was not ideally suited for linking specific data and sound parameters using a discrete metaphoric approach. Hermann (2011) also commented on a drawback in the sound design process when stating that 'since MBS constructs virtual sounding objects from the data, their sound synthesis is as complex as the numeric physical modelling of acoustic instruments, and full-quality rendering of this may exceed the available computer power for many years.' (p. 423).

2.5 Interactive sonification

One of the initial studies for interactive sonification (Hunt & Hermann, 2004) provided a framework that emphasised the need to rethink our interaction methods with information technology and make them more akin to our actions in the real-world. This included looking at the manner in which humans physically interact with the world and how sonic feedback was an important part of this process. The same authors organised the first meeting specifically for this discipline in 2004 under the title of ISon, from which some selected works were later reissued. Since then there has been a small number of additional conferences whose proceedings (Hermann & Hunt, 2009) provide many examples where there is a focus on real-time interaction with the AD.

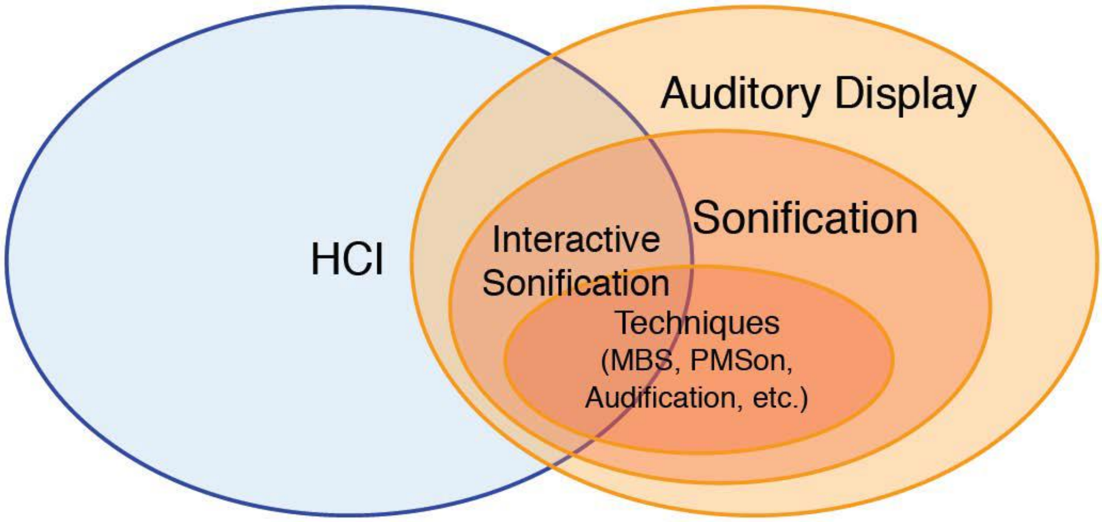


Figure 2: A Venn diagram showing the related fields for interactive sonification

Hermann and Hunt (2005) define interactive sonification as 'the use of sound within a tightly closed human-computer interface where the auditory signal provides information about data under analysis, or about the interaction itself, which is useful for refining the activity' (p. 20). This puts the user at the heart of a real-time interactive control loop for the purposes of exploring and interpreting data through its transformation into sound. Here, the model encapsulating the data produces an immediate sonic response when subjected to excitation or interactive exploration, thus creating the tightly closed loop between human, data and sound. Hunt and Hermann (2011) believe this research area to be a subdiscipline of sonification, placing it within the overlap of AD and HCI while portraying the various mapping techniques as core members of the sonification process (Figure 2).

Interactive sonification shares much of the theoretical framework behind the MBS methodology, coming from the same group of authors responsible for introducing and defining both areas. In both cases there is a reliance on the user to interact with and manipulate the data, placing emphasis on the interface. In the context of HCI, Hunt and Hermann (2011) consider the most important types of interaction as 'physical manipulations, where sound is a highly informative by-product of human activity' and 'the intentional use of objects to actively create sound, as in musical instruments' (p. 275). Regarding musical instruments as highly developed and successful models of interaction with sound, and an ideal hypothesis for HCI optimisation, would find sonification seeking inspiration in the form of music once again. Recent developments in interactive sonification design have continued focusing on creating a direct and dynamic interaction that strike a balance between functional and aesthetic experiences (Weinberg & Thatcher, 2006). This would see work from designers who initially employed a more aesthetic approach in creating a compositional environment (Barri, 2009) as suitable candidates for adaptation to more scientific purposes (Snook & Barri, 2010). Irrespective of their location on the art-science continuum, much of the work demonstrates a metaphoric grounding of data parameters through recognisable physical sources. In other words, interaction may affect the sound in a complex manner but since the sound generation process is represented by a physically inspired intermediary model it can make the connection between sound and data appear more natural. In this regard, if sonification aims to close the semantic gap between the user and the data through sound then interactive sonification aims to close the semantic gap between the human interface device and the sonification process.

2.6 Summary

The brief overview for this chapter was to serve as a preface for the next where we will look at the design aspects of an AD. Having established the location of interactive sonification within its parent fields of HCI and AD we are provided with a research area on which to focus and explore. In doing so, it will become apparent that there are common trends in both fields that are important to contemplate when considering interaction, data representation, aesthetics and musicality.

CHAPTER 3 - DESIGN

3.1 Introduction

Video game engines are freely available and often come with an extensive code base that includes, or can readily incorporate, physics engines. Nguyen (2012) favoured their use in the AD design process stating that 'many game engines offer developers a level editor, sandbox or toolkit that enables interaction in 3D virtual environments and thus interaction with real-time dynamic data streams and multimodal feedback.' (p. 107). While this description may ostensibly depict an ideal medium for interactive sonification it is important to note that the development of physics engines is an ongoing process where formative decisions are not necessarily focused on musical activities or concerned with AD. As such, there is still the need to place many of these positive elements within the framework of a design process which aims to enrich both interactive sonification and musical aspiration.

There are many approaches one might adopt in the creation of an AD (Frauenberger, 2009). While useful as strategies supporting the sonification process, none of them provides a complete framework for a comprehensive design. This prompted Frauenberger to suggest that 'the guidance in this area is as diverse as the application domains and the context of the work often dictates the stages in the design process.' (p. 38). The same sentiment can be argued for the design described in this thesis. From a scientific point of view a rigid body simulation lends itself to a model-based approach employing a mediating metaphor, used in the analogical sense of a virtual object, from which we can characterise the data and tasks involved. On the other hand, when considering Mhyusics as a creative tool, a greater freedom in the sonic interpretation of the data would imply that a parameter mapping solution would be more suitable. Hermann et al. (2011a) also appreciated that each of the sonification techniques 'has its favoured application domain, specific theory and logic of implementation, interaction, and use.' (p. 5). For this reason they preferred the idea that 'none of the techniques is superior per se, and in many application fields, actually a mix of sonification techniques, sometimes called hybrid sonification, needs to be used in cooperation to solve an Auditory Display problem.' (p. 5).

The approach laid out here has much of its focus on the user. From this perspective it seeks to characterise the interactive process in terms of data and its corresponding representations. Similar approaches to interactive design will often find authors drawing inspiration from traditional musical instruments (Hunt & Wanderley, 2002; Levitin et al, 2002; Hunt & Hermann, 2004; Jordà, 2004).

Musical instruments provide a rich heritage of tools which allow humans to produce complex and expressive sounds. More pertinently, they are capable of demonstrating how information is perceived from distinct sensory modalities where multimodal inputs are combined in a coherent way through interaction. Kramer (1994a) outlined their similarity in structure to an AD when viewing the act of performing on a musical instrument as special case of data generation where the data is supplied in the form of physical movements. This notion was later echoed by Hunt and Hermann (2011) when they distinguished interactive sonification from musical instruments. Here, they regard the former as transforming 'data into sound (modulated and controlled by human gestures) for the purposes of data analysis' (p. 275) while the latter transforms 'human gestures into sound for the purposes of expression.' (p. 275). In both cases we are dealing with a transformation, or mapping process, in order to achieve a certain goal. As such, this process can be seen as the crucial factor of general AD design where the key is to understand how the user makes sense of a data set through technological interaction.

Technology brings with it a sense of disconnection which, when viewed in terms of an AD, creates a gap between the data, its representational modalities and the user. This presents us with a freedom of choice during its conception when attempting to forge a link between the comprising elements. In the context of musical interface design Hunt and Wanderley (2002) understood the significance of mediating this gap when stating that 'explicitly defined mapping strategies present the advantage of keeping the designer in control of the design of each of the instrument's component parts, therefore providing an understanding of the effectiveness of mapping choices in each context.' (p. 98). In the more specific sense of a physics simulation, Cadoz et al. (1993) were also aware of its presence from the outset of their design when they wrote:

> 'any computer simulation of a physical object relies on a digital calculation process. To make this process useful, we need a bridge to the world of perceptible phenomena.' (p. 20)

What follows is an analysis into the process of building this bridge by defining and outlining its integral elements in terms of mapping, metaphor and magic. In turn, this anticipates examining the role of each element in the relationship between data, audio and video where careful consideration of the user helps to establish a solid foundation underpinning the interface and its principal components. Each component is subsequently evaluated in order to ascertain the scope of their application by providing relevant examples, some of which are accompanied by external demonstrations. Advantages of the current implementation of this design are then ascertained and discussed within the context of a comparison to similar work.

3.2 Understanding the gap

3.2.1 Mapping

Barrass and Vickers recognised that all perception of data, including visualisation and sonification, involved some level of mapping. They stated that 'any time something is represented in a form external to itself, a mapping takes place; an object from a source domain is mapped to a corresponding object in the co-domain (or target domain).' (p. 153). As such, any AD that maintains an audiovisual approach should consider mapping as fundamental to all aspects of its design.

The mathematical definition of mapping is synonymous with that of a function, which is described as uniquely associating members of one set with members of another set (Arfken, 1985). When thought of in computerised terms it can be seen as how two sets of parameters relate to each other. Hunt and Wanderley (2002) devised three generalised categories of mapping configuration based on similar classification in the literature. One-to-one expresses cases where one parameter is mapped to another parameter, one-to-many relates one parameter to two or more parameters, and vice-versa for a many-to-one arrangement. In all cases there are several important factors that should be taken into account. Decisions must be made as to which parameters should be exposed and how many of them should be involved in the mapping process. From a composer's perspective Doornbusch discussed the relative cognitive weights of increasingly complex mappings stating 'there is clearly a limit to how far such complexity in mapping can be taken before all sense of the original data is lost in the mapping.' (p. 155). As a general guideline for sonification design de Campo et al. (2004) believed that a balance between complexity and immediate intelligibility should be sought where the designer would bestow the user with fine control over the parameters. While this ideal may appear to grant total freedom it also comes with a large caveat as Grond and Berger (2011) made clear when stating that 'the wide variety of mapping possibilities poses a challenge in terms of consistency and comprehensibility.' (p. 364). Indeed, the process of assigning acoustic parameters is a complex task that relies on extensive empirical testing in order to determine the best choice in any given scenario. Some of the applicable elements comprising such a task will be addressed when progressing throughout this chapter, while a more detailed general inquiry can been found elsewhere (Worrall, 2009b; Grond & Berger, 2011; Walker & Nees, 2011). A prominent point in this regard is that the transfer function, or rule of transformation, must be carefully considered as factors such as range, polarity and scaling play an important role between conceptual data dimensions and perceptual acoustic dimensions.

In the context of AD design a good mapping can be described as that which facilitates the communication of information and advances understanding. As such, it is important to be aware of how much information is being acquired by the recipient in the manner that is intended its architect. This is a particular problem in the sonification process as 'mappings of conceptual data variables and auditory percepts are rarely homogeneous, i.e., judged similarly by different people.' (Vogt, 2010: pp. 90-91). Therefore, in order to assess what can be deemed illuminating rather than obfuscating, we need to seek common ground on which to assess the data transformation process. One accepted solution to this problem is the use of metaphor. Metaphor is often employed as a means of comparison and in the subsequent section it will become apparent how it is applicable to the architecture of this AD. Here, it can be seen as building on the more mathematical relationship of mapping by placing it in a recognisable context, thus strengthening the bridge between the data, audio and visual elements.

3.2.2 Metaphor

In their design of a generalised sonification environment de Campo et al. (2004) recommended the employment of metaphor for data with less immediate connections to audible qualities. When examining the process of metaphor it becomes clear why this should be the case. Metaphor and data transformation share a similar aim of elucidating meaning in one phenomenon by describing it through another. Lakoff and Johnson (1980) showed that metaphors are a fundamental part of our thought processes whenever we try to think abstractly and are instrumental in shaping our perception of reality. They claim that all meaning is grounded in our corporeal experiences and that many connections between two disparate ideas can be shown to depend on mental images imparted through our vocabulary of spatial and physical metaphors. Furthermore, these metaphors persist in the lexicon when they excel at comparing and conveying information, emerging as what is termed a dead metaphor. Dead metaphors are conventional in their use and are not generally ascribed to any particular person but instead to common sense, reason or intuition. As such, a collection of these abstractions can be seen as a general consensus of meaning when shaping the path of communication. For this reason there is a general agreement across the appropriate fields that by expanding the knowledge base we can introduce more people to new concepts of connection through the medium of metaphor in the hope that they may become equivalent to a dead metaphor.

Ballas (1994) considered metaphor to be at the core of sonification techniques where he proposed that an engineer should aim to produce dead metaphors. Blackwell (2006) also calls for common sense during his reflection upon the relationship between the designer and the user stating that 'constructing effective metaphors is to some extent a "literary" skill.' (p. 494). With respect to historical views, he showed that guidelines for software interfaces are often in favour of metaphor since the commercial success of the Apple Macintosh in 1984. This has since led to a wealth of implementations which employ metaphor, though not always successfully due to a tenuous link with their physical counterpart. Consequently, he submitted that 'the best advice to a designer would be to follow his or her common sense, as it is common sense rather than analytic theory that will correspond most closely to the user's own natural understanding.' (p. 520). This opinion of seeking a common ground between two disparate parties was also expressed by Fels et al. (2002) when they determined a methodology to overcome problems associated with the physical separation of control from sound. These authors considered musical expression as an act of communicating meaning where both the player and the listener are involved in the understanding of the mapping between action and the resultant sound. They understood that new controllers required both the audience and the player to learn the connection between unfamiliar interaction gestures and interfaces. This motivated them to believe that both the mapping and ease of understanding are critical in determining the success of an instrument and that metaphor is crucial in this regard. From this standpoint they propose that initial mappings based on strong, intelligible metaphors can establish a good base on which to diversify. Put more succinctly, they determine that:

> 'metaphor enables device designers, players and audience to refer to elements that are 'common knowledge' or cultural bases which we call literature. By grounding a mapping in the literature, it is made transparent to all parties.' (p. 109)

With this notion comes the association of culture with common knowledge, indicating that we must be aware of its role in structuring the communication of ideas. Anacleto et al. (2006) described culture as a 'shared meaning system which forms a framework for problem solving and behaviour in everyday life.' (p. 1). Here, 'individuals communicate with each other by assigning meaning to messages based on their prior beliefs, attitudes, and values.' (p. 1). In this light, they view common sense as knowledge possessed by most people from various cultures, where it refers not to good judgment, but to information learned from the everyday interaction with our environment. Ma et al. (2010) advocated the understanding of cultural sensitivity when establishing that age and personal experience have a profound influence on our understanding of sound sources. They recognise that some sounds are timeless while others decline over time. This is especially true in relation to technology where sounds from older appliances, for example a modem, can prove unrecognisable to younger generations. Additionally, regional based sounds, such as radio jingles, highlight the significance of personal experience as they will often only be recognisable for a limited demographic.

When exploring the potential causes of conflict across cultural boundaries, the designer themselves should also be acknowledged as their decision making will often be guided by their own cultural perspective. In the field of HCI a mental model uses structural analogies to the real world in order to provide predictions for understanding the interaction process. Therefore, the designer will bring their own set of preconceptions for these analogies based on personal experience. Blackwell (2006) submits that 'when a company designs a "mental model" as an engineering surrogate for an actual person, this company is really designing the user to be a better customer for their products.' (p. 513). Similarly, Barrass and Vickers (2011) wrote that the 'pragmatist aesthetics perspective reconciles us to the assertion that user experience may only be designed for, that we must do all we can to maximize the opportunities for meaningful dialogue with our sonifications, but recognizing that the experience will not be universal.' (p. 160).

Many of these arguments demonstrate that we should refer to established theoretical findings In order to seek a successful model for shared interpretation of design. Arguably, one such model involves the implementation of simulated physics within software interfaces. Furthermore, it is apparent that this activity is undergoing considerable changes that are encouraged, in part, by the ongoing developments of the physics engine. When applied to the architecture of a virtual interface, there is evidence to show that physics provides a persistent consolidation of experience that is inherent to many cultures. Hermann and Ritter (2004) argue that the universal relationships between events and sounds in the physical world 'provide important design guidelines for the creation of auditory displays in such a way that they facilitate an immediate and natural perception of meaning event without prior training.' (p. 734). Jacob et al. (2008) shared this view when they relied on the users' knowledge of the ubiquitous non-digital world to act as the framework behind their idea of reality-based interaction. Kramer (1994c) argued that fabricating a Newtonian reality would produce a more comprehensible AD, referring to it as an ecological validity where auditory variables change in natural ways. Similarly, Fels at al. (2002) proposed that grounding a computerised musical interface in a mechanical system made the 'mapping between instrument control and sound production psychophysiologically transparent for the player.' (p. 110). They believed that increasing such transparency reduced the cognitive load placed on the user and therefore allowed them to perform a greater number of concurrent tasks.

While there is further evidence to support this argument throughout the course of this chapter, we must exercise caution before confiding in this ubiquitous metaphor. Blackwell (2006) observes that 'it is difficult to disentangle researchers' own experiences of using metaphorical interfaces from their creative experiences of imagining new metaphors.' (p. 511). He argues that dead metaphors such as the desktop can impose the creative experience of the user but believes that while good ideas will persist, they can always be built upon. Indeed, when consulting the literature we find the progressive enhancement of a core metaphoric principle concerning simulated physics. Since the introduction of direct manipulation (Schneiderman, 1983) and the desktop metaphor we have been portraying data files as virtual objects, or icons, capable of being transposed via "drag 'n' drop" gestural interaction and buttons that can be pressed when clicked. As technology improves we continue to seek further embodiment of physical principles into interface design including velocity induced gestures with inertial response. Be that as it may, if the full implications of the computer's power are to be acknowledged, we must also encourage going beyond physical imitation. This was made clear by Jacob et al. (2008) when they wrote:

> 'simply making an interface as reality-based as possible is not sufficient. A useful interface will rarely entirely mimic the real world, but will necessarily include some unrealistic or artificial features and commands. In fact, much of the power of using computers comes from this multiplier effect—the ability to go beyond a precise imitation of the real world.' (p. 205)

As the potential of the computer is too vast to cover in a general sense, it is instead defined here in terms of enriching and augmenting a metaphor for the sake of this design. This strengthening element is regarded as magic and in the proceeding section it will be outlined through the deliberation and confirmation of an appropriate set of boundaries.

3.2.3 Magic

Blackwell (2006) expressed that 'the abstract world of computation empowers users with creative and magical potential.' (p. 511). In particular, he noted that in presenting a simulated copy of its real-life counterpart, the virtual environment metaphor 'extends the supposed power of the computer from the mental realm into the physical world.' (p. 517). In this regard, Keislar (2009) gave a generalised overview of the potential data processing implications when he wrote that 'the computer's increased precision, power, fidelity, flexibility, extensibility, and reproducibility augmented analog technologies' impact on music.' (p. 21). A summary of the qualities that the magic of computation can provide was also given by Jacob et al. (2008) in the form of expressive power, efficiency, versatility, ergonomics, accessibility and practicality. Their reality-based interface framework viewed these elements as tradeoffs for less realistic analogies where 'designers must strike a balance between the power of their interface and its level of reality.' (p. 205). This in itself can be viewed as an axiom of the design process in the sense that we are starting with a computerised emulation of reality, the physics simulation, and adapting it for the purposes of interactive sonification.

During the adaptation process we can seek to apply these attributes as a means of enhancing our perception of reality or diverging from emulation and venturing outside the realms of what we normally encounter. This is not without its limitations as we are still constrained by our physiological limitations in both the physical and virtual worlds where 'our experiences in them are dictated by what we can perceive and understand.' (Kramer, 1994a: p. 56). However, when mimicking physical objects through simulation there is still plenty of scope in which they can embrace extraordinary properties. For instance, in the desktop environment we are presented with an office filing cabinet metaphor where we can find files associated with named folders. With the addition of computational power these folders are granted instant, often automatic, sorting of their files based on a particular rule. Crucially, the metaphor provides control parameters which are less abstract and more closely related to our experiences, while the computer goes beyond what we would normally anticipate. When simulating more advanced behaviour of physical objects the same idea still applies. Menzies (2008) noted that when empowered with the freedom of control over the data defining rigid bodies, sound synthesised through physical modelling algorithms could be greatly personalised:

> 'Although grounded in physical behavior, and therefore naturally appealing to human psychology, the intimate interactions can be tailored to more unusual simulations that would be difficult or impossible in the real world. For instance very deep resonances can be easily created that would require very heavy objects, and unusual resonances can be created. Likewise, the parameters of surfaces can be composed to ensure the desired musical effect. The physical behavior of objects can be matched to any desired scale, of distance, time or gravity.' (p. 75)

Kojs et al. (2007) attested similar ideas in their conception of Cyberinstruments, describing them as 'inherently responsive to modifications, thus facilitating a myriad of abstract sonic identities.' (p. 65). They showed that by characterising a musical instrument in terms of computational information we can extend the possibilities of material forms while closely relating to musicians' expectations. Again, certain aspects of this can be thought in terms of metaphor and magic. The physical model presents users with a customary analogy for assessing action and response. However, the possibility of computer-based control allows the user to overcome their physiological limitations and embrace complex, or normally unplayable, compositions.

This discussion has included a small selection of examples to demonstrate that depicting physical objects within virtual domains can be thought of in terms of metaphor and magic. As we continue to define the AD, further evidence will emerge to support this argument. Within this framework rigid bodies can be customised for more unfamiliar purposes in order to communicate a wealth of fresh and dynamic ideas through the medium of sound.

3.2.4 Summary

Keislar (2009) stated that 'the general-purpose nature of computer technology enables abstractions, disjunctions, and remappings at all points in the music-making process.' (p. 17). In order to bridge these gaps we have an incrementally structured method in the form of mapping, metaphor and magic. While the mapping process is fundamentally a mathematical procedure, the introduction of metaphor presents a common ground for its understanding which, in turn, is enhanced by the additional element of magic. The next section will deliberate a number of concepts that can be upheld by these central factors when analysing the relationships between data, visual and audio. Here, the relevance of the user is a prevailing factor as it is their comprehension of these relationships that will influence the design of the AD's interface and its comprising components.

3.3 Establishing connections

3.3.1 Establishing a connection between data and video

As one embarks on the design process of a sonification, understanding the chosen data set becomes a significant area of interest. This can bring with it some considerable challenges as Worrall (2009b) points out when he describes it as being the first major bottleneck:

> 'In data sonification, whilst the input data can be thought of as eventually controlling the sound rendering, the transformations it has to undergo in the interim can be considerable. Such data processing can reasonably include multidimensional scaling, filtering and statistical analysis which itself may itself become the subject of sonification.' (p. 5-2)

According to Hunt and Hermann (2004) we face at least two fundamental problems in the domain of data exploration. First, 'the data often inhabit a high-dimensional data space that is very different from the 3D space we are familiar with' (p. 6) and secondly, 'these data spaces have no intrinsic means of interacting with them.' (p. 6). Both of these issues can be addressed when the data is derived from a simplified analogy of physical objects and placed in the framework of a game engine. In most game engines the correlation between data and video has been well established by allocating polygonal models to depict this computational information (Table 1) in a faux, three-dimensional Euclidean space. At the same time, Hunt and Hermann have proposed that 'model-based approaches may offer the chance to bind together different modalities into a useful whole, both for display and interaction purposes.' (p. 6). Their method requires the designer to map the data to an intermediary model rather than directly to sound. Accordingly, we can regard these polygonal models as one of the most fundamental forms of MBS, employing a data set which is commonly accepted as being defined in this manner. This brings many of the benefits of MBS but places the intermediary model in the virtual environment, allowing it to be viewed from any number of angles and at an arbitrary distance. In that respect the models not only contextualise the data and reduce its complexity by means of pre-interpretation, they also instantiate an elementary means for direct manipulation which permits both the user and an audience to perceive the direct relationship between action and effect.

Property	Type	Example
Position	Vector	(11.2, 4.3, 1.0)
Orientation	Vector | Quaternion	(0.0, 0.0, 3.14) (1.0, 0.0, 0.0, 0.0)
Angular Velocity	Vector	(0.0, -0.8, 0.0)
Linear Velocity	Vector	(0.2, -9.4, 0.4)
Angular Momentum	Vector	(0.8, 34.0, 4.8)
Linear Momentum	Vector	(9.0, 1.5, 12.5)
Mass	Scalar	1.8
Scale	Vector	(1.0, 0.5, 1.0)

Table 1: Rigid body data - Common properties of a simulated rigid body alongside their corresponding type

In order to understand the underlying data for the purposes of an AD, Barrass (1998) proposes the combining of task analysis and data characterisation methods. His design approach draws attention to how an AD will be used through the identification of task related goals. In this sense, Hunt and Hermann (2004) demonstrated that we perform tasks and data analysis through our everyday interaction with objects. It would therefore be reasonable to assume that we can apply the same techniques to their simulated equivalent when they are portrayed via a common channel of perception. This was substantiated by Blackwell (2006) who gave the simulation approach to metaphor (Pirhonen, 2005) as an example to suggest that 'safety for the designer lies in mimicking a nondigital artifact in order that the user's actions are predictable.' (p. 514). The environment of Mhyusics Auditory Display is designed so that computer-mediated tasks are determined, in part, by the manipulation of the visually simulated bodies. In turn, these actions are directed by gestures that are mediated through our hand movements. Continued investigations by Cadoz (2009) have observed several gestural methods for the performing of functions in a virtual space. He divided the approach for accomplishing tasks involving simulated bodies into two broad typologies. The first, described as the ergotic function, has similar connotations to that of direct manipulation entailing 'exchanges of energy between a human body and material objects or a material environment.' (p. 217). Conversely, the second, or non-ergotic, function occurs when 'forces, displacements or exchanges of energy are involved only and exclusively with the body of the subject.' (p. 217). It is this second type that leads us to consider the influence of the computer on gestural interaction.

Kojs (2009) digitised human gestures when formulating what he called action-based music. While his cyberactions all derived from the same mechanical principles, they were extended by the computer to create increasingly incongruent actions resulting in unpredictable energy flow and extrinsic sonic response. The virtual environment provided by GENESIS (Castagne et al, 2009) allowed for the reconfiguring of physically modelled entities, such as exciters and resonators, formed new actions and galvanised the creation of novel compositions in cyberspace. These achievements serve to indicate that object related tasks within similar mental models are still open to user interpretation and encourage the prospect of further work. Collateral evidence emerges from Blackwell (2006) who concludes that when depending on metaphor as a visual communication channel 'the design ideal in this case is to provide effective access to this information, while allowing expert users to bypass the metaphor if they already have the necessary information.' (p. 508). Certainly, the display of physically simulated models brings a sense of familiarity, inviting us to rely on our common sense knowledge when interpreting and communicating ideas. However, the typical physics engine also allows us to circumvent the standard model behaviour. For example, let us consider that in reality we witness objects governed by the laws of our environment, the actions of which are exemplified by the physics engine as the dynamic rigid body. Each time the engine updates, it recalculates the body's world transform. This is specified through its scale, rotation and translation, which changes in response to the forces exerted from environmental factors to give the sense of realism we come to expect. Kinematic bodies also comprise the same world transform but are capable of ignoring environmental factors, such as gravity and resistive forces against motion, thus denying our expectations. They permit a one-way transfer of energy when interacting with dynamic objects by admitting an influence upon them but remaining unresponsive in return. Both of these core types can be animated by the user and the computer but only the dynamic objects have the capacity to act after a gesture has concluded. This poses a number of conceivable scenarios where kinematic objects can be employed to restrict or guide gesture and motion relating to their dynamic counterparts. Additionally, changing between the two types during the course of a simulation can incite abrupt changes in movement and drastically alter the energy flow of neighbouring objects.

By relying on established visual relationships pertaining to the real world we are looking to create an ideal foundation for sonifying unique constructions. This is helped, in part, by working solely with a single data set whose properties are constrained by the simulation algorithms of rigid body dynamics. The entities of the natural world, which incorporate the behaviour that these algorithms seek to replicate, have a longstanding synergy with humans. This would suggest that less mental bandwidth is required to comprehend the visual activities and events. Instead, the attentive capacity of the user can focus on the audio, along with its governing mapping configurations, as a means of encouraging sonic exploration. Worrall (2009b) characterises this situation as 'data SONIFICATION' where 'the primary focus is on sound rendering whilst input data is constrained so it can be dealt with adequately by the rendering software.' (p. 5-3). This avoids the alternative method, known as 'DATA sonification', which has an 'emphasis on data-processing tools at the expense of sound rendering flexibility.' (p. 5-3).

3.3.2 Establishing a connection between data and audio

In the design of the AD presented as part of this thesis the connection between data and video requires a polygonal model in order to encapsulate the data set of a simulated object which, in turn, facilitates inherent methods for task analysis. However, it can be argued that the process of task analysis may be enhanced by the introduction of sound, which can coincide with the visual modality to reveal the current state of each model. Hermann (2002) illustrated that the majority of our everyday tasks involving physical objects are accompanied with sonic feedback, which acts to aid us in their processing. This led to the conception of analytical everyday listening (Hermann & Ritter, 2004) which suggested that we could analyse the properties of an object under investigation due to the unique sound each interactive task procured. Both these concepts were based on the understanding that 'the meaning of an acoustic event is primarily rooted in conveying information about important physical properties of an object or process.' (p. 733). This proves to be of some importance when we consider that:

> 'compared to other contexts, the context given by physical laws was stable all the time, so that evolution had ample time to adapt our brains extremely well to the ways how physics

> links sounds and their causes. This is reflected in a number of rather "universal" relationships that are deeply engrained in the way we - usually subconsciously - pick up meaning from sound events.' (p. 733)

Indeed, the idea of invariance between sound and data also supports the act of everyday listening (Gaver, 1993) which upholds the proficiency of the auditory modality for the comprehension of object data arising from physical events, rather than tasks. Nevertheless, while both these modes of listening are relevant to rigid body interactions, they tend to rely on known associations that have been ingrained through experience. Consequently, they may be less suited to more variable cases where the connections between information and sound are yet to be forged.

Kramer (1994a) understood that the user will not necessarily deem data and audio to be inherently linked during their initial experience of a typical AD. He argues that when placing the user in this situation they will automatically begin to identify structures and patterns which emerge from a more complex array of sounds. This perceptual event is regarded by Kramer as an auditory gestalt, which implies a prominent feature of the underlying data and is therefore a crucial factor in the forging of a link between data and sound:

> 'As we learn to use a representation technique and become familiar with its gestalts, we recognize gestalts as signatures of specific events. If the universe of possible events is sufficiently limited, we in effect learn the "language" of the display, wherein each class of gestalts symbolizes a general category of data event or system state.' (p. 27)

His argument is supported by the theory of auditory scene analysis (Bregman, 1990) which distinguishes a number of grouping cues to fuse and segregate streams of sound as a means of identifying their source. Likewise, these cues are instinctively employed in our everyday listening and often resemble those classified by the gestalt principles for perceptual organisation. For instance, both the law of similarity and fusion cues can refer to common audio elements such as timbre and frequency, while the law of common fate and segregation cues can be influenced by periodic components such as those with amplitude modulation. Both sets of principles can be applied to any AD that relies on the user's discretion to form their own connections as their brain naturally learns to recognise circumstances which associate certain sound states with equivalent structural states. Over time, the user will find that less cognitive effort is required to link the data to the sonic representation and that the audio becomes more efficient in categorising underlying data events. Kramer (1994a) compares this process to musical training where 'a student learns to identify certain chords or intervals from what was previously a sea of notes rising and falling.' (p. 29).

Hermann and Ritter (2004) regarded auditory gestalt perception as an important mode of listening which complemented their idea of analytical everyday listening. They asserted the concept of auditory gestalts as analogous to visual gestalts calling them 'a subset of acoustical elements perceptually bound together into a "unit" as a result of a particular coherence, characterized by one of the "gestalt laws".' (p. 738). With this in mind they put forward that sonification models can support both their acquisition and the learning process 'by supplying an invariant process to be used in the same manner for very different data sets.' (p. 739). The same can be said for the polygonal models in this design which reflect the internal parameters of a rigid body in a consistent manner. Accordingly, the search for gestalts becomes more efficient as the visual evidence not only serves to reinforce where gestalts occur, but distinct model behaviour and interaction can predicate their emergence.

3.3.2.1 A foundation for data transformation

De Campo (2009) understood that 'the most general task in data sonification designs for exploratory purposes is to detect auditory gestalts in the acoustic representation, which one assumes correspond to any patterns and structures in the data one wants to find.' (p. 41). In order to ascertain guidelines for the conception of these auditory gestalts he produced the Sonification Design Space Map (SDSM). The groundwork for its design involved the abstraction of well-established sonification strategies 'based on how many data points are rendered into the basic time interval, how many data dimensions are being used in the representation, and how many perceptual streams are in use.' (p. 43). To that effect, the audification, parameter mapping and model-based methods were redefined as continuous, discrete-point, and model-based data representations, respectively. Each of these representations is guided by the fundamental notion of a time frame for the formation and attention of auditory gestalts:

> 'In auditory gestalts (or sound objects) of 100 milliseconds and less it becomes more and more difficult to discern meaningful detail, while following a single gestalt for longer than say 30 seconds is nearly impossible, or at least takes enormous concentration; thus, a reasonable rule of thumb for single gestalts is to time-scale their rendering into the duration of echoic memory and short term memory, i.e. on the order of 1-3 seconds.' (p. 42)

While this specifies a period in which data should be represented, the various representations indicate there is a great deal of flexibility in the number of data points involved. For instance, granular synthesis relies on a high density of short audio events which, via the process of sonification, can derive from an equally high number of data events. Conversely, a less complex tone could express a single data point mapped only to its pitch. In essence, both the definite time frame and the adjustable quantity of data points are regarded here as essential for underpinning the foundation for data transformation.

The connection between data and video established that data values inform the state of a visible model and therefore indicate a MBS methodology. This setup can prove effective for discerning and analysing auditory gestalts. The sonification design distinguished by this thesis permits the allocation of one voice per data set which is manifest as a single rigid body or object. Allocating a subsequent voice to the same object ensures that the previous instance ceases to exist in the same manner that is typical of a monophonic instrument. Each voice is controlled by an ADSR envelope which remains at the sustain stage while the conditions for the existence of this voice are met. The immediate reasoning behind this is to counteract extraneous clicks caused by an abrupt onset, or offset, of audio while adhering to a standard amongst hardware and software synthesis implementations. However, this practice also guarantees that an object will only generate a single sound stream for any given moment in time which can assist the listener in identifying each object to sound relationship, particularly if distinct timbres are used. De Campo (2009) observed that 'when the dimensions in a data set are directly comparable (like here, where they are all power consumption measured in the same units at the same time instants), it is conceptually convincing to render them as parallel streams.' (p. 62). Every rigid body in a physics engine is described from another instance of the same data set, with more complex derivatives adding varying levels of abstraction. As such, each object can produce an audio stream that is directly comparable to another object's stream where multiple instances will generate parallel streams. When recalling the implications of auditory scene analysis we can accept that parallel streams of generated sound may fuse or separate based on a perceptual context. In the context of this design the user can rely on the auxiliary feedback of the visual channel to assist comparisons as to what fuses, segregates or has no discernible effect.

Although it is apparent that MBS has distinct qualities which support the search for auditory gestalts, De Campo (2009) reasoned that 'assumptions built into models may introduce bias leading away from understanding the domain at hand.' (p. 40). One important assumption made by the MBS approach is that sound is generated in response to user interaction. Under typical circumstances objects in a simulation will act in consequence of the calculations performed by the computer. Dynamic objects in particular will continue to act after long after any form of human interaction has set in place a sequence of events. During the course of their actions the data values are continuously informed by the environment, along with other bodies, to produce a wealth of context sensitive information. When factoring in some key elements from a discrete point data representation, or PMS, this data can be directly mapped to sound in a more explicit manner that no longer overlooks the absence of user interaction with the model. De Campo distinguished this representation as follows:

> 'Discrete Point Data Representation creates individual events for every data point. Here, one can easily arrange the data in different orders, choose subsets based on special criteria (e.g. based on navigation input), and when special conditions arise, they can be expressed well.' (p. 40)

By applying this approach to each model the parameters for their voice can be continually controlled by a series of events which derive from their corresponding data set. This is particularly salient when considering that the data set itself is ideally suited for an event based method due to the manner in which it is generated. De Campo acknowledges that 'mapping data time to listening time is metaphorically very close and thus easy to understand.' (p. 40). This was echoed by Hunt and Hermann (2011) who stated in their assessment of interactive PMS design that 'if the data are themselves time-stamped, it is straightforward to map the time value onto the sonification time.' (p. 289). A typical physics engine will generate a new set of parameters for each object every time the engine is updated which, in turn, can be denoted by the designer. In order to achieve a sense of realism the value of this time interval is generally low enough to incur updates at a frequency that either matches or exceeds the visual frame rate. Nonetheless, it is unfeasible to suggest that this data rate could match that of the high audio sampling rate recommended for a continuous data representation as the demands it places on the processor would make real-time interaction impossible. Instead, a discrete point data representation provides a configurable arrangement that can stream parameters at a rate in keeping with the physics engine while effectuating an interactive and dynamic audiovisual environment. It also offers explicit control over the dimensions of each data set making it possible to define a mapping configuration which involves any number of comparable attributes, or subsets, while maintaining that omitted parameters have no influence over the audio. In that respect, the user is given the opportunity to observe de Campo's recommendation that 'generally, when multiple streams are used in a sonification design, the individual streams can and should use fewer dimensions.' (p. 43).

Moody et al. (2006) accepted that 'our experience with objects which do emit sound when we can see they're in motion seems to make our brain much more receptive to linking otherwise unconnected audio and visual phenomena when it perceives a certain similarity in the temporal information of the two streams.' (p. 149). One fundamental aspect of this foundation for data transformation is that both the audio and visual data representations will evolve simultaneously as they both rely on the values obtained from the periodic updates of the physics engine. This creates an ideal environment that can be configured to benefit from this heightened perceptivity. In this case, the speed parameter, calculated from the normalisation of the linear velocity, will inform the model to visually describe the motion of the object, while mapping speed to the amplitude of the object's voice will allow the same parameter to simultaneously act as sonic metaphor for the same motion. Serafin et al. (2011) warned that 'there may be sound categories with very salient sonic parameters which are perhaps very intuitive, yet the sound would be less pleasant for long-term use, or even irritating or provoking an unwanted emotional reaction.' (p. 105). A lengthy time frame also presents an issue for the attendance of auditory gestalt as it demands an unfeasibly high level of concentration from the user. This can be compensated for by the current mapping configuration. By mapping speed to amplitude the user is able to control the length of each voice through a motion based metaphor which, at the same time, highlights the number of active parallel streams.

When incorporating a parameter mapping approach to sonification each data set can become sensitive to special conditions that delegate the flow of information and determine the presence of sound without being solely dependent on, or predicated by, human interaction. By assigning a condition to an interactive model the user is provided with a distinct mechanism over the timing of the model's voice and, in consequence, a period in which to monitor and evaluate the data. This is guaranteed by the envelope accompanying each voice as it maintains a definite level of control that is independent of any mapping configuration. At this point only one condition has been described and it assumes the continuous presence of a voice to be characterised by any combination of data parameters via a mapping configuration. With the introduction of a proximity condition the user has a means of punctuating continuous audio on the grounds that each voice will only remain active while its corresponding object is in close proximity to another object. In this sense, although the voice can be regarded as continuous while the condition is met, the user now has a greater degree of control over which data sets will meet this condition. Moreover, the visually apparent nature of each model facilitates the viewer's perception of the spatial relationship between objects where the automatic sorting and analysis of their data sets can be supported by the gestalt law of grouping by proximity.

While the proximity condition admits a certain level of control over an otherwise continuous signal, the signal still lends itself to a prolonged time frame. However, in the interest of a more concise period of data exploration we can refer to our understanding that sound generally accompanies interaction between two or more rigid bodies. At its most fundamental level this interaction can be described as an impact, or collision. This was understood by Gaver (1993) when he put forward that:

> 'impacts are a basic-level event in the sense that they are produced by a simple interaction of objects; combinations of impacts may produce more complex events such as footsteps, hammering, or bouncing noises. Because they are basic-level events, understanding the information they convey is useful in understanding a great many more complicated events.' (p. 292)

The same notion applies to physics engines where interactions are described through a series of impulses (Millington, 2007). Typically, the duration of each impulse is so brief that any corresponding voice would be hard to distinguish as its time frame would fall short of the criterion for the formation of an auditory gestalt. For this reason, the collision condition allows the voice's envelope to be configured by the user, where its overall length is bound by the same criterion. Each envelope operates as a dynamic tool for encapsulating the acoustic consequence of the kinetic events which, in turn, signify impulsive exchanges of momentum and energy within the simulation. These exchanges can be considered significant as they represent a discontinuity in the data which is discernible from the visual behaviour of the objects involved. This behaviour is also supported by the gestalt principle of good continuation which specifies that abrupt changes along a mapped dimension are indicative of a notable event (Bregman, 1990). Furthermore, when Levitin et al. (2002) identified two distinct scenarios for the manifestation of musical events they affirmed that our auditory perception is consistent with this principle. The first of these scenarios, termed as an explicit beginning, stemmed from the listener's perception of an intensity discontinuity whereas the second, referred to as an implicit beginning, was formed on the basis of a perceived spectral discontinuity. Based on their findings it is evident that the user should be acutely aware of this principle should they wish to modify either the intensity or the spectral properties of a voice. This should apply to all conditions in which a voice is present and not just to the collision condition described here. In that respect, any mapping function that allows for an erratic fluctuation in the value of an audio dimension could unintentionally encourage the formation of an auditory gestalt and serve to mislead the listener, particularly if the underlying data suggests nothing of interest.

Menzies (2008) brought attention to a prominent issue concerning the generation of impacts in a physics simulation when stating that:

> 'even for a surface that is completely solid and smooth, the excitations do not necessarily correspond very well with the surface profile. A contact may jump creating a small micro-impact, due to the blunt nature of the contact surfaces.' (p. 72)

Taking this into account it becomes plausible that the designer must try to compensate for any anomalous collisions detected through engine instability. Offering direct control over collision sensitivity ensures that unwanted detections can be filtered out at lower values whereas higher values can facilitate a greater density of sonic events.

Voice Condition	Description
Continuous	Voice is continuously active
Proximity	Voice is active while the object is within a given proximity of another object
Collision	Voice is active at the onset of a collision (duration is denoted by an envelope)

Table 2: Conditions for a voice to exist

The data transformation design currently presents the user with a choice of three conditions to determine the existence of a voice (Table 2). While all conditions denote a loose time frame for the listener to attend to the audio, their current stage of implementation assumes that the data is to be constantly streamed within this period. In turn, this indicates that a mapped data parameter will continuously update its corresponding voice parameter while the voice is active. As previously established with amplitude, there are voice parameters which may require dynamic control throughout the duration of the envelope. Conversely, dynamic control may not be appropriate for audio dimensions, such as frequency, if the user wishes to generate a series of stable and discrete tones in keeping with traditional Western music. For this purpose, a single value sent once at the onset of each voice presents a viable alternative. Although this streaming condition treats a parameter as static, the underlying data remains dynamic. The static value therefore derives from the current state of the dynamic data each time the voice condition is met. For example, every time an object is involved in a collision it will generate a new voice instance that is informed by the current static state of the dynamic data. In this case, a new collision event would indicate a potential change in frequency. From a creative standpoint each collision can now generate enough variety to be considered musically dynamic despite the frequency being more discrete and controlled. From a more analytical perspective the static value can also be sustained by the voice parameter. This makes it ideal to obtain an auditory snapshot of the data parameter at the time of the event. However, it is worth considering that the condition most appropriate for either scientific analysis or musical ambition will be dependent on both the parameter and the preference of the user. For that reason each data streaming condition (Table 3) can be independently applied to a parameter in order to determine its rate of transfer.

Data Streaming Condition	Description
Voice Synchronous	Send data while the voice is active (data is sent each time the physics engine is updated)
Voice Onset	Send data once when the voice becomes active

Table 3: Conditions for data streaming

This section has described a foundation for data transformation (Figure3) which is supported by a hybrid sonification method known as model-induced parameter mapping (Hermann, 2011). To better understand why this particular approach has been chosen it will now be compared to a more established process for generating the sonic counterpart of a physically simulated object. By demonstrating a fundamental difference with respect to the connection between data and sound it can be argued that the sonification method presented here holds some key advantages.

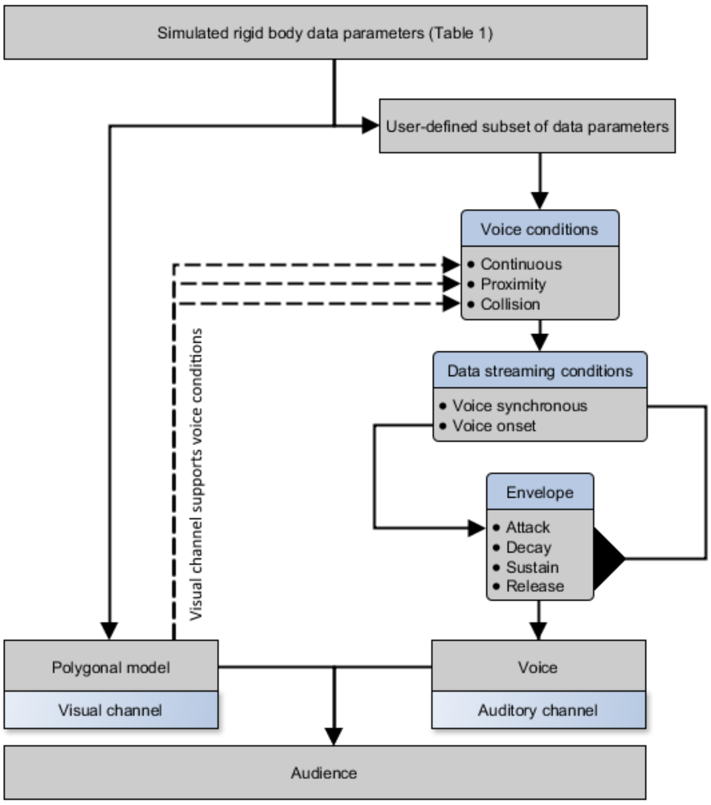


Figure 3: Diagram explaining the flow of rigid body data which derives from a model-induced parameter mapping approach

3.3.2.2 Comparison with the physical modelling approach

A physical model uses mathematical formalisms to simulate the sound source of an object based on the understanding and implementation of its sound production mechanism. Accordingly, a computer can be used to generate convincing replications of existing sources, such as musical instruments, while presenting the means to extend their sonic functionality beyond the limitations of the physical world. Although this methodology can be considered to provide a more intuitive link between data and sound, it is perhaps restricted by its imitative nature.

For both physical modelling and PMS, synthesis is derived from the internal state, or set of parameters, that describe the object. However, the physical modelling approach differs in the fact that both the mechanical and acoustical systems are governed by classical physical laws, denoting the intrinsic link between motion and sound. Cadoz (2002) explained that, with a mass-interaction approach to physical modelling, mechanical movements, or vibration modes, must be constrained to exclusive frequency bandwidths in order to separate audible objects from their non-audible equivalents. By contrast, any entity in Mhyusics Auditory Display can be a source of energy for performing an action, or triggering a chain of events, while providing the user with full control over their sonic activity. This implies a means for a greater potential of gestural interaction while circumventing the need for some imposing restrictions. For instance, Cadoz uses the metaphor of the maraca for the simulation of free particles in a box. This could be seen as a restriction brought about by the tight coupling of the two systems where the number of vibrating bodies involved alludes to the generation of noise. Likewise, Gaver (1993) identifies that there is a one-to-many mapping between the sound and the parameter under consideration, which he perceives as a flaw:

> 'A change in an object's length produces a change of the fundamental frequency of the sound it makes. But other attributes of an object, such as its shape, density, and hardness, also determine its functional frequency. Thus, an impact sound's fundamental frequency does not specify its length.' (p. 309)

The resultant sound is therefore unable to reveal particular information about the event without the use of specified constraints, that is, all other attributes must remain constant.

Ma et al. (2010) found that 'similar materials or textures of the sounds sources, similar effects of the interactions, and similar events that take place can all be the cause of confusion.' (p. 360). In order to better facilitate the communication of ideas they proffer that 'distinctive environmental sounds can effectively evoke concepts (nouns and verbs) commonly used in everyday communication.' (p. 360). Indeed, physically modelled objects that have a large number of parameters in common can produce indistinguishable sounds. This is particularly true when the sound is modelled from a limited subset of data, such as those pertaining to the contact point (Menzies, 2010). However, all these restrictions need not apply to the synthesis routines associated with PMS as they are not determined by the resonating structures themselves but by sonifying any combination of the parameters that attribute to the simulated object's behaviour. In other words, the user can determine the number of object properties that are mapped to sound and thus the level of sonic complexity over which they have control. Ergo, when recalling Gaver's object to sound correspondence it is possible to argue that an impact sound's fundamental frequency can specify the object length should the user choose to link that particular attribute to the relevant sound dimension. Furthermore, given the plethora of available synthesisers, along with their wide range of sounds, the user is presented with a rich palette for communication. This includes the fulfilment of more environmentally accurate material sounds when sending the underlying data parameters to those of a physical modelling synthesiser.

In the same way that electronic musical instruments physically decouple the control interface from the sound generator a PMS would furnish a decoupling of the sound from the acoustic properties of a simulated object. Considering this alternative approach we are presented with a visual metaphor that fits our everyday observations but the sound representation has become more subjective. As such, the user must have greater involvement in understanding how the data dimensions should be linked to the audio dimensions. Menzies (2010) regarded exact physical modelling as a limiter of expression stating that 'explicit physical models are often difficult to calibrate to a desired sound behaviour although they are controlled directly by physical parameters.' (p. 3). He advocates that 'the sound designer is often more interested in the freedom to shape the sound how they would like, rather than exactly matching a real behaviour that may not be quite suitable.' (p. 3). In contrast, PMS provides a more explicit link between data and sound where the user is granted with full control over each stage of the transformation process in order to facilitate a greater capacity for expression. This follows the advice given by Hunt and Wanderley (2002) who suggested that 'explicitly defined mapping strategies present the advantage of keeping the designer in control of the design of each of the instrument's component parts, therefore providing an understanding of the effectiveness of mapping choices in each context.' (p. 98). Understandably though, the user should be aware of the consequences such freedom may bring regarding the transformation of data to sound.

3.3.2.3 The mapping problem

While model-induced parameter mapping sonification can be put forward as a reasonable alternative to physical modelling, it is not without its drawbacks. In context of this design it is framed by a cogent visual model which supports the perception and control of data events behind the formation of auditory gestalts. The additional integration of the PMS paradigm promoted a more explicit level of control over the data transformation process leading to the arrangement and classification of information flow. Despite this, the current framework cannot be viewed as complete as there is still the need to decide which audio dimension may best represent a particular data value, along with a suitable function than can relate the two.

Kramer (1994a) stated that 'spatialized sound can, with limitations, be used to analogically represent three-dimensional volumetric data.' (p. 26). If we recognise that both objects and sound coexist within the same Euclidian space then it can be viewed as a highly intuitive representation. Indeed, video games often exploit this technique and process sound in order to strengthen related illusions of space, distance and motion in a virtual world. However, while this process may work well when expressing object location, it becomes counterintuitive when representing more abstract data dimensions such as mass. This often calls for the employment of an abstract sonic space, such as pitch or timbre, introducing what is widely regarded as "the mapping problem":

> 'The main limitation of PMS is thought to be the known co-dependence of psychophysical parameters: linear changes in one domain produce non-linear auditory effects in another. These perceptual parameter interactions can also produce auditory artefacts that obscure data relations and confuse the listener regarding the parametric origin of the effect.' (Worrall, 2011: para. 4)

Being mindful of this issue Worall (2009b) has previously suggested that 'though conceptually simple, in practice, parameter mapping requires a working knowledge of how the parameters interact with each other perceptually.' (p. 2-16). Walker and Nees (2011) outline many of the salient points in that respect when referring to growing evidence from the field of psychoacoustics. They discuss the detection of sound, one of the more critical determinants for AD design, for which the field holds extensive descriptions of minimum thresholds across various auditory parameters in terms of just-noticeable differences. The concurrent presentation of audio also plays a vital role regarding the detection of sound. McGookin and Brewster (2004) demonstrated that our ability to both discriminate and identify individual sounds decreases in response to an increase in the number of simultaneous sounds. Bregman (1990) helped in this regard by determining what acoustic properties support or inhibit stream segregation. Similarly, Kramer (1994c) proffered further techniques such as scaling down the range of distracting, or forceful, variables. Nevertheless, Flowers (2005) argues that there is most likely a limit to the number of concurrent channels of auditory information before additional streams become impractical for a listener to discern and analyse.

Barrass and Vickers (2011) emphasise that we should consider aesthetics in sonification design while summarising its main issues as 'intrusiveness, distraction, listener fatigue, annoyance, comprehensibility and display resolution and precision.' (p. 57). Much of their work derives from Vickers and Hogg (2006) who maintained that 'once sonification designers have learnt to listen like composers, sound designers, and recording engineers they will be much better placed to create sonifications that maximise the communicative potential of the auditory channel.' (p. 215). Correspondingly, Coleman et al. (2005) carried out a participant observational study of sound designers working within the entertainment industry which highlighted some important design choices for ADs. They discovered that issues of annoyance are prevalent in video games where a major challenge is to ensure that the ambience will remain interesting for multiple hours of playtime and continue to be immersive. From a mapping perspective their findings showed that the right associative sound can make a game more involving and that a purposeful balance is required in order to prioritise game related events. Kramer (1994a) commented from experience that an unpleasant auditory computer interfaces are often turned off even if they serve a useful purpose. He acknowledges a more musical approach when postulating that 'a composer's skills can contribute to making auditory displays more pleasant and sonically integrated.' (p. 52). In this manner, the skill of orchestration demonstrates how to balance the timbres of each instrument in order to present a certain idea or mood. Being mindful of the connotations aesthetics shared with more musical aspirations Vickers and Hogg (2006) made the distinction that what is 'important for sonification designers is not how 'musical' their work sounds, but how easy they have made it for the audience to listen to it, and by listen we mean 'attend carefully'.' (p. 215). In that respect both Kramer (1994c) and Hermann (2008) agreed that artistic license should be restrained for the sake of clarity. They both use the example of reverberation as a parameter which may improve musicality but can obscure the data behind the sound source. Additionally, Hermann warned that 'by using too much noise, it is possible to generate useless sonifications in the sense that they garble interpretation of the underlying data.' (p. 3).

Nguyen (2011) opted to tip the balance between musical aesthetics and scientific accuracy in favour of musicality. He felt that a loss of resolution in the data counteracted for an increase in clarity of data relationships through the listener's understanding of consonance and dissonance; a process first described by Worrall (2010) as soniculation. In this case, Nguyen (2011) suggested that aesthetics 'becomes a mechanism to provide clarity, because it is the discord between data relationships (that is its dissonance) which enables a listener to recognize differences, changes and trends and improves the communication of relational information.' (para. 22). He later employed the same principles in order to highlight the variance between two data values (Nguyen, 2012). The notion that using the organising principles of tonal music to structure the mapping process can produce more aesthetically coherent results has been discussed by Vickers (2006). He puts forward that tonal music is easy to learn, organise cognitively and discriminate, where innate skill permits us to recall melodies without formal training. However, he warns that repeated use of such harmonic tonal sequences and rhythmic structures could lead to mental fatigue.

The cases involving more musical tendencies serve to highlight another important consideration for AD design as the designer must now rely on both a trained musical ear and user subjectivity. Individual differences have a profound impact on the interpretation of ADs and include aspects such as listening skills, learning styles, perceptual abilities and cognitive faculties involving memory, attention, mood and affective association. While some of these factors are ascribed to physiological differences, a large number can also be regarded as a product of cultural background and training. This prompted Kramer (1994a) to note that one of the main disadvantages for an AD is that some users will have less acuity for the display variables, such as timing and pitch. Indeed, Neuhoff et al. (2002) observed that a small number of musical novices 'found nothing compelling or even comprehensible about directional frequency change.' (p. 5). They highlighted the need for training so that accurate labelling of pitch change could be acquired before a user could explore conceptual issues such as metaphor. Additionally, they suggested that changes in frequency 'should be sufficiently large in order to minimize errors in judging the direction of change' (p. 5) but also recognised that perceptual deficits, such as congenital amusia, may limit the role of pitch as an audible dimension.

Kramer (1994c) initially wrote that 'the fact that our attention is drawn more to certain variables than other makes the design of a balanced, or unbiased, auditory display virtually impossible.' (p. 200). Since then the sonification community has continued to encourage experimentation that draws upon knowledge from the aforementioned fields in order to alleviate the factors involved with this problem. Be that as it may, Worrall (2009b) attests that the most common solution to the mapping problem 'is to test empirically which of a number of fine-tunings or "tweaks" of parameter space mappings is the least problematic; perhaps in the hope that eventually, over time, a generalised model may become evident.' (p. 1-2). He notes that such a conventional solution could only be supported by further empirical evidence which, by itself, is a rather demanding task beyond the extent of the work presented here. Instead, as the data under investigation seeks to replicate observable physical behaviour, we can depend, in part, upon our inherent understanding as a guide for comprehensible mapping. By seeking out natural correspondences between audio and visual modalities we are able to ascertain a path towards a more generalised solution within this context of this thesis.

3.3.2.4 Physically informed mapping

Ballas (1994) submits that 'effective delivery of information through sound should produce responses that are quick and accurate.' (p. 86). His studies revealed that listeners could identify typical environmental sounds faster than a non-typical equivalent. Correspondingly, the speeded classification paradigm refers to congruent mappings as those which elicit a quick response. Evans and Treisman (2010) utilised this method to measure reaction time when presenting users with a number of bimodal presentations. Their investigations referred to several experiments in the field that exemplified natural correspondences between visual features of physical objects and abstract auditory dimensions. They, too, believed that such intuitive crossmodal correspondences 'may arise through frequent associations in everyday experience.' (p. 2).

Vogt and Höldrich (2010) advocated that the notion of experience can be integral to AD design:

> 'For a good sonification design, it would thus be enough to know about the underlying metaphors of a scientific theory and the metaphors for sound of these basic experiences. By mapping, e.g., higher energies to what people in our culture perceive as higher in sound, a completely intuitive sonification could be created.' (p. 271)

They sought the advice of physicists in order to create more immediate metaphoric mappings for simulation data and discovered that strong metaphors emerge from their professional knowledge. Their findings unveiled that more mapping associations were offered for the well-known particles and fewer for the rarer. This also applied to everyday physical properties such as mass, which was cited more often than comparatively abstract examples. They proposed that perhaps these observations arose from the subjects' lack of interaction with these properties which, in turn, resulted in them being less prominent in their minds. When constructing an elementary metaphor between particle physics and sound, Sturm (2001b) used common sense as a guide for the correspondence of physical phenomena with auditory dimensions. In keeping the number of arbitrary decisions in the mapping process to a minimum he believed that 'such a direct mapping of both multi-dimensional fields not only allows a cleaner interchange of concepts to enrich both, but also enables an additional level of comprehension of the underlying physical concepts.' (para. 3). This guided him to propose that a sonification would be easier to perceive 'by creating a scheme that requires a minimum of new skills one must possess in order to understand the underlying generative schemes, i.e. the user does not have to train his ears.' (para. 9). Furthermore, it can be argued that the acquisition of these listening skills is heavily influenced by our acoustic ecology (Dietze, 2000), that is, the relationship we have with our sonic environment. As such, this holds important guidance for design which Kramer (1994a) substantiates when referring to our aural experience in order to heed a balance of sonic qualities guided by their natural prevalence. With this in mind he maintained that 'extremely high pitched and loud sounds, which occur infrequently, may strike us as unpleasant or even hurt our ears.' (p. 55). This would also suggest that frequencies should be kept to an audible range so that the listener can carefully attend to them.

Hermann and Ritter (2004) also took a perspective motivated by ecological acoustics, viewing the fundamental laws of physics as connecting 'geometrical properties of our environment to sound characteristics in a rather universal manner.' (p. 731). They discussed many strong correlations that derive from our multimodal perception of physical phenomena. Much of their work expanded on observations made by Kramer (1994c) who gave several metaphorical mapping recommendations based on a similar frame of reference. A suggestion from both parties stemmed from the argument that we expect louder sounds in response to the stronger excitation of a physical structure. Musical instruments are crucial in this regard as they create a shared model for understanding this phenomenon. Musicians are familiar, on some level, that an instrument's amplitude is directly proportional to the energy of their input gesture. This also leads to the anticipation of a steady reduction in sound once energy is removed. Additionally, this model possesses the ability to express other natural links between physical properties and sound such as illustrating that the natural oscillation frequencies of an object decrease with its size. In this case a musician can equate it to a larger string instrument producing a lower pitch range in comparison to a smaller member from the same family, for instance, a double bass in contrast to a violin.

Although it is possible to seek further affirmation that physics-based metaphoric association can devise strong mappings based on our intuition, we must be aware of the significance it holds. By collating what is widely regarded as common ground for interpretation we can inform the design of a standard set of audio mapping rules within an AD concerned with rigid body dynamics. This configuration represents an ideal starting point for the exploration of the data set where the corresponding manifestation can be appreciated by a wide audience. Fundamentally, it serves to provide the recommended criterion in accordance with this design framework as part of a course of action towards a potential solution for the mapping problem.

3.3.2.5 Towards a potential solution

Kramer (1994c) accepted that by facilitating a structured empirical process 'the system user may listen to the same data a number of times with different mappings and decide which mapping most satisfactorily displays the meaningful contours he is looking for in the data.' (p. 201). Walker and Nees (2011) indicated that this idea still endures today when stating that 'sonification designers need to be aware that not all mappings are created equal, and must use a combination of empirically-derived guidelines and usability testing to ensure the message they are intending to communicate is being received by the listener.' (p. 24). However, Kramer (1994c) was also aware that a variety of mappings approach was the most unbiased user-orientated way of dealing with the mapping problem. In this sense, one could argue for a heuristic approach more applicable to a compositional process that encourages experimentation and creativity.

It has been shown that sonification was practiced in experimental music as a metaphor for natural phenomena long before its concept was established (Straebel, 2010). In this manner sonification becomes a blank canvas for individual interpretation, based on a series of mapping decisions, where an audience can reflect on how the final arrangement was informed by cultural and physical experience. Furthermore, when seeking composers' views on such matters, Doornbusch (2002) revealed that compositional structures and mappings were unique to each composer, with no set method for data transformation from the domain of the conceptual, gestural or structural to the musical. He concluded that 'the nature of composition and composers means that there will never be a 'solution' to the mapping problem in algorithmic composition, that it will remain a part of the exploration.' (p. 155). Be that as it may, Walker and Kramer (2005) have determined that this mindset is not entirely suitable for more objective sonification. They showed that relying on a designers' perception alone could lead to worse performance where less desirable choices for metaphoric mapping often performed better. This served as a notable case to prove that, like any mental model, intuitive mappings can not necessarily be predicted a priori, although we could argue that common sense pertaining to physics explains why their less obvious mapping of size to onset worked well when attributed to inertia inducing slower changes in larger objects. However, they put forward that 'guidelines may emerge, based on this kind of study, so that designers will not need to guess about effective mappings.' (p. 412). In turn, this will enable them 'to deploy more truly "intuitive" mapping ensembles that actually allow a listener to understand the message the display is meant to convey.' (p. 412).

When contemplating these arguments it would stand to reason that there is a need to strive for a balance between strict guidelines and free interpretation in order to facilitate a progression of knowledge. Kramer (1994c) originally proposed that 'the metaphorically or affectively associated map may be generalizable to different sets of data and various sound generation techniques, flattening the learning curve and making the sonification process more efficient.' (p. 217). His idea that the sharing of mapping strategies could expedite the learning process has since been explored by a number of authors. Whereas some provide a considered list of citations for guidance (Walker & Nees, 2011) others have chosen to deliberate and create online databases (Barrass, 1996; Walker and Lane, 2001; Ma et al, 2010; Schertenleib & Barrass, 2010). Such work has shown that a number of factors must be taken into account in order for a database to be authoritative including categorisation with respect to cultural background, sensitivity to context and, where possible, further detailed annotations pertaining to conception and effectiveness. It is through such a collective agreement of well-organised findings that we attain a picture of how mappings can then be effectively applied within the appropriate AD designs.

Converse to such arguments, Kramer (1994a) remarked that prior knowledge could conceivably 'act as an obstacle to fresh insights.' (p. 17). Blackwell (2001) also discovered that allowing users to construct their own meanings from dissimilar elements using metaphors, rather than being shown systematic and consistent mappings of the kind recommended in HCI textbooks, resulted in improved learning. As a consequence it is worth remembering that these mappings are informed suggestions and should not be regarded as a panacea for a particular purpose. However, when presented in an informed manner they can elicit guidance for further experimentation while granting us insight into the techniques behind their architecture. Doornbusch (2002) endorsed this concept when preparing a collection of compositional mapping strategies for the prospects of sharing knowledge. He submitted that 'in this way, mapping in algorithmic composition could be demystified and more complex, varied and musically appropriate practices could be developed by building on the work of others.' (p. 155). Similarly, Clarke (2005) introduced an interactive aural framework to elicit basic components of compositional techniques for others to analyse. While serving as an insight into the mind of the composer, his approach allowed one 'to develop a deeper understanding of the compositional process and of the way in which the work is structured.' (p. 87). He noted that this would not only allow for a means of orientation within an otherwise abstract collection of sound but also presented an opportunity for the user to construct unique interpretations. Kramer (1994a) contributed a corresponding notion for AD design when stating that 'becoming familiar with specific system states may improve the data exploration process by providing verifiable system states as references.' (p. 17). This was later reiterated with a recommendation that the ideal AD would 'be easy to get started with, providing a graphical user interface (GUI) and "canned displays" with standard hooks for different types of variables.' (Kramer et al, 1999: p. 12).

Certainly, pre-configurations of such systems have already provisionally demonstrated that the sonification of physics engines can immediately engage users with the creative process of musical ideas (Dolphin, 2009a; Dolphin, 2009b). However, these examples place heavy restrictions on the range of objects involved, along with their mapping configurations. In contrast, Tache (2008) delineated how a more complex system, comprising virtual objects, could support similar conventions by presenting an organised collection of constructions. He referred to his ontology of examples as an "instrumentarium" which aimed to familiarise non-experienced musicians with the architecture afforded by the CORDIS-ANIMA system. Tache argued that a unified presentation of possible constructions, alongside parameter configurations, would improve a user's prowess for design. This was supported by Cadoz (2009) who declared that it allows for others to expand upon and perfect original ideas while taking advantage of the digital instrument's duplicable nature. Although such a comprehensive task would be beyond the scope of a thesis observing a theoretical background to an original AD design, a number of demonstrations in a similar vein are included throughout the text. These examples are consistent with the fundamental principles described here, serving to highlight notable ideas, reflect the potential capabilities of the components, and form the basis for future research.

3.3.3 Establishing a connection between video and audio

It can be understood from a mapping perspective that the connection between video and audio has already been established due to both modalities being informed by the same data. Indeed, by employing a model to present this data as a cohesive whole it becomes likely that any parameter which influences the audio will also be responsible for the same model's visual behaviour. With this in mind, the deliberation of further mapping specifics will not be the main focus for ascertaining a connection between the two modalities. Instead, this section will present several cases where the video can support the sonification process while building upon the framework ascertained so far.

Lokki and Gröhn (2005) found that navigation of virtual reality was faster and more accurate when the user could rely on both the audio and visual modalities in lieu of either one alone. Their experiments showed that the omnidirectional nature of our hearing overcame the perceptual limens of the visual modality when drawing attention to important features located outside our main field of view. Video games that attempt to recreate an immersive environment will also combine visual representations with spatially distributed sound to support the visual channel and make the listener aware of surrounding information. This is typical of the first person shooter genre which Grimshaw (2007) distinguished as an employment of sonification for the player to contextualise themselves 'within the spaces of the game world or in relation to events and other characters.' (p. 120). He found that data from the virtual world not only informed sound sources to be sensitive to the location and actions of a controllable avatar but also signalled the game engine to process the sound and recreate audible phenomena in keeping with our perception of the real world. Cues for filtering would inform the player about the distance or occlusion of a sound source, reverberation cues expressed the acoustic space in which either the player or source reside, while Doppler effects would coincide with their relative movement. The drive to produce more convincing sonic cues has seen the continued development of various open source solutions which have been assembled into dedicated audio processing libraries such as OpenAL (Creative Labs, 2013). These libraries are designed to be integrated into gaming engines and can be used to sonically enhance other integral parts of the immersive gaming experience such as the physics engine. For instance, their inclusion can present each rigid body as a dynamic sound source when their relevant data is shared with audio processing techniques such as vector based amplitude panning (Pulkki, 1997). With this in mind it is possible to claim that the game engine has the capacity to facilitate the navigation of data in the AD designed as part of this thesis by automatically distributing and processing the voice of each model in a manner that is readily understood.

Hunt and Hermann (2011) suggested one of the shortcomings of PMS was that 'the sound of such an interactive sonification will only make sense to the one who is in the control-loop, since others do not know whether sound changes are due to the data or due to parameter changes performed in the interaction loop.' (p. 289). When introducing a visual representation, as part of a model-induced approach towards PMS, both the control loop and the current state of the audio are exposed for other people to engage with. This was appreciated by Jordà et al. (2007) who expressed that 'the 'audio' feedback presented to the performer in a visual form, intuitively helps the understanding and the mastery of the interface, enabling the simultaneous control of a high number of parameters that could not be possible without this visual feedback.' (p. 141). Moreover, they argued that the same concept applies when the control loop is extended to include the delegation of control from performer to instrument. In this scenario the performer assumes more of a supervisory role by guiding the actions of the instrument instead of attending to every nuance through physical actions. As such, in order to facilitate the monitoring requirements pertaining to this new role, the digital instrument can employ a visual display to 'dynamically 'communicate' the states and the behaviors of their musical processes.' (p. 141). This reasoning also proves remarkably germane for the design presented here when considering that the states and behaviours of the model are governed by a physics engine capable of responding to either human or computer intervention.

Kramer et al. (1999) advise that 'analysis of multivariate data may sometimes require focused attention on individual variables, and at other times require divided attention to allow the listener to detect similarities and contrasts in trends of different variables or streams.' (p. 11). The foundation for data transformation identified that combining the configurable qualities of parameter mapping with the visualisation of the streams that it designates can assist the user in this process. However, those methods were concerned with the data that described the physical behaviour of the model. To avoid any direct influence over this dynamic state of the data set the user can instead choose to exploit the colour of each model as a means for segregation and fusion. According to research pertaining to the gestalt laws of grouping there is a stronger tendency to group local elements by common colour than by similarity of shape (Quinlan & Wilton, 1998). It is therefore likely that the involuntarily grouping of similar coloured entities would draw the attention of the viewer to a single contrastingly coloured object whereupon it is perceived as being outside of the group. The user could employ this visual phenomenon to highlight the sonic characteristics of a particular data set. Alternatively, a more creative interpretation would see it function as the introduction of a solo theme in a musical arrangement.

Another ramification of the foundation for data transformation was that the audio and visual modalities could be synchronised due to the periodic nature in which the data was generated by the physics engine. This type of concurrency was understood by Spence (2007) as an important element for optimising the localisation of data, adding further credence to the use of the game engine's architecture for that purpose. However, he also determined that 'whenever two or more sensory inputs are perceived as being highly consistent (i.e., as being related in a way that they appear to 'go together'), observers will be more likely to treat them as referring to a single audiovisual event.' (p. 66). This phenomenon is also observed by Chion (1994) whose notion of synchresis recognises that our brains can intuitively create a connection between the two modalities despite the lack of any further logical evidence. Nonetheless, he cautions that:

> 'it is not totally automatic. It is also a function of meaning, and is organized according to gestaltist laws and contextual determinations. Play a stream of random audio and visual events, and you will find that certain ones will come together through synchresis and other combinations will not.' (p. 63)

These ideas are applicable to the design of Mhyusics Auditory Display when considering that any meaning in the data is not only perceived through sound but also implied through the visual actions of the objects. The foundation for data transformation was also aware that both the organisation of sound, and the conditions for its interpretation, would be supported by gestalt laws. However, the notion of contextual determination proves to be more flexible when taking into account that the visual representation of the environment, along with the objects that act within its constraints, can assume many forms. Robertson (1990) proposes that imagery from natural scenes would exploit innate visual mechanisms and optimise the process of navigating and analysing data variables. Indeed, the software infrastructure behind Mhyusics is capable of rendering such environments in real-time while enlivening them with rigid body dynamics. Be that as it may, there is evidence to suggest abstract visual representations that are less commonplace would be more appropriate when coupled with a parameter mapping approach that has the aptitude for a wide range of sonic interpretation.

Gaver (1993) demonstrated that the intrinsic link between motion and sound allows us to recognise a number of visible features associated with everyday objects by listening to their involvement in physical events. The implications of this relationship were also understood by Jensenius (2007) who believed that an embodied approach to cognition would suggest that we mentally hear the sound of a sonorous action when observing it through vision alone. He reasoned that 'it is our life-long experience of acoustical and mechanical properties of objects and actions that makes us able to predict the sound of an object-action-object system even before it is heard.' (p. 22). In this case, his notion of an object-action-object system recognised our ability to determine the objects and actions involved in producing a sound. However, our ability to predict was further characterised by what the author refers to as an action-sound palette. This proves to be of some concern as it comprises what we consider to be the complete range of sounds that the actions of a particular object can afford. With this in mind it becomes counterintuitive for us to view the actions of an identifiable object and associate them with sounds outside this range. For example, Jensenius wrote that 'if we heard a baby's scream when the glass hit the floor, it would be beyond the laws of nature and we would assume that the scream was coming from somewhere else.' (p. 23). Fortunately, as the laws of nature informing these objects are easily manipulated in the computer realm it is possible to design counterparts which take on more abstract qualities. Such objects can continue to enact actions which remain familiar to our everyday experience while carrying an appearance that does not imply a limited range of sounds. In that regard, Cooley (1998) found that abstract computer-generated images could accommodate a wide variety of sound, whereas recognisable images of everyday objects would stimulate a critical response from the listener when conflictions with their experience created a lack of sonic verisimilitude.

While complex polygonal models in virtual worlds can closely resemble recognisable entities, it is generally the case that the rigid body responsible for their actions is derived from simpler geometric forms. The most common of these forms are part of a subset of solid geometry, generally referred to as geometric primitives, which includes boxes with uniformly flat surfaces and perfect spheres. The behaviour of these shapes is described using Newtonian mechanics in an environment that is described as a Euclidean space. These idealised components of the simulation affirm that is it is built upon mathematical conceptions more in keeping with Platonic idealism. In that regard, Bernays (1935) submitted that:

> 'the value of platonistically inspired mathematical conceptions is that they furnish models of abstract imagination. These stand out by their simplicity and logical strength. They form representations which extrapolate from certain regions of experience and intuition.' (p. 53)

These points stand to reason when considering that a model-based representation of the rigid body was adopted to provide insight into the data based on our everyday experience. However, the same representation can also be thought of as an embodiment of the mathematical laws behind the dynamics of the data. Forbus (1997) determined that various causal theories of physical phenomena can be better understood through qualitative embodiment providing it is tightly coupled to its quantitative concepts. Consequently, it is possible for the same tight coupling between the model and the data to extrapolate from certain regions of our experience with general physical behaviour. When realised in the context of the AD designed as part of this thesis, a single abstraction could intuitively represent a specific behaviour that derives from rigid body dynamics. This representation could then be combined and arranged with further abstractions to compartmentalise the structure of a wide variety of algorithms which, with the process of sonification, are simultaneously expressed through sound.

Schroeder et al. (2011) claimed that the complex mathematical definitions associated with the physical modelling of a string would become more comprehensible when presenting them alongside an interactive audiovisual representation. Their design was driven by the notion that, during the process of analysis, the attention of the user could shift between the mathematical descriptor behind the motion of the string and the visual model in which it is enacted (Figure 4). In practice, the model would actively respond to adjustments made to this descriptor in order to form an intuitive and dynamic observation of an otherwise abstruse and static equation.

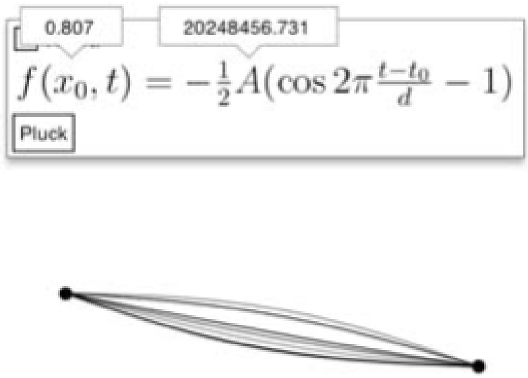


Figure 4: A model of a string enacts the mathematical equation describing its motion

More complex physically modelled structures comprising a series of interconnected springs have also relied on audiovisual models to delineate a language for their construction and behaviour. Kontogeorgakopoulos and Cadoz (2007) submitted that modular representations of computational algorithms make it possible for the average user to design a sonorous structure based on intuition. They contended that 'as the basic building elements actually have a strong physical counterpart they remain pertinent to human senses and create a very realistic mental model.' (p. 277). Furthermore, they observed that it was 'possible to guess and predict the general behaviour of a model by examining its network without the use of mathematical analysis tools.' (p. 277). Comprehending the dynamics of the data in this manner was described by the same authors as a physical thinking approach. Its conception stood in comparison to what they considered to be a more traditional and complex mathematical decomposition, or signal thinking approach, offered by digital signal processing blocks.

Electroplankton (Nintendo, 2012) demonstrated that a framework for musical algorithms based on visual abstractions of rigid bodies can be designed for a more general audience. In the Hanenbow algorithm, the manifestation of sound relies on collision events involving a dynamic body whose appearance comprises simple abstract shapes. The cause and structure of the dynamic body's movement are assigned to abstractions of kinematic bodies which either control its initial trajectory or redirect its path upon collision. By interacting with these representations of logic the user can grasp how each condition manipulates the timings of the musical events. The strength of this design was remarked upon by Dolphin (2009a) who wrote that 'the symbolic representations of the algorithmic musical processes allow quick comprehension of the compositional controls, allowing a reasonable level of compositional intervention by even a novice player.' (p. 213). However, it can be argued that this design is further enhanced by its direct ties to the quantitative underpinnings of the algorithm. This connection is revealed to the user when they enable the display of the angles at which the kinematic objects are placed (Figure 5). With this additional level of information the user gains more accurate control over the exact timing of the sound. Moreover, they can consistently reproduce a particular rhythmic idea by memorising and recalling a sequence of numbers.

These three examples indicate that visual abstractions derived from simplified simulations of physical phenomena can be used to describe algorithms pertaining to both sound and music. They also demonstrate a range of ideas for how the arrangement between the behaviour of these abstractions and their mathematical underpinnings can be portrayed. The next section will argue that such arrangements are also feasible within the framework of Mhyusics Auditory Display by outlining the foundation for one possible interpretation. In this case, the flexible connection between data and audio ensures that the objects can express the close ties to their quantitative underpinnings in more imaginative ways resulting in a combined rendition of information that supports the creation of a dynamic score.

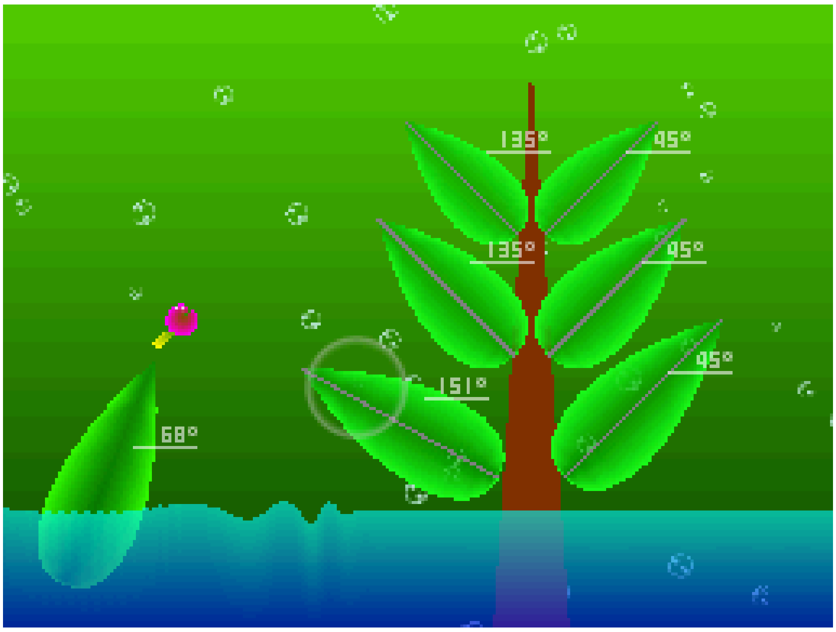


Figure 5: The Hanenbow algorithm is presented using both symbolic and quantitative representations

3.3.3.1 A foundation for a dynamic score

Walker (1987) identified that although cultural and environmental factors can inform visual metaphors for sound, subjects with more exposure to a Western lifestyle predominantly chose crossmodal metaphors ascribed to musical training. These metaphoric relationships included the linking of frequency with placement, amplitude with size and duration with horizontal length. Such a widely recognised body of evidence can also be considered pertinent to sonification based on the understanding that, as part of musical training, the reading of a traditional score teaches the pupil to associate visual representations with particular attributes of sound. In that respect, Grond and Berger (2011) found structural similarities between a sonification and a traditional score when considering that 'music notation is, in a sense, a mapping from visual representation into pitch, loudness, duration and timbre instrumental space with time progressing according to an independent tempo map.' (p. 369). Accordingly, each model instance in Mhyusics could symbolise a traditional note when the data it represents is mapped to the same musical elements. On the other hand, as this analogy is restricted by traditional design it is perhaps less appropriate for expressing the wide range of musical intention associated with computer music. Clarke (2005) noted that a major limitation of working with traditional notation for the analysis of computer music is that it is:

> 'designed to place events on a two-dimensional lattice-work (striated pitch and rhythm). Other dimensions (e.g. timbre) and continuous, smooth data are not so easily represented.' (p. 85)

He was also aware that the discrete nature of the data conveyed by a score made it difficult to identify many of the continuous events that are common to computer music, such as the transformation of sound. Although alternative printed abstractions could somewhat improve on these shortcomings, Clarke felt that 'two-dimensional black and white drawings are rarely able to capture the subtle detail of the sound and can frequently trivialise the musical structure.' (p. 85). With the introduction of the model-induced parameter mapping sonification approach that is distinguished by the foundation for data transformation there is reason to suggest that these issues can be addressed.

There are three fundamental elements for constructing a dynamic score in Mhyusics Auditory Display, the kinematic rigid body, the dynamic rigid body and the simulation environment. While all these elements are capable of generating forces within the system, only dynamic rigid bodies will respond to them. Kinematic rigid bodies and the simulation environment are therefore considered as unidirectional force types, whereas dynamic bodies are bidirectional. The dynamics of the score are controlled by adjusting the numerical properties associated with each of these types. These types also serve as the basis for further abstractions which can determine the conditions of the system in a more specific manner. Examples of those that are currently implemented by this AD are discussed further ahead in the objects section. An algorithmic framework can now be thought of as an arrangement of kinematic and dynamic bodies within a single environment. This framework places constraints on the simulation system and encourages the dynamic objects to act in a unique manner. The reader of the score therefore depends on the intuitive behaviour of each dynamic object in order to ascertain how their physical attributes develop. With the addition of a versatile parameter mapping system the score is now in a position to classify subtle details of the sound as its characteristics can be informed by any of the aforementioned attributes. Furthermore, the introduction of the data streaming conditions yields an opportunity for the score to represent smooth or continuous data transitions.

While the idea of the score has been delineated in the context of this AD, the preparatory stage of its creation still raises some issues. Doornbusch (2002) considered the process of mapping in algorithmic composition as a discrete stage of practice that requires time and planning. His interviews revealed the compositional practice of a number of composers would generally involve experimentation with various configurations in order to explore and contemplate the musical implications they held. The mapping configurations in Mhyusics invite the same conventions. However, the notion of them taking place at a more considered pace is contravened when they are tied to constructional elements of the score that continuously act under the timings of a simulation. For example, when assembling a composition comprising a sequence of events that rely on gravity, the user is forced to create additional structures in order to hold the dynamic objects in place and prevent unintentional incidents prior to its completion. Such circumstances call for a mechanism that allows the user to take the simulation in and out of an active state, thereby guaranteeing a mode of operation where dynamic objects remain fixed. The consequences of their modifications can then wait until the user decides that the simulation should become active again.

Barrass and Vickers (2011) took an artistically conscious approach toward the creation of an AD and described design as an 'iterative practice-based discipline involving cycles of hypothesis testing and critical evaluation that aims for solutions to specific problems in context.' (p. 146). It is clear from their description that a pausing mechanism by itself does not satisfy the design process. One feasible solution in that regard would be to integrate it with the concept of a beacon, which was first introduced by Kramer (1994c) as 'a means by which one can identify particular states of a system.' (p. 202). In the context of Mhyusics Auditory Display a beacon can be generated by saving the state of the simulation to a file. This file would contain all the values that describe the structural parameters of the score at the precise point in time the save operation was requested by the user. Loading the file would then recall the same parameters while placing the simulation in a paused state. This provides the user with an opportunity to make modifications before choosing to simulate the results. Given that this process can be repeated any number of times until the desired outcome is achieved, it can be regarded as suitable for an iterative approach to design.

Much of the rationale for adopting this type of state based reference point system has been touched upon when deliberating a potential solution to the mapping problem. In addition to mapping configurations, the states of the system can now offer a deeper insight regarding the construction of a composition. The iterative process of recalling reference points serves to familiarise the user with the audiovisual behaviour of the score. During this process the user can make changes to the simulation data, along with the mapping configurations, which will lead to variations in the timing and timbre of the audio events. By comparing various iterations they can appreciate the influence of their decisions in a particular context. Furthermore, when these contexts are categorised and stored they create a compositional narrative that is presented in its original framework for others to analyse, interpret, modify and develop.

3.3.4 Summary

The incrementally structured method of mapping, metaphor and magic can be regarded as fundamental to each connection which characterises this design. Mapping the data to video created a model which reduced the complexity of the data by acting as a metaphor for the familiar visual behaviour of everyday objects. The extraordinary qualities of the computer meant that this behaviour could be manipulated, allowing the design of Mhyusics Auditory Display to interpret this metaphor in a novel manner. This was illustrated when establishing the connection between data and audio. Here, the same metaphor was regarded as an indicator for both the dynamics of the data and the emergence of gestalts. The addition of PMS introduced an element of magic to the metaphoric model-based approach. In this case, the dynamics of the data became individual parameters which could be expressed through the audio under specific conditions, predicated by visual gestalts, and no longer solely dependent on user interaction. Consequently, this offered an explicit level of sonic analysis for any combination of data parameters attributed to a model. It was also put forward that this approach presented a number of advantages over the strict ties between data and audio that are typical of the physical modelling approach. Metaphors grounded in our common knowledge of physical causalities, as demonstrated by the model, were also submitted as an intuitive method to strengthen the mapping between data and audio. Such mappings could then be stored and categorised to promote successful configurations and demonstrate the circumstances in which they were applied.

The connection between audio and video further characterised how the visual nature of the model could support the sonification process. It argued that the video game engine is an ideal platform for both data exploration and navigation as it can dynamically process the model within an audiovisual environment so that it is presented in a manner that imitates our perception of the real world. The mapping between video and data was submitted as an ideal method for exposing a control loop, determined either by the user or the computer, while acting as dynamic feedback for both the audio and the data. It was then proposed that the appearance of the model could be manipulated to serve as a metaphor for an important process or solo theme. Furthermore, it was put forward this appearance could manipulated to such an extent that the model could take on more abstract qualities which were subsequently suggested to better afford the expansive sonic palette furnished by the PMS approach. Despite these changes in appearance, the close ties between the data and video ensured that the model would continue to enact actions familiar to our everyday experience. This proved to be essential when reclassifying the model as an abstraction to denote particular logics associated with the inner processes of various algorithms. The combination of this intuitive visual logic with a flexible audio configuration led to the outlining of a framework for the creation of a musical score. Here, the computer provides a sense of magic by bending the rules of the simulation to benefit an iterative process for design and analysis.

It is clear that each of these connections have been forged upon cogent theories which have since been interpreted to inform the design of Mhyusics Auditory Display. The next section will examine the interactive element of the sonification process, adding further concepts to support and extend the theory of this design, which is then consolidated in the implementation of the necessary components.

3.4 Defining the components

3.4.1 User control

Interactive sonification calls for a real-time interactive control loop between a human and a computer in order to both manipulate data and sonically interpret any subsequent changes that are informed by our actions. In reality, any action performed on an object elicits an instantaneous reaction which is implicitly linked to sound by the object's physical attributes. With the advent of computational technology this link is now explicitly defined where some form of human interface device (HID) converts our actions into information that the computer can recognise and process before responding accordingly. As a consequence of this distinct separation between action and sound we have found ourselves embracing a great diversity of control structures for virtual instruments (Roads, 1996) which still continues to expand to this day (New Interfaces for Musical Expression, 2013). This separation has also presented the opportunity to process and map the information coming from these recently discovered input devices in a multitude of ways. Unsurprisingly, this has raised a number of methods for the general classification of music control (Levitin et al, 2002; Jensenius, 2007; Cadoz, 2009) ushering in a vast area of research and debate that is difficult to cover in any depth within the scope of this thesis. Instead, the next section looks to highlight a number of significant theories associated with the study of our everyday interaction which are particularly congruous to the design laid out here.

3.4.1.1 Everyday interaction and common knowledge

Hunt and Hermann (2011) considered that one of the fundamental guidelines for the design of interactive sonifications is to 'respect natural physical coherences and to be aware that interfaces that deviate from them may give decreased performance by not connecting the users so well with physically expected linkages.' (p. 295). This echoed a previous argument which emphasised that 'the more one studies the ways that humans interact with the everyday world, the more it becomes obvious how our current computing technology uses an unbalanced subset of possible interaction techniques.' (Hunt & Herman, 2004: p. 7). They hypothesised that, until this point, sonification had neglected some of the innate skills acquired from interactions with everyday objects resulting in a loss of quality along with the potential to miss out on important aspects of diagnosis and analysis. Their solution was to delineate and implement a model-based approach which strove to capitalise on our interactive knowledge of the unchanging laws of physics and regain a more familiar means of connecting our actions to sound.

When comparing their underlying principles to trends in HCI it becomes apparent that they share several implications with the philosophical stance of embodiment. In particular, they tend to align with Dourish's (2001) idea of embodied interaction or 'the creation, manipulation and sharing of meaning through engaged interaction with artefacts.' (p. 126). This approach to HCI asserts a coupling of human action and perception where our mental processing is inseparable from the body and knowledge is gained through our actions in the world around us. By considering that user interaction is a fundamentally embodied phenomenon, the designer can maintain that a creative environment, populated by objects, provides an excellent foundation for communicating, interpreting and sharing ideas. For example, creativity can arise from our subjective understanding, which is often influenced by other people. When engaging with an unfamiliar object the user may be aware of their own activity, rationalising the actions they perform while discovering novel methods of control. With further experience the user grows accustomed to these actions and the object becomes more familiar and embodied. Similarly, when another person perceives the same actions they begin to interpret and rationalise the manner in which it is being used. When combining this process with their own interactive experience they can share their observations before both users arrive at some level of consistency, or common sense, for object interactivity. The same idea can be applied when trying to comprehend a data set. In this case the designer can employ a model-based approach to sonification where the data set becomes embodied within an interactive object.

One of the underlying concepts of embodiment is that of embodied cognition which emphasises a connection between sensory-motor experience and the mind. One pertinent aspect of this philosophy is demonstrated by our automatic ability to determine possible actions that can be performed on an object, thus creating a link between perception and action. Gibson (1997) termed this the affordance of an object which he defined as 'a combination of physical properties of the environment that is uniquely suited to a given animal - to his nutritive systems or his action system or his locomotor system.' (p. 79). He believed that there is an inherently creative aspect in the way we understand these entities as they can have multiple affordances which are dependent on the individual, the culture, and the context in which they are located. In effect, we are continuously expanding our knowledge about the affordances of the objects from our everyday interaction while retaining those learned through experience. It is therefore reasonable to suggest that we not only share a common knowledge of affordances for familiar everyday objects but that these same objects also hold a latent potential for less expected use.

Another significant benefit of data embodiment is that the designer can rely on the user's enactive knowledge for data exploration, that is, a set of motor skills which they acquired through interaction with their environment. Bruner (1966) postulates that our direct physical interaction with the world is a key component of cognitive development. He proposes that, irrespective of age, when presented with an unfamiliar task our intellectual development begins with an enactive stage which includes touching, feeling and manipulating objects. As such, we can acquire new skills when engaging in any interactive task. With this in mind it is arguable that we have already acquired a basic formalism for control of primitive shapes due to various childhood activities involving the manipulation of simple physical objects, such as stacking blocks. Indeed, this simple scenario proves to be remarkable for the development of human intelligence (Hirsch, 1996) and an effective benchmark for the design and analysis of artificial intelligence, or the creation of a machine with common sense (Minsky et al, 2004). Furthermore, if we consider that we typically possess an inherent understanding for the manipulation of primitive shapes then it is feasible to assume that this set of skills is transferrable to sonorous derivations, or abstractions more analogous to musical instruments. This has been supported in the design of new electronic instruments by authors such as Garnett and Goudeseune (1999) who wrote that 'problems associated with physical retraining can be minimized by retaining a physical interface that is functionally very similar to the practiced instrument.' (pp. 268-269). Indeed, in a more general sense, Cadoz (2009) anticipated that interacting with any object as a musical instrument would create a scenario where we could rely on both our enactive knowledge and the object's affordance.

When simulating a physical environment it is reasonable to assume that many of the skills acquired from our everyday interactions are transferrable. Kontogeorgakopoulos and Cadoz (2007) implied the principle of object affordances in the virtual realm when stating that 'since physical models enable an intuitive representation of the action we perform with real objects we can imagine several physical gestures to manipulate and control our model: dumping, pulling, pushing, etc.' (p. 277). From an embodied cognition perspective, Diniz et al. (2010) reasoned that 'given its inherent multimodal nature, a virtual reality based framework presents itself as an appropriated setting for the investigation and development of interfaces between body and music in which the natural communication tools are covered through the immersion of the actors involved.' (p. 66). In this context the actions of the users can be communicated through the behaviour of an embodied avatar which portrays the experience of data interaction from a familiar first person point of view. Others have claimed that the gestural aspects of this type of avatar could rival those we normally associate with the performance of a musical instrument (Hamilton, 2008; Berthaut et al, 2011). Both of those studies demonstrated that game engines designed for a first person shooter genre are adaptable for more expressive purposes by providing prefabricated actions such as running, jumping, strafing and shooting for the improvisatory creation of sound. Moreover, virtual environments of this kind are becoming increasingly popular and are host to many gamers' pseudo-self (Dolphin, 2009a). This fosters the emergence of a new generation of users who have already embodied the various requisites for interacting and communicating within this realm.

3.4.1.2 Establishing a connection between gesture and sound

Under the conditions of this design a gesture is a means of controlling the transformation of data into sound. For this reason gesture and sound are linked through the embodiment of the data, or the sonification model, which is realised in the form of a virtual object. Jensenius (2007) advocated an embodied approach to the cognition of sound and music when reasoning that 'music-related movement is a key component in all musical cultures, and that our life-long experience of relationships between objects, actions and sounds is an essential part of our bodily and mental capacities.' (p. 5). His stance suggested that musical features of sounds, such as amplitude and tempo, can be understood through a variety of physical metaphors based on movement, the basis of which stems from our ecological knowledge. In that regard, the most fundamental connection between gesture and sound, referred to by Jensenius as an action-sound coupling, is that which occurs naturally and thus determined by the laws of physics. As a consequence these connections are experienced as part of our everyday interaction with objects helping them to become so ingrained that they form a body of dependable evidence, identified by the same author as an object-action-object system, for forging a link between gesture and sound. Furthermore, as our perception of these couplings is essentially multimodal, Jensenius reasons that this experience allows us to 'predict the sonic result of a sound-producing action we only see, or imagine the sound-producing action of a sound we only hear.' (p. 32).

Fels et al. (2002) recognised that discerning the connection between gesture and sound could be ascribed to the conventions of embodiment when writing that 'the transparency of a mapping for the player depends both on cognitive understanding and on physical proficiency.' (p. 111). Specifically, experience can be acquired through interaction with a device, either directly (physical proficiency), or from the observation of others (cognitive understanding). In this case the role of the device was depicted as any traditional musical instrument that predicates our cultural knowledge in addition to the anticipation of physical causalities. They argued that such instruments not only support the inference of action from sound, or vice versa, but also serve to exemplify an action-sound coupling as a carrier for musical expression. In this sense the performer could exploit the transparency of the link to express a range of emotions through sound. The audience could then infer these emotions from the performer's actions without necessarily being well versed in performing with the same instrument. Kojs (2009) took these same basic principles and applied them within the framework of a virtual environment to demonstrate their feasibility in communicating a more expansive form of musical expression. He proposed that simulated entities which embodied the rigor associated with an action-sound coupling could underpin the conception of action-based music or 'the artistic exploration of mechanical actions which are used to control all aspects of composition, including its conception, form, instrumentation and instrumental design, performance and score.' (p. 286). As such, it is implied that the discernment of these compositional elements is to be guided by a collective knowledge of actions and their consequences.

While action-sound couplings are based on an enduring link, Jensenius' (2007) concept of an action-sound relationship promoted a more flexible method of defining the same relationship. This ranged from practical designs that closely adhered to the rules of action-sound couplings, to more creative solutions that would intentionally violate them. For the latter approach he noted that acting on volition alone could result in an incongruous design that failed to engage with the user in any meaningful manner. For this reason he believes 'that designing better action-sound relationships that either simulate or exploit the potential of action-sound couplings is of crucial importance.' (p. 33). This assertion recognises that a balance between intuitive interaction and innovation can be achieved by grounding the connection between gesture and sound in common knowledge relating to physical causalities, where it is more likely to be understood by the general user. In that regard, his conventions for forging a more acceptable link coincides with those of the physically informed mapping approach described in this thesis. This is to be expected if we consider that a virtual object not only embodies the data but also affords the actions that can be actuated by the user. When these actions are enacted they introduce energy into the virtual environment which can be described solely in terms of data values. For example, when performing a throwing gesture the energy applied is reflected by the relevant data values of the object it is performed on, that is, its linear and angular velocities. A greater amount of gestural energy will result in higher values for at least one of these velocities which can then be intuitively communicated through sound using the physically informed mapping approach. In essence, this creates a direct link between cause and sonic effect, serving as the dynamic control loop which underpins the interactive sonification. However, the crucial difference here is that we acknowledge another means of contextualising the objects by appreciating an alternative frame of reference that applies user control to the gap between the data and sound. In this light, the representational models for each data set can now be thought of as an ontological metaphor (Lakoff & Johnson, 1980) in that an abstraction, such as an activity, emotion, or idea, is represented as something tangible.

3.4.1.3 Determining an appropriate human interface device

In an effort to maintain a somatically intimate connection between musician and virtual instrument Cadoz et al. (2003) proposed that interaction with a physically modelled simulation should rely on force-feedback gestural transducers. Their bespoke setup (Figure 6) maintained a bidirectional continuum of energy between gesture and sound in order to reproduce the tactile sensation associated with producing a particular sound. While it proved that a multi-sensory motor experience close to that of playing a traditional instrument is attainable, other authors have maintained reasons why this approach may not be feasible and why simpler alternatives can often hold several advantages. De Campo (2009) observed that implementing elaborate devices requires more development time due to their complexities and may result in limiting the transferability of a user's interaction skills should their design prove too impenetrable. For this reason he expressed that 'even when there is a strong reason to use a special interface device, including fallback variants with standard UI devices is recommended.' (p. 69). Howard and Rimell (2004) were also aware of the problem posed by such bespoke interfaces when identifying that they were relatively expensive and only operable when paired with specially coded drivers. Similarly, Fiebrink et al. (2007) maintained that while standalone music controllers can be useful, they often invite several caveats such as 'exacerbating the long set up/tear down time, complicating transportation, requiring expensive sensors or components and expertise in their construction and maintenance, and presenting steep learning curves to players.' (p. 164). Their remedial stance was to advocate the innate capabilities of well-established HIDs, such as the mouse and keyboard, which are readily available as part of the conventional computer setup. In this sense the keyboard and mouse have become an accepted method of control, and thus embodied, for a whole generation of users. In consequence, the average user can transfer their associated skills to these everyday devices to make sounds without too much thought or practice before conceivably realising more musical ambitions (Wang et al, 2009).

Figure 6: This bespoke force-feedback gesture transducer allows for bi-directional interaction with the computer

Fiebrink et al. (2007) implied that the computer could enhance a generally understood metaphor for control when stating that 'although input devices such as keyboards and trackpads are simple and not physically configurable, software flexibility presents unexplored possibilities for using these devices in new and musically interesting ways.' (pp. 164-165). One notable example demonstrated this potential flexibility by implementing a unique hardware and software combination (Figure 7) to convert our interactions with commonplace objects into messages that could be interpreted by the computer as mouse and keyboard commands (Silver et al, 2012). In turn, this provided general compatibility with existing software to which the user could connect a host of conductive items in order to rapidly prototype new tangible interface ideas. While this setup allowed for the mapping of actions to button pressing commands without the need for programming, more complex interpretations alienated many users by requiring a greater level of coding knowledge. However, those who possess the knowledge required for more involved implementation strategies have already demonstrated that tangible objects can be an effective mechanism for interacting with computational models of real-world phenomena with the intentions of generating sound (Fels et al, 2002). The architects of this design upheld that any form of interaction can be better understood by choosing a physical controller that closely matches the actions and characteristics of the virtual object it is intended to control. Unfortunately, this concept implies that the appropriate interface device is often dictated by the virtual object which gives rise to a couple of limitations in the context of this design. First, a physical counterpart is not always guaranteed to be attainable, especially when considering the sheer number of possible forms that a virtual object might take. Second, it is not necessarily possible for the counterpart to dynamically replicate all the actions of a simulated rigid body, particularly if they are abstract from the physical world. Because of these issues it is perhaps more appropriate to pursue an alternative method for interaction. In particular, this method should be able to accommodate the many forms and behaviours that are possible within the realms of a rigid body physics simulation while remaining familiar and embodied for a wide range of users.

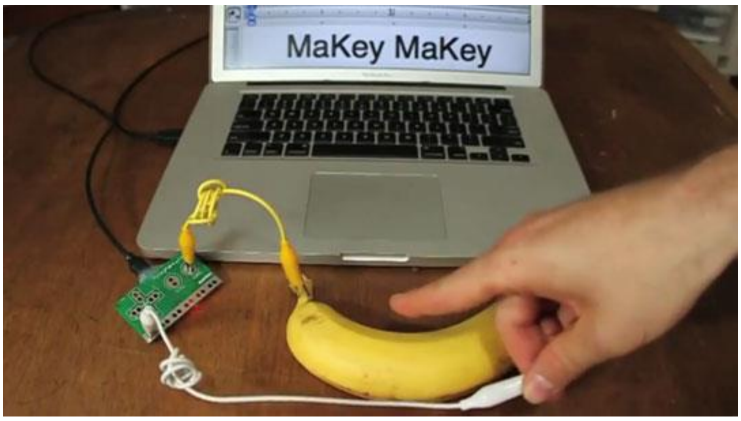


Figure 7: Makey Makey ensures that the user can employ everyday objects to send commands to a computer

Blackwell (2006) argued that one of the most dramatic user benefits in HCI arose from the development of direct manipulation (Shneiderman, 1983). Direct manipulation moved user interfaces closer to real world interaction by allowing users, through the use of a pointing device, to control visual abstractions of computational processes in real-time rather than typing a series of verbose instructions for the computer to execute. Since this remarkable step towards a more natural interaction process, a trend has emerged that encourages the characteristics of these abstractions to be enhanced with a greater degree of simulated physics, making them more in line with the mechanics of the real world. This was recognised by Jacob et al. (2008) who believed that many post direct manipulation interaction styles 'draw strength by building on users' pre-existing knowledge of the everyday, non-digital world.' (p. 201). These authors cited iOS (Apple Inc., 2013) as an example of a prevailing skeuomorphic interface (Figure 8) that will 'employ physical metaphors that add the illusion of gravity, mass, rigidity, springiness, and inertia to graphical widgets.' (p. 202). Accordingly, there are now an increasing number of users who are more accustomed to this manner of interaction and regard it as a convenient way of controlling data. For this reason the design of Mhyusics Auditory Display seeks to adopt many of the established direct manipulation techniques for two-dimensional abstractions and place them in the context of three dimensions. For instance, the user can perform a "drag 'n' drop" action with a rigid body in the same manner as they would with a two-dimensional icon, that is, utilising a pointing device to select and translate the body. However, while this presents a conventional interaction method that may be used for a wide range of conceivable objects, a typical pointing device, such as the mouse or touch screen, can only recognise movement along two axes. In the case of many virtual spaces this can be mitigated by rotating the camera to determine a new plane in which this movement can take place. Nonetheless, true three-dimensional operations will require a more appropriate device.

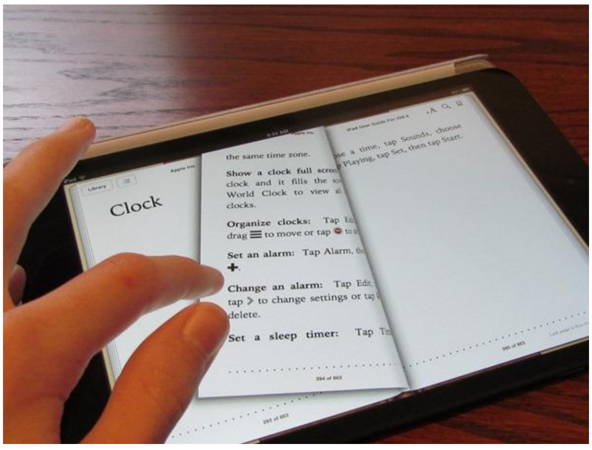


Figure 8: This skeuomorphic interface presents a manual in the form of a virtual book with pages that can be turned using a "drag 'n' drop" method of gestural interaction

The Wii Remote (Figure 9) colloquially known as the Wiimote, is a well-known, inexpensive and portable pointing device that can be conveniently connected to a computer via the Bluetooth specification (Haartsen, 2000). It shares a number of components that are common to game controllers such as a directional pad, several buttons, and basic haptic technology in the form of a rumble device. The original model featured a number of notable additions including an infra-red camera and a three-axis accelerometer, with the subsequent motion plus iteration adding a gyroscope that can read angular velocities along the x, y and z axes. These integral components render it suitable for the direct manipulation of objects in three-dimensional user interfaces providing that translation and orientation commands are performed separately (Chadwick et al, 2010). This is feasible when considering a direct manipulation approach as user interfaces generally treat these commands as two independent functions. As such, the Wii Remote can be regarded as a natural fit for this particular approach to interaction.

Figure 9: The Wii Remote along with its main interactive features

With the appropriate use of its additional components the Wii Remote can realise a greatly enriched experience in comparison to more modest pointing devices, such as the mouse. For instance, when examining the move based "drag 'n' drop" command it becomes apparent how the Wii Remote can procure a more intuitive metaphor for control. Quality of tactile interaction is arguably improved due to the rudimentary grabbing sensation that accompanies the initial selection command when enacted by simultaneously pressing the opposing A and B buttons. This quality could then be improved further with an additional degree of haptic response performed by the rumble device. The actions performed during the translation, or drag, stage are more evident to both the user and audience as the object path can correspond directly to the path of the device, providing a clear indication of how the two movements are related. The same idea applies to the release, or drop, stage of the command where the subsequent path and velocity of the object are directly linked to the gestural forces that govern the values of the accelerometer. Moreover, these values are susceptible to some level of variability that can add a sense of expression to the corresponding gesture. This was remarked upon by Kiefer et al. (2008) when they wrote that 'the physical nature of accelerometer control lends a natural randomness to its output, and the addition of acceleration motions such as 'flicking' the controller add an extra and slightly unpredictable dimension to the possibilities.' (p. 95).

3.4.2 Computer control

Keislar (2009) understood that continually evolving technology, such as the computer, has redefined the role of the performer. In particular, 'the process of disjoining the control interface from the sound generator means that human musicianship need not be so focused on developing physical dexterity and instead focus on higher level control with flexible mapping of gesture to sonic result - the instrumentalist tends to become a conductor.' (p. 39). Likewise, Jordà et al. (2007) observed that:

> 'in traditional instrumental playing, every nuance, every small control variation or modulation (e.g. a vibrato or a tremolo) has to be addressed physically by the performer. In digital instruments nevertheless, the performer no longer needs to control directly all these aspects of the production of sound, being able instead to direct and supervise the computer processes which control these details.' (p. 140)

The idea of instructing a computer to control details traditionally reserved for human interaction has also been realised within a framework concerning simulated objects. Cadoz (2002) orchestrated specific object parameters in order to simulate a broad range of human gestural vocabulary which, in turn, could forge musical constructions. He submitted that:

> 'position, movement, velocity, inertia, elasticity, viscosity and non-linearity are states and properties that can be applied to a wide variety of entities (even to certain life-like features or beings, not only in metaphoric senses). Thus we can use these concepts to create virtual gestures presenting quite rich properties.' (Cadoz, 2002: p. 309)

This notion was recalled as part a conceptual framework (Cadoz, 2009) for interaction with constructions derived from a mass-spring system (Figure 10). While both instances advocated the potential of a virtual instrumentalist, it was also disclosed that 'there is no question of replacing real instrumentalists with simplistic and naïve simulations.' (p. 228). Instead, this approach served to make the creative process more transparent where modelled actions are 'a tool to represent, understand, treat, communicate and teach the gesture.' (p. 228). In this sense, they not only grant insight into a myriad of potential techniques for playing virtual instruments but also provide the grounds for further inspiration. Kojs (2009) followed these sentiments when formalising a typology of cyberspace actions that borrowed from their physical counterparts. He reasoned that 'establishing, sustaining and developing actions stimulates creation of the musical content' (p. 288) where 'the syntax of action-based music governs the formulation and order of all-encompassing performance techniques.' (p. 288). Similarly, Hunt et al. (1998) believed that scoring a sequence of actions was well suited to an audiovisual environment and equivalent to 'a set of notions and concepts existing within the mind of a composer for the structure of a composition.' (p. 200).

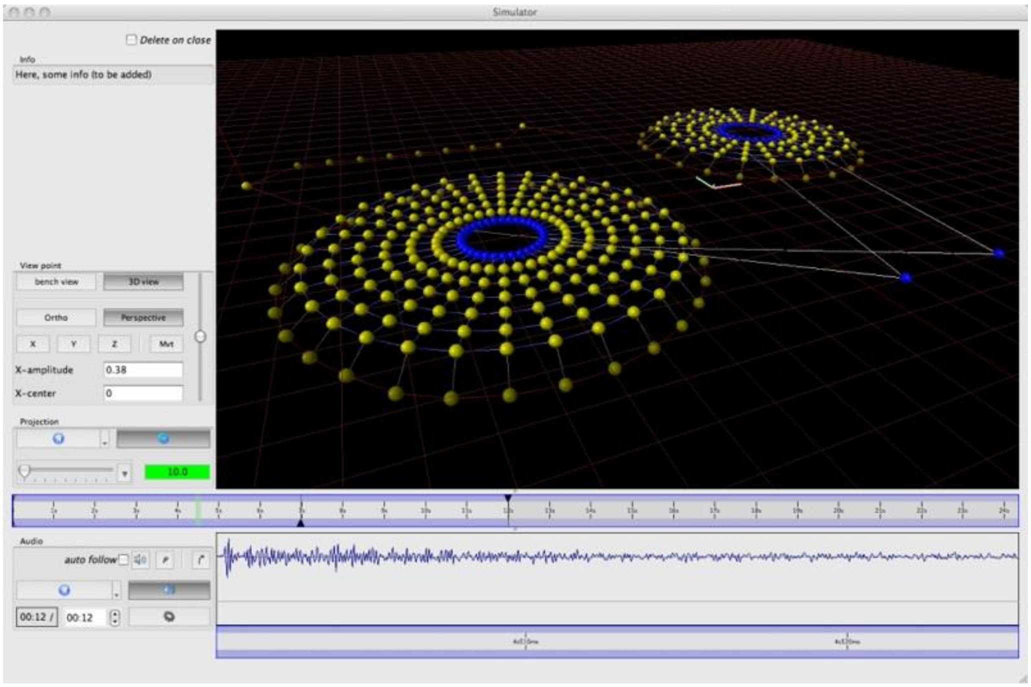


Figure 10: The simulation window used by GENESIS to describe a mass-spring system

While this model for interaction clearly supports musical creativity, it can also prove effective for analysing data. Hermann (2008) accepts that 'repeatability is essential for a technique to be scientifically valid and useful - otherwise nobody could check the results obtained by using sonification as an instrument to gain insight.' (p. 4). This can prove unattainable when circumstances dictate that only gestures enacted by the user can be relied upon. Fels et al. (2002) note that 'since humans do not reproduce their gestures very precisely, natural gesture recognition is rarely sufficiently accurate.' (p. 116). Indeed, in reality it is impossible to consistently perform a particular gesture in order to guarantee an exact object response, especially when its parameters are under strict scientific scrutiny. We can instead rely on the computer to trivialise such a process but it must first interpret a number of explicit instructions in order to perform the task at hand. In that respect, it is feasible to create a set of coherent actions, or commands, that form the basis of a rudimentary scripting language.

The introduction of a scripting language ensures that the user has access to a fine level of control over the parameters associated with each object. The current implementation of the language incorporates a number of basic commands (Table 4) that describe a particular state of an object at any given moment in time. By combining these commands the user encodes a unique gesture that will introduce energy into the system. As such, scripted events can be used to maintain a tight scheduling of data transitions and thus a controlled evolution of sound. Any number of actions are permitted to take place concurrently providing they are not of the same command type. Should this situation arise, successive commands will annul the actions of any previous commands based on a top to bottom list execution order. When compared to the physics network scripting language (Castagne et al, 2009) this approach opts to encode the state of a given model rather program the modelling activity itself, that is, using the language to automate the creation of structural entities. It is likely that such a method would prove too verbose when both a description for a specific object type and corresponding mapping configuration are required. Instead, an individual script is intended to be a fundamental component of each object and thus integral to the AD.

Command (case-insensitive)	**Description**
AngularVelocity(*x, y, z*)	Set the angular velocity of the object
GravFieldStrength(*strength*)	Set the strength of the gravitational field (a negative value will repel objects)
Hide()	Hides the object when the simulation is active and ensures that it is unable to exert or respond to any forces
LinearVelocity(*x, y, z*)	Set the linear velocity of the object
Orientation(*x, y, z*)	Orientate the object based on the yaw, pitch and roll values
Position(*x, y, z*)	Relocate the object at the position specified
Show()	Ensures that the object active and visible (only applies to hidden objects)
Spawn(*spawnObjectID*)	Position the object at the location of the identified spawn object

Table 4: A selection of commands for determining object behaviour

Levitin et al. (2002) believed that one of the most challenging goals in the design of computer music controllers was to feature 'a climate of individual expressivity and musical nuances, rather than quantising and otherwise impoverishing these gestures with the rigidity and strict repeatability that the digital computer so easily produces.' (p. 185). So, while scripting can offer precise actions, it becomes clear that the designer should facilitate the means for these actions to take on more humanlike qualities. In the case of this design the user can use nested commands to modify parameters which are otherwise static (Table 5). Current implementations include randomising values within a given range to allow for a controlled level of uncertainty, and linear interpolation, which encourages a less rigid transition between two values.

Parameter Modifying Command (case-insensitive)	**Description**
Lerp(*x, y, time*)	Perform linear interpolation between two numbers over a period of time (ms)
Random(*min, max*)	Calculate a random number between a given range

Table 5: A selection of subcommands for modifying the parameters of a script command

An important point to make here is that the designer must find the desired proportion between explicit control and ease of use if both computer and user interaction are to coincide. For example, if an object is engaged in direct manipulation, one should endeavour to prioritise the user's actions over scripted commands. This prevents the user from feeling disconnected from the control loop when unrelated computer actions counteract any intrinsic response.

3.4.3 The graphical user interface

In terms of HCI, Blackwell (2006) considered that a user interface could be described as 'a representation created to help the user understand the abstract operation and capabilities of the computer.' (p. 494). In this context of this design, the graphical qualities imply that many of these capabilities were deduced when establishing the video channel's connection to both data and audio. However, we must be aware of the learning process that ensues when a user is presented with a new set of acquirable skills particular to the lexicon of a novel AD. This requires accommodating for a range of competency where the learning curve is a crucial formative element for the architecture of the interface.

Blackwell (2006) implied that any creative use of metaphor in interface design can alienate a particular user depending on their proficiency. He refers to an account of scientific users appreciating a complex level of interpretation which served to overwhelm those who were less proficient. The same notion can be applied to interfaces concerned with sonification and AD. From their analysis of trends in the field of sonification, Kramer et al. (1999) believed that the tools at the time where too complex and specific to be widely applicable. For that reason they urged them to be more approachable so that 'researchers who are neither composers nor audio engineers would be able to produce detailed and intelligible data-driven sounds, manipulate sonification designs, and evaluate human performance associated with these designs.' (p. 11). Since then, this issue has been largely addressed by the interactive sonification paradigm (Hunt and Hermann, 2011) which has its focus on the needs of the general user. From this line of thought the authors call for a more inherent metaphor for interaction with complex data sets, predicating much of their guidelines for design with analogies to musical instruments, which they judge to be exemplar interfaces. When focusing their attention on human learning capabilities, and the means to structure a suitable learning curve, the authors suggest that:

> 'designers of such interfaces perhaps should consider how to engage the user in practice and learning. This is possibly best achieved by creating sonifications which contain information on multiple levels: a coarse level gives useful information even when the interaction is not mastered well; whereas a more subtle information level may be accessed with growing interaction competence, which furthermore motivates the user to engage in the interaction and in learning.' (pp. 294-295)

Naturally, this has been indicated in musical instrument design where Jordà (2005) revealed a consensus amongst the literature to seek a balance between initial low frustration, or usability with little training, and enough diversity to offer long term interest and thus associated skill development. Menzies (2010) also felt the concept of layers could be applied to a more pertinent interface concerning audiovisual rigid bodies when stating that 'a layered structure is desirable in which more complex features are accessible, but this can be overlooked initially.' (p. 3).

While this gives some indication for a general approach to interface design, the levels of information, and thus the learning curve, still need to be defined within the framework of the AD designed as part of this thesis. If we refer back to the sonification methodology, namely model-induced parameter mapping, we can distinguish the basic levels of information that are required to be displayed by the GUI. Hunt and Hermann (2011) discerned two types of interactions regarding PMS. The first, described as interactive data selection, involves 'controlling what subset of the data set under exploration is to be sonified.' (p. 290). The second, termed mapping interactions, concerns the adjustment of the mappings and their related parameters. For the former they advocated that 'a very intuitive and direct form of interacting with the data is to provide a visual user interface where the dataset can be visually inspected.' (p. 290). This idea correlates with the model-induced aspect because the data selection is manifest through polygonal models which, in turn, require a virtual Euclidean space in order to be fully realised. The GUI (Figure 11) is therefore required to incorporate this space so it may be populated by these interactive representations of the data.

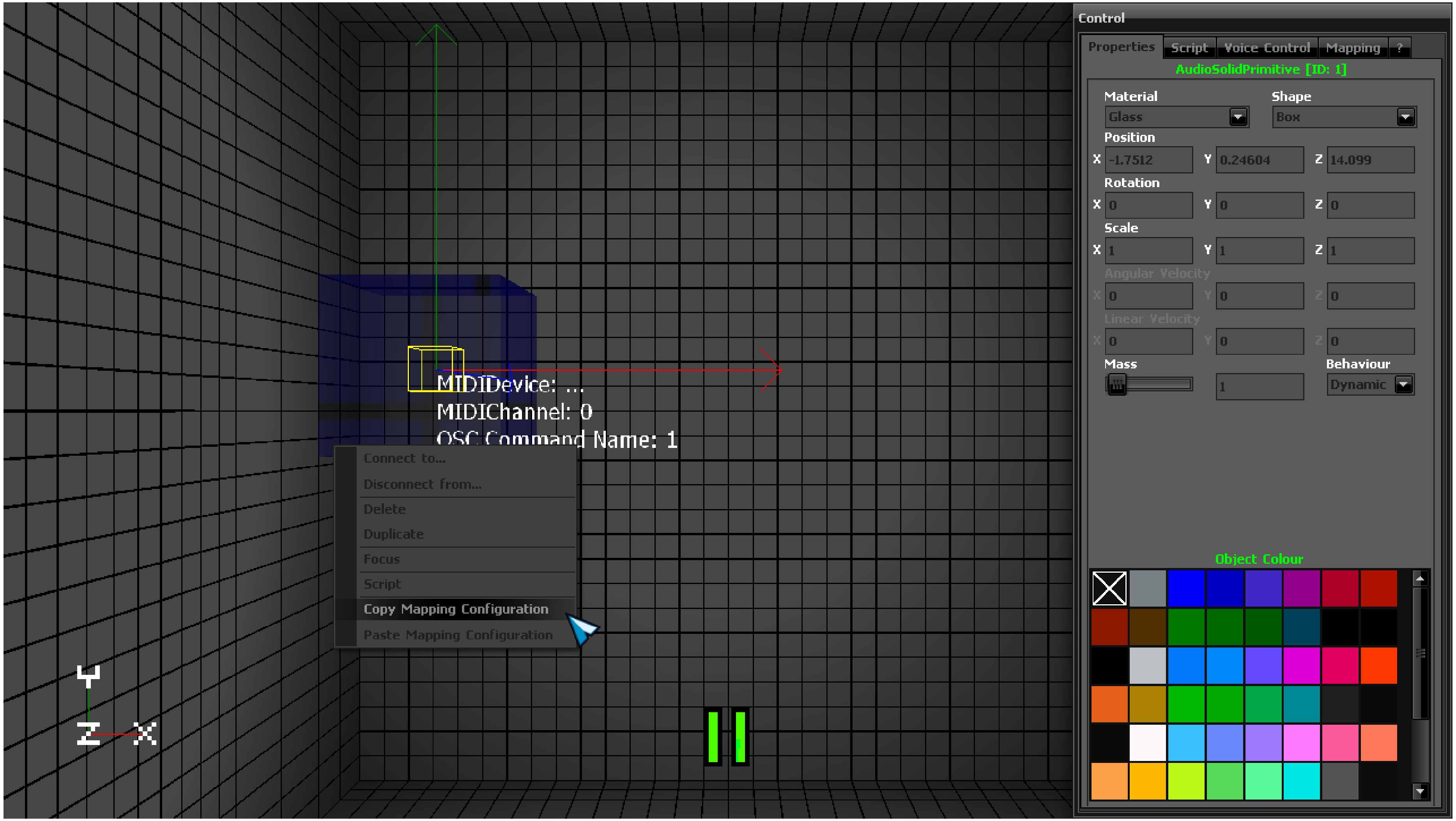


Figure 11: The GUI for Mhyusics Auditory Display

As the characteristics of each model are attributed to the dynamics of rigid body physics engine they act under a set of prevailing rules maintained by a simulation space, or environment. Under typical circumstances this space comprises three dimensions and is of a known volume. As such, it functions as an overview of the data set while representing an encompassing metaphor for interaction. For example, the rules of the environment not only govern the actions of each model, they also provide conditions for the dissipation of gestural energy, where the adjustment of forces, such as damping, can augment or diminish its effect. As part of a foundation for interaction this equates to the least complex level of information display and is realised by the GUI's environment window. It invites the direct manipulation of data without the need for coding keywords, operators and the understanding of basic programming paradigms. In this manner, data characteristics are portrayed through object behaviour and sonification involves loading preconfigured environments with corresponding synthesis patches. Consequently, the user can readily explore an arrangement of data models which they can then rearrange in a manner of their choosing to procure a unique flow of information, and evolution of sound, over the simulation period.

Although this metaphor can be seen as an inherent starting point, it serves to restrict those who wish to promote data interpretations beyond that which is offered by a fixed schema for objects and sonic response. This is, in part, is substantiated by the mapping interactions category, where various mapping configurations can be configured by the user. However, Hunt and Hermann (2011) warned of a major issue introduced by this technique accepting that 'it is quite difficult to program a good interface with which to influence the mapping' (p. 290). They put forward that, in most cases, it 'often results in a demand for too much knowledge about the software system (e.g., Pure Data or SuperCollider) for the user to change the mapping in an intuitive way.' (p. 290). While this is undeniable, the GUI contains a number of tools, such as the function editor and messaging system, which attempt to optimise the learning curve by conforming to agreed standards for the mapping process. These tools constitute part of a second window, known as the control window, which permits a more precise means of data control and analysis. Here, the GUI controls offer a discrete approach to the adjustment of parameters as part of a composite numerical analysis of each subset of data, or object, in the environment. Whereas this can be argued to require a greater degree of interaction competence, the effect of changing each parameter can be dynamically reflected by the environment window which works in tandem to present the same information at a more elementary level. This correspondence also serves to aid the user in comprehending the most advanced level of data interaction which is delivered by a bespoke scripting language. By engaging in this interaction method the user gains tight control over a number of strictly timed commands, allowing them to prioritise particular changes in parameters. Because the environment window presents a more transparent view of the actions that ensue, the user can quickly learn the effects of each command.

3.4.4 Objects

From the delineation of the connections comprising the design of Mhyusics Auditory Display we can understand that the objects present the user with a mental model for the manipulation of multiple modalities under one unifying abstract structure which, in turn, allows them to comprehend the underlying data. When considering control of high-dimensional spaces Garnett and Goudseune (1999) argue for a mental model simpler than brute-force awareness of every detail in order to avoid cognitive overload. One can make a case that this is achieved here when we consider that the high-dimensional data, encapsulated by each object, allows us to intrinsically understand how the data parameters vary together. For instance, throwing a rigid body would imply a change in velocity that is influenced by its mass. Likewise, a spherical body is naturally inclined to roll along a surface, providing smoother changes in angular velocity as opposed to the sudden, erratic changes of a square shaped equivalent. The idea of a single gesture controlling multiple parameters is also consistent with what we expect from a traditional musical instrument. Indeed, the mental model becomes more analogous to such a device with the addition of the mapping framework. Here, a body can stream its parameters to sound dimensions within a designated extent of time that is framed by a particular type of condition. At the beginning of this period the object is assigned a single voice whose length is controlled by an envelope. In this fashion the object now symbolises an individual event on a dynamic virtual score where emergent musical structures are determined by the simulation.

When placed within the framework rendered by the user interface, the objects can respond to multiple levels of control. This ranges from the direct manipulation approach found in the environment window to the more precise control afforded by the control window. In all cases the objects rely on a source of energy to provide the impetus for movement. This energy is governed by three factors which can be categorised as environmental, object interaction and gestural control. The environmental factor concerns global parameters which regulate the actions of all objects through forces, such as gravity and damping, and the volume in which all simulations can take place. Gestural control relates to either the user's direct manipulation of objects, such as grabbing and throwing with a mouse, or scripted actions that are enacted by the computer. Object interaction applies to every object in the simulation as they all play a role in guiding the transfer of information, or energy, within the system. This type of interaction comprises exerting and responding to forces that derive from collisions, vector fields or constraints. As these determinants are realised within the context of a computer, the possibilities for the control of this energy goes beyond that of the physical and cognitive human limitations we associate with materialistic equivalents. Hence, while an overwhelming number of configurations could be proposed, we can instead refer to the discretion of the design framework to compartmentalise a number of suitable control methods. Here, further derivations of our mental model can denote a specific set of actions while adhering to an inherent metaphor for comprehension.

The modular approach to the object design calls for the user to construct logic for directing the flow of energy using of a limited number of building blocks. Each of these objects has intelligible physical behaviour which serves to make both data analysis and compositional processes more accessible. The classification of the six basic modules (Figure 12) comprises one object that is capable of producing sound and five objects that are designed to influence its actions. Whereas the single sonorous type has comparatively simplistic behaviours pertaining to physical phenomenon, the subsequent types seek to build upon this basic metaphor with further abstractions. Together, these objects form a common language for establishing a variety of data structures, instruments, scores, and musical algorithms of varying complexity. In this manner, the user is invited to make use of the synergy of physics to achieve something greater than the individual components where their assembly is guided by freedom of choice.

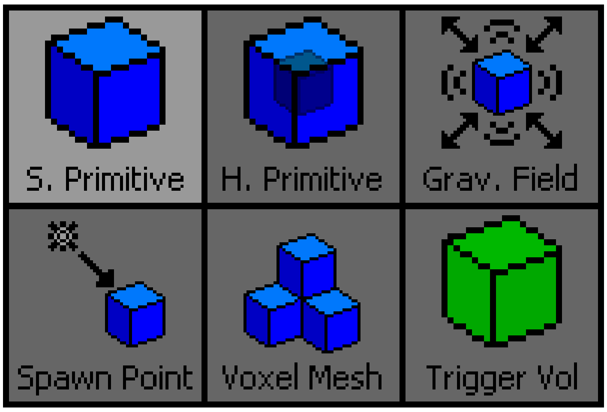


Figure 12: Modular object icons (a lighter background colour indicates a potentially sonorous type)

3.4.4.1 Solid primitive

A solid entity that denotes a rigid body in the form of a box, capsule, cone, cylinder or sphere. This object exposes a number of parameters, based on its apparent characteristics, which can be mapped to any number of audio dimensions.

3.4.4.2 Hollow primitive

A hollow shape that can assume the same forms as the solid primitive. These forms allow other objects to be contained inside their volume while preventing those outside from entering. Simulation stability is a necessary consideration with this object type as heavily nested hollow objects tend to be unstable under a number of conditions. This includes extreme angular and linear velocities brought about through vigorous gestures and interactions between two objects that have a vast difference in mass. It should also be noted that some physics engines may suffer from lack of stability due to a loss in numerical precision when objects are located at a great distance from the world origin, that is, an x, y and z value of zero in Euclidian terms.

3.4.4.3 Gravitational field

A solid shape which, again, can assume the same forms as the solid primitive. This shape represents a gravitational vector field which is capable of pulling objects towards, or pushing them away from, its central location. Any forces exerted by this field will be applied in conjunction with the environmental gravitational force parameter. It is also possible to constrain this field to an arbitrary volume so that dynamic objects are only affected should they interact with a defined area of influence. In reality it is not as feasible to limit a gravitational field to such a precise area, yet in the virtual world this is relatively straightforward.

Jordà (2005) submitted that predictability is important in instrument design because the performer needs to understand the consequences of their actions under various circumstances in order to gain a sense of confidence. However, he also felt that this should be juxtaposed with a sense of non-linearity, or more random behaviour, where a user-controlled balance between randomness and determinism should be sought. When placing this argument into the context of a musical performance he believed that:

> 'non-linearity should not mean uncontrollability nor unpredictability! In a good performance, the performer needs to know and trust the instrument and be able to push it to the extremes, bringing it back and forth from the zones of non-linearity.' (p. 216)

Through the audiovisual display of an arguably deterministic field of physics (Hutchison, 1993) we can comprehend classical mechanics algorithms that lend themselves to stability and predictability. Indeed, both the solid and hollow primitives will adhere to these principles under typical circumstances. However, with the addition of the gravitational field object their behaviour can become more erratic. This was indicated by Sturm (2001a) when he evaluated the potential dynamics of a particle simulation and wrote that 'any number of mysterious forces, can be applied to a system, creating drag forces and keeping the system under, or out of, control.' (p. 134). With a gravitational field object we can introduce these forces to a defined area, which can be equated to Jordà's zone of non-linearity. If we consider a scenario where a number of dynamic solid primitives are contained within a larger kinematic hollow primitive it can be demonstrated that, under typical circumstances, the dynamic variants will eventually come to rest. With the direct control of a gravitational field object the user can make numerous gestures to bring the dynamic objects in and out of its vector field. The objects respond by becoming more animated, leading to interactions which lend themselves to more chaotic behaviour that is unique to each gesture [Demonstration 1].

3.4.4.4 Spawn point

A spawn point is a means of visualising the position at which an object may be generated at a specified time during the simulation. It behaves in the same manner as a solid cube-shaped rigid body but its appearance is that of a two-dimensional sprite. When acting dynamically it becomes a convenient method for assigning an active point for object creation based on the forces within the system. With kinematic behaviour it becomes a fixed point for the same creative process. In both cases the user can employ direct interaction to grab and guide the spawn point through the environment.

Keislar (2009) postulated that one could assess the success of a technique when gauging its results against a familiar reference point. He refers to examples such as physical modelling synthesis generating a sonic complexity akin to traditional musical instruments and the Turing test (Turing, 1950), where a computer is programmed to provide a believable imitation of a human being. From his argument one can put forward that Mhyusics Auditory Display uses a deterministic simulation of rigid body dynamics as its familiar point of reference. This makes it viable for audiovisual assessment of procedures, providing certain conditions remain constant. Hunt and Hermann (2011) suggested that 'it is good practice to render short, looped sonifications, so that the effect of mapping and parameter changes become clear with the next few seconds at most.' (pp. 290-291). Using a kinematic spawn point we have a fixed location at which to spawn an object with a specified configuration. Introducing scripting to this scenario permits the user to define regular intervals at which this relocation event takes place. Providing both the environmental parameters remain constant, and any obstacles it encounters remain unresponsive, the dynamic object will follow the same predetermined path each time it appears. This creates a looping scenario where the visuals and underlying data remain constant. The user is now free to actively adjust the mapping configuration at any point during the simulation which will then cause the AD to immediately reflect the effects of their changes [Demonstration 2].

3.4.4.5 Voxel mesh

A voxel is a volumetric pixel which assumes the form of a box, typically with uniform dimensions. Any number of voxels can be combined to form a hollow mesh which can then contain other objects. This provides a visually defined constant for quantifying virtual Euclidean space, making it easier for the viewer to assess relative distance under various environmental conditions. A small selection of these meshes are supplied with Mhyusics along with a bespoke voxel mesh editor for creating and exporting further instances [Appendix 1].

It can be argued that one can procure tighter rhythmic scenarios by using this uniform approach to determine object movement. If we align a number of kinematic meshes of various lengths and place a single dynamic object inside each one we can establish the maximum relative distance for which these objects can travel. By creating corresponding gravitational field objects, and restricting their field volumes to envelop a particular mesh, the system grants individual control over each dynamic causing it to travel towards the boundaries of its container, whereupon a collision event is generated. With the ability to script this parameter the user can command the timing of each collision, and subsequently, the rhythm of the corresponding sound [Demonstration 3].

3.4.4.6 Trigger volume

The trigger volume is another object that derives from a primitive shape but unlike other primitives it allows for intersection. When intersecting with another object the trigger volume sends out a message that activates affiliated scripts. Such an event can be used to initiate a number of computer controlled actions which manipulate the flow of energy in the system. This can range from small changes in an object's behaviour to a dramatic shift in the state of the simulation.

As previously demonstrated the spawn point object allows for the realisation of repeatable loops. With the addition of the trigger volume to this scenario it is possible to introduce a controlled degree of randomness within this definite framework. For example, the user could create a script for each kinematic obstacle that randomises their position when triggered. When the dynamic object reaches the end of its path it will come into contact with the volume and the loop will begin again. This will result in a virtual space that is populated with the same objects as the looping demonstration but, due to their scripted characteristics, the dynamic object will act in a unique manner each time it is spawned [Demonstration 4].

It can be suggested that such a scenario has similar connotations to that of a manifold composition (Kaper & Tipei, 1998) where levels of indeterminacy are added to an otherwise deterministic framework. Kaper et al. (1999) stated that as 'a nontraditional way of composing, the manifolds show how high-performance computing provides the composer with new means to try out compositional strategies or materials and hear the results in a reasonable amount of time.' (p. 49). Indeed, with arrangements involving this object type the composer has more immediate access to a potentially unlimited number of variants on a compositional idea. Likewise, the sonification process can profit from this situation as the data itself undergoes randomisation while the mapping arrangement remains constant. This permits the listener to test a controlled range of conditions in which the same mapping may be applied and assessed.

3.4.5 Function editor

The mapping, or transfer, function follows the formalism set out by Hermann (2002) where the data values are located along the domain and the sound attribute values are along the range (Figure 13).

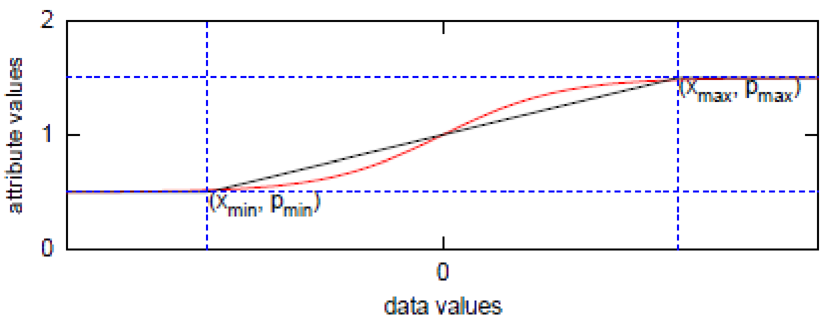


Figure 13: A typical transfer function as denoted by Hermann

The function is presented inside an editor (Figure 14) which is designed to accommodate the non-linearity between these values and our auditory perception by allowing the user to define various rules for range, polarity and scaling. This is accomplished through a breakpoint system where there are two mandatory points to denote the corresponding limits of the data domain. These points respond to "drag 'n' drop" gestures but are constrained to movements along the range only. Further breakpoints can be created and deleted between these permanent points allowing the data to audio relationship to take on many forms including linear, exponential, sigmoid and step functions. Other controllable factors include snap to grid, interpolation between breakpoints, and real-time curve following, which displays the value of the curve in relation to the position of the cursor. The user can also zoom in and out of the function, adjusting the resolution of the data values to best reflect what numerical detail they require.

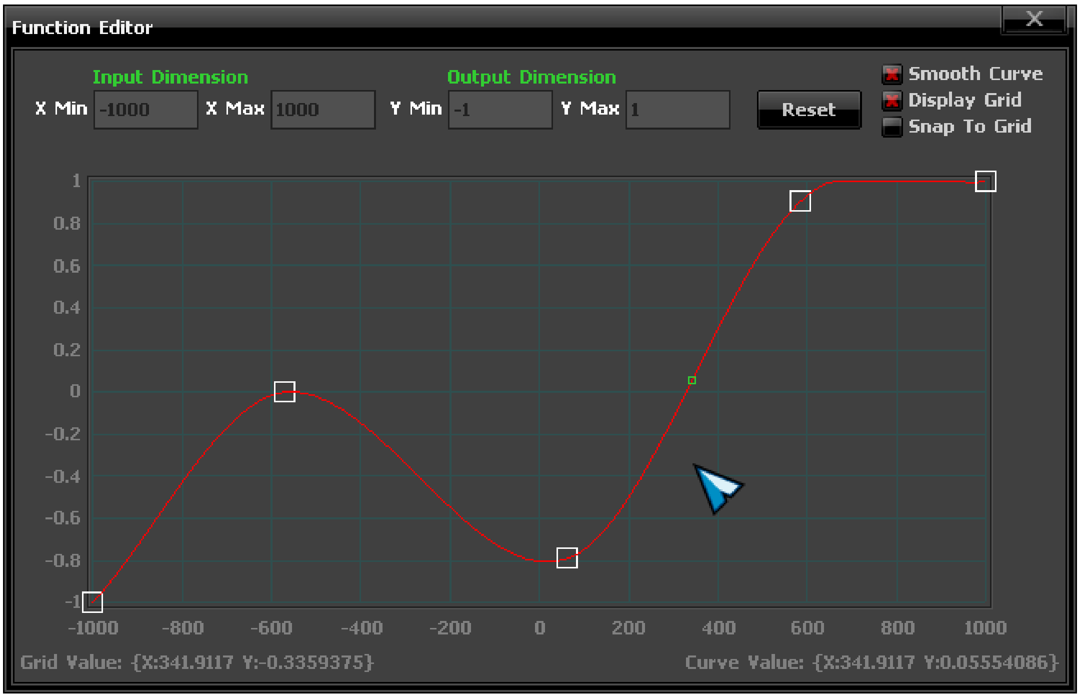


Figure 14: The function editor in Mhyusics Auditory Display

In addition to these controls, the data transformation tools allow for the saving and loading of both the mapping configurations and functions. This provides the means to assess a particular function in combination with a large variety of mapping arrangements. When loading a function it will retain its original shape by automatically scaling to the range of the input and output dimensions which are, in turn, dictated by the current mapping configuration. With this system in place it becomes straightforward for the user to apply a series of functions to a particular data to sound arrangement in order to analyse and evaluate the successive changes in audio.

3.4.6 Messaging system

Due to the great diversity of control structures for musical instruments it would make sense to utilise protocols that have attempted to standardise communication, or bridge the disconnection, between control and sound generating sources. The separation between controller and generator was first standardised by the MIDI protocol (Loy, 1985), bringing computer music into the mainstream interactive domain, but has since been criticised for its limitations (Moore, 1988). Wright and Freed (1997) established the OSC protocol to address many of these problems. Likewise, their approach has a number of shortcomings (Fraietta, 2008) and is not integrated to the same extent in both commercial hardware and software synthesisers, with only a few notable exceptions such as Reaktor (Native Instruments GmbH, 2013). Consequently, those who wish to benefit from OSC will either need to depend upon non-commercial examples, which often come with limited instructions and support, or design their own compatible synthesis algorithms.

```
private void OscMessageReceivedHandler(object sender, OscMessageReceivedEventArgs args)
{
    OscMessage message = args.Message;

    // The following line is not required if 'FilterRegisteredMethods' is enabled
    // The filter address can be set in the 'OSCSettings' class (Default: "/oscReceive")
    if (message.Address != IncomingOSCAddress) return;

    // Can also handle bundles if required
    //if (message.IsBundle)

    string address = message.Address;
    int length = message.Data.Count;

    string messageAddress = message.Data[0].ToString();

    // Examples of functions controlled by messages based on the OSC protocol

    //=============================================
    //---[Camera Control]--------------------------
    //=============================================
    var c = game.ActiveCamera as Camera_FreeLook;

    if (messageAddress == "/camera_TranslateForward")
    {
        // Try parsing the input value as a floating point value
        float value;
        bool isFloat = float.TryParse(message.Data[1].ToString(), out value);

        // Only a floating point value will allow the function to be called
        if (isFloat)
        {
            // Translate the camera forward by the value associated with the message
            game.ActiveCamera.Move(game.ActiveCamera.Forward * value);
        }
    }
    else if (messageAddress == "/camera_TranslateBackward")
    {
        float value;
        bool isFloat = float.TryParse(message.Data[1].ToString(), out value);

        if (isFloat)
        {
            game.ActiveCamera.Move(-game.ActiveCamera.Forward * value);
        }
    }
    // etc...
```

Figure 15: Handling input messages via the OSC protocol

The messaging system in this design utilises these two protocols to accommodate a standardised solution for the customisation of data flow both in and out of the software. Input messages are currently handled by OSC as it facilitates a more direct and flexible addressing of the compliant functions (Figure 15). Opting for an integrated messaging system to control the outgoing data ensures that it is straightforward to separate the audio synthesis from the AD. Implementation of these protocols grants access to network communication where the data can be passed on to computers running on separate operating systems and mobile platforms. In turn, this provides the opportunity to communicate with a vast range of audio applications, and synthesis tools, which offer a greater level of sonic diversity when compared to an inbuilt synthesis engine. Dolphin (2009a) drew attention to this fact when stating that 'the notable restriction of some game engine software systems is their in-built sound generation capabilities, with these often limited to basic sample playback triggering with pitch adjustment control.' (p. 214). He recommended that 'the use of a network connection to communicate between a game engine and an external sound and synthesis software application bypasses this issue, allowing any OSC capable external application to be integrated.' (p. 214). In return, this provided 'significant creative possibilities for a sound artist who may then use familiar tools for the development of sound and synthesis systems.' (p. 214).

It could be argued that using the MIDI protocol is more likely to provide a link to more commonly used, and therefore widely understood, solutions. However, it has been my experience that the MIDI specification permits a limited range of freely assignable controller commands for mapping purposes. In addition, many of the standard commands are not universally recognised across MIDI enabled devices, often eliciting no response when performed. Another prominent difficulty from a sonification point of view is that MIDI does not easily allow for both a continuous and precise determination of pitch in the same manner as OSC. A compromise can be attained with a dynamic pitch algorithm (Figure 16) but this method is highly dependent on the user knowing the exact pitch bend range of their MIDI-enabled instrument which, again, is not a universal standard. Alternatively, some plug-ins expose coarse and fine tuning parameters which can be used to similar effect and are often assignable via a MIDI learn function.

The routing of parameters does not necessarily need to be restricted to virtual or physical instruments. With the appropriate messaging setup objects can become a metaphor for a virtual mixer with control over various effects, or a sampler with the means to trigger and manipulate sound files. This would necessitate mapping the data to intermediary functions as part of a more complex routing process in a similar manner to the aforementioned handling of the input data. With this in mind the system also supports streaming data values directly to any OSC compatible software, bypassing the inbuilt mapping function, so the original values can be processed through custom mapping functions or routing algorithms.

```
//============================================
//---[MIDI Note & Pitch Bend Method]-----------
//============================================
// • Convert frequency value to MIDI note (Hz to MIDI)
// • Find the upper interger value (ceiling)
// • Pitch bend the remaining difference (with pitch bend mantissa value)
float midiNote = finalValue == 0 ? 0 : MIDIDevice.MIDINoteFromFrequency(finalValue);
float roundedMIDINote = (float)Math.Round(midiNote, MidpointRounding.AwayFromZero);

// Update MIDI note for correct note on retrigger (the pitch is bent from this anchor note)
parent.AudioObject.MIDINote = (int)roundedMIDINote;

// [Pitch Bend Information]
// • Maximum of 14-bit pitch bend range = 16384 = 2^14
// • Range is above and below MIDI note therefore range is [-8192, 8191]
// • Pitch bend range in MIDI notes is NOT standard across all devices

// Get the current pitch bend range for the device (this can't be queried so default value is 4)
float pitchBendRange = game.MIDIDevices[parent.AudioObject.MIDIDevice].PitchBendRangeSemitones;
pitchBendRange *= 0.5f;

// 2D linear interpolation (method returns interpolated x2 value)
float pitchBendAdjustment = ToolBox.Lerp2D(
    0f,                                     // x1
    16383,                                  // x3
    -pitchBendRange,                        // y1
    midiNote - roundedMIDINote,             // y2
    pitchBendRange                          // y3
    );

int LSB;    // Least significant byte
int MSB;    // Most significant byte

// Uncombine a 14-bit value to get two 7-bit values
MIDIDevice.Uncombine14Bit((int)Math.Ceiling(pitchBendAdjustment), out MSB, out LSB);

// Generate a new MIDI message with the 7-bit values
ChannelMessageBuilder cmb = new ChannelMessageBuilder();
cmb.Command = ChannelCommand.PitchWheel;
cmb.Data1 = LSB;
cmb.Data2 = MSB;
cmb.MidiChannel = parent.AudioObject.MIDIChannel;
cmb.Build();

// Send message to MIDI device
game.MIDIDevices[parent.AudioObject.MIDIDevice].Send(cmb);
```

Figure 16: MIDI dynamic pitch algorithm

While the amount of data generated by each virtual instrument is extensive, the GUI controls permit the system to select and filter any number of data dimensions. The user is therefore able to concentrate on any particular subset of data streams for the purposes of analysis or musical effect.

Furthermore, the GUI grants quick access to, and routing of, the input and output dimensions, providing the listener with the means to assess contrasting aural perspectives on the same data. However, it should be noted that the combinations of data to sound parameters are restricted to one-to-one and one-to-many mapping configurations. There is no allowance for a many-to-one setup because it requires some method of interpolation for the multiple input values in order to produce the final output value. By interpolating several data streams we introduce a less direct relationship that is not as metaphorically apparent due to the confusion over the true target domain which, in turn, undermines the bridging of the gap between data and sound. In contrast, the one-to-one and one-to-many configurations do not require interpolation of either the source or target values and are easily defined by a single mapping function.

3.4.7 Camera

When discussing data navigation design Brazil and Fernström (2011) concluded that 'the provision of a general overview or gist, a navigable data space, dynamic filtering and details on demand about items or groups has been shown to be successful.' (p. 521). While the overview and details are provided by the AD's environment and the control windows respectively, the camera ensures that the environment is navigable while proffering a further mechanism for the dynamic filtering of information. As a visual tool it can be employed to locate specific regions of data by framing a group of objects. From an aural perspective this can involve focusing the viewers' aural attention on a particular subset of objects in a similar manner to the "cocktail party effect" (Cherry, 1953). However, facilitating this idea requires the introduction of camera relative parameter mapping and in this instance, a dynamic range of aural perception can be realised by attenuating a sound source based on its relative distance from the camera.

A camera relative mapping approach also proves advantageous should we consider that any spatially dependent properties, such as position and velocity, are ordinarily relative to the environment. This was understood by Sturm (2001b) when he deliberated the correspondence between the virtual space of simulated particles to the aural space of the sonification. He believed that 'any sonification of a particle system is dependent on the state of the observer; each observer with a unique position and/or velocity will hear the system in a different way.' (para. 12). Under these circumstances we can deduce that a change in perspective will contravene any former metaphorical links to sound. To this effect, Johnson (1987) claimed that metaphors are unconsciously formed from experiential gestalts, or embodied schemata, that derive from our body's orientation in space and its interaction with objects. Correspondingly, Lakoff and Johnson (1980) demonstrated how the familiar upright position of our body in space underpins a number of spatially relative metaphors. This includes a general belief that up is more because we can relate it to our interactions with the physical environment, for example stacking more objects on a pile to increase it vertically. As such, the user can choose to employ camera relative parameter mapping in order to help maintain these intuitive relationships.

Saue (2000) was among the first to propose a first person point of view for the navigation of an AD. He described a foundation of ideas for exploring large data sets via the metaphor of a virtual space which he believed to accommodate a number of essential features including an 'egocentric point-of-view, immersion, user-centered interaction, multi-sensory feedback and three-dimensional perspective.' (p. 106). Indeed, many FPS games comprise these same elements in an attempt to provide immersive audiovisual simulations of reality and have been shown to be a good match for sonification systems (Grimshaw, 2007). From their design we can understand that the player perceives, navigates and interacts with the game world through the operation of a camera. The same basic principle also applies to Mhyusics Auditory Display where physical control of the camera follows that of a typical FPS setup, utilising the W, A, S and D keys for translation, and a mouse for rotation. This configuration not only invites the means to generate highly dynamic gestures (Hamilton, 2008; Berthaut et al, 2011), but also introduces a sense of bimanual control that is common to most traditional instruments. Cadoz (2009) submits the notion of an energy continuum arising from our interaction with traditional instruments where 'the energy of the sound is a transformation of the energy of the gesture.' (p. 218). He describes the two functions of gesture that influence this continuum as production and modification which he illustrates through the playing of a violin. In this scenario, the right hand produces the energy by either bowing or plucking the string while the left hand modifies the energy by adjusting the length of the vibrating string. When demonstrating bimanual camera control in conjunction with the available objects we can appreciate how this can be realised within the context of Mhyusics [Demonstration 5]. Here, the left hand controls the volume level by translating the camera towards and away from an object while the right hand adjusts the orientation of the camera to pan the audio based on the object's relative location.

When describing how physical camera navigation could be used to browse through a data set Saue (2000) argued that 'free mouse search is advantageous during the first exploratory phase, where orientation is important.' (p. 109). More readily, it enables the user to find areas of interest at their own leisure. However, it was stressed that this method proved strenuous, lacked precision and was less appropriate for tasks where 'the timing of a data-driven sound sequence is crucial, as in pattern recognition and comparisons.' (p. 109). As a means for a solution he proposed the implementation of automated data-driven movement via predefined paths. This notion forms the basis for the automatic control of the camera within the context of Mhyusics Auditory Display where motion is defined through a series of interconnected nodes. The properties of each node (Table 6) furnish an increase in the response time and accuracy of the camera while freeing the user from direct control so they may concentrate on other tasks. When created, their initial properties are based on a snapshot of the camera's current location and orientation, which can then undergo finer adjustments via the control window. Additional controls in this window enable the user to alter the fluidity of motion by changing the interpolation method between nodes, advance to any particular point along the path, and choose from a number of path types including open, closed and oscillating. Spherical linear interpolation is also enforced to ensure smooth changes in camera orientation and abate the likelihood of viewer disorientation.

Node Property	Description
Position	Location of the node
Rotation	Camera's orientation when reaching the node
Speed	The constant speed of the camera until the next node is reached *A change in value induces a recalculation of Time*
Time	The time in seconds at which the camera will arrive at the node *A change in value induces a recalculation of Speed*

Table 6: Configurable properties of an automated camera path node

From the perspective of data navigation in an AD each node can be viewed in the same regard as a dynamic beacon (Kramer, 1994a), a time-varying point of interest in the data set that is initially found through a free mouse search. Similarly, a musician might wish to use this viewpoint to draw attention to a specific structure within a composition. The timing of the camera is essential within this context as it adds to the overall sense of musical structure by guiding the viewer to focus on visual snapshots of the environment during designated periods where precise values for motion can support various tempi. This principle has since been explored to some degree by real-time compositional works (Barri, 2009; Climent, 2012) which employ a camera as a virtual microphone to be directed through an auditory scene. In that respect the camera is manifest as the audience's guide through a dynamic visual score where choices made in its direction become part of the creative process. The design of Mhyusics Auditory Display also follows this interpretation. Here, experimentation with the properties of each node, along with the various mapping configurations, enables the user to establish exactly how the data and sound will be perceived. Fundamentally, the objects will continue to behave in a deterministic manner regardless of the camera's presence but the perspective dependent mappings invite the audience to observe them through the cognitive lens of the composer.

3.4.8 Summary

This section started by examining several theories which delineated how the real-time interactive control loop, fundamental to interactive sonification, corresponds to the context of the AD designed as part of this thesis. It was put forward that we have automatically acquired a set of skills from our everyday interaction with physical objects. Moreover, it was argued that these same skills are transferrable to the virtual realm and that, to some extent, the game engine has already expedited this process for many of us. In consequence, there is substantial theoretical evidence which affirms that the model-based embodiment of rigid body dynamics is a worthwhile method for connecting our actions to sound through the direct manipulation of data. From the user's perspective, the interactive model allows actions to be conceived, performed and conveyed, while the addition of parameter mapping permits these actions to express specific elements of the underlying data. In that respect, it was confirmed that the link between gesture and sound could be seen as an extension of the link between data and sound where both relationships are supported by the same notion of a physically informed metaphor. This metaphor was also shown to be a prominent trend in HCI during the exploration of HIDs for mapping gestures from the user to the computer. It was then proposed that the gestural metaphor could be augmented by what was regarded as computational magic. This led to the implementation of a rudimentary scripting language which instructed the computer to orchestrate the parameters of a data model in more precise and reproducible terms. As such, the design could now accommodate a broad range of gestures to be embodied by a simulated object as part of an intuitive methodology for the direction and analysis of both data and music.

Outlining the many integral components of Mhyusics Auditory Display determined how the real-time interactive control is both managed and presented. The GUI comprised qualitative and quantitative representations of the data corresponding respectively to the model-based and parameter mapping elements of the hybrid sonification approach familiar to the design of this AD. As an integral component it realised increasingly complex levels of control in terms of customising data flow, or energy, in the simulation. The most accessible and direct level of control would always be present in the form of the environment window. This qualitative representation of data supported the more involved levels by dynamically reflecting changes made to the quantitative elements found within the optional control window. The exact nature of the qualitative data representation was ascertained when categorising and describing the objects. Each abstraction represented a unique behavioural mechanism which, in several cases, was demonstrated to support a number of techniques for composition and data analysis. Notable quantitative aspects of the GUI concerning the data transformation process were then explained in more detail. The function editor revealed exactly how the link between data and sound could be defined. The messaging system assessed the benefits of sending and receiving the data from inside and outside Mhyusics Auditory Display while suggesting why a many-to-one setup was not appropriate for a metaphoric approach to mapping. Finally, the camera provided the means to navigate and filter the environment, creating a dynamic perspective on the data which could be directed by the bimanual control of the user or the automated actions of the computer.

The next chapter opens with a brief rationale behind the choice of synthesis environment which corresponds with the AD that is presented as part of this thesis. A list of software requirements for implementing Mhyusics Auditory Display are then presented before a comparison with other relevant work reveals a number of advantages which are held by this particular design.

CHAPTER 4 - MHYUSICS AUDITORY DISPLAY

4.1 Implementation requirements

The initial design of Mhyusics Auditory Display (Perkins, 2011) took a number of key ideas from the reacTable (Jordà et al, 2007) where each object 'represents a modular synthesizer component with a dedicated function for the generation, modification or control of sound.' (p. 142). The objects were no longer tangible but were instead realised in a computational form through the combination of a physics engine and a virtual environment. This introduced the scope for more abstract behaviour and forms the basis for where each model could be automated or user-controlled. Relationships between each model also followed that of the reacTable where 'a simple set of rules automatically connects and disconnects these objects, according to their type and affinity and proximity with the other neighbors.' (p. 142). Realisation of the dynamic patching principle required a separate environment for real-time audio processing as there were no available implementations that were supported by the game engine. SuperCollider (McCartney, 2002) was originally selected for this purpose but required considerable work which included the programming of specific functions for determining a signal processing order in addition to synthesis and effect algorithms.

When assessing tools for the purposes of sonification, Worall (2009b) notes that the user has to make a decision as to whether they have the required expertise and time to dedicate to the implementation of specific data transformation algorithms. He argues that 'while SClang is a very elegant and powerful composition language that can support the development of data-processing solutions, being unique, it lacks the transportability that more general and widely available tools afford.' (p. 5-3). In view of his claims we should consider that the general user is unlikely to be a dedicated programmer with expertise in a specific language. One example that warrants such advice is RedUniverse (d'Inverno & Olofsson, 2006) which serves as an audiovisual toolkit for sonifications of simple, two-dimensional dynamic systems. This software held an assortment of functions which featured the potential for an extensive customisation of data conversion and sound synthesis. However, as it relied on a command-line interface for direct user interaction, its audience was limited to those who could employ their understanding of the SuperCollider programming language in order to take advantage of what was provided. Based on this perspective it was deemed more appropriate to undertake a more user friendly approach by defining the information processing methodology within a new framework (Perkins, 2012a; Perkins, 2012b). While this approach has been expanded upon and discussed in more detail throughout this thesis, there now follows a more practical discussion which introduces the tools (Table 7) used for the development of Mhyusics Auditory along with a brief explanation for their selection.

Component	Purpose
ASIO4ALL audio driver	Low latency audio
BEPUphysics	Physics engine
Bespoke.OSC	Communication using the OSC protocol
C# MIDI Toolkit	Communication using the MIDI protocol
Max/MSP	Audio synthesis environment
Neoforce Controls	Controls for the graphical user interface
Visual Studio 2010	Integrated development environment
XNA Game Studio 4.0	Game engine

Table 7: A list of core components used for the development of Mhyusics Auditory Display

Given that real-time audio synthesis was not available within the game engine, and SuperCollider had proven to be unsuitable, Max/MSP (Puckette, 2002) was chosen to fulfil the role of producing audio based on the dynamic state of the rigid body data. Max/MSP was deemed an appropriate choice as it offers a graphical environment with interactive controls that rely on the same direct manipulation interaction metaphor as those found in the Mhyusics Auditory Display control window. In addition, Puckette (2002) revealed that 'a central design principle of Max is that, to the extent possible, the patch that appears on the screen should provide a full documentation of what is happening.' (p. 36). The corresponding patch for Mhyusics (Figure 17) follows this principle by using concisely labelled objects to both define and hide the more complex underlying structure of the message routing and sound synthesis algorithms. The visually compact nature of these objects avoids the abundance of textual information found in SuperCollider and presents the user with only the necessary information required for the purpose of the patch. In this case, toggling the audio on and off, adjusting the gain, testing the synthesiser voices and displaying the valid OSC commands that can be used to dynamically adjust the state of each voice.

During the initial stages of development it was found that there was a sizable degree of latency between the visual events in Mhyusic's environment window and the corresponding changes in Max/MSP's audio synthesis when using the standard Windows audio driver. By relying on a hardware independent ASIO (Steinberg Media Technologies GmbH, 2013) driver (Tippach, 2013) the latency could be significantly reduced and an apparent synchronicity between the auditory and visual channels would be gained.

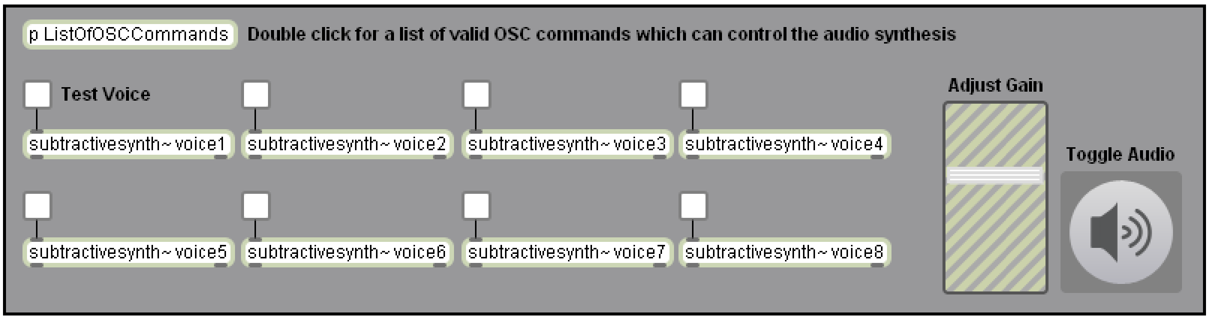


Figure 17: The GUI for the corresponding synthesis environment to Mhyusics Auditory Display is simplified to only present the necessary information

Mhyusics Auditory Display was predominantly written in the C# programming language using Visual Studio 2010 as its integrated development environment (Microsoft, 2013c). The design incorporated Microsoft XNA Game Studio 4.0 (Microsoft, 2013b) which provides a number of tools for facilitating video game development including a rudimentary game loop with update and draw functions as well as a number of classes for handling mathematics, hardware accelerated graphics, user input and multimedia. A number of external libraries were imported into the project during development in order to add further functionalities to the core code. As a principle of the design process each of the additional libraries had to be both open source and written in C# in order maintain compatibility with the existing codebase and allow for a straightforward means of modifying and developing the imported code. This proved worthwhile when accommodating specific features of the auditory display. For example, the function editor was created by combining and extending a number of basic user interface objects found in Neoforce Controls (Shane, 2013). Likewise, the gravitational field objects derived from BEPUphysics (Nordby, 2013) using a combination of a basic rigid body entity with a vector field that was constrained by a bounding volume. The integration of the remaining libraries, Bespoke.OSC (Varcholik, 2013) and the C# MIDI Toolkit (Sanford, 2013), required no further changes to their core functionality.

An important design consideration was to keep the code as platform independent as possible. In particular, saved configurations use the highly portable Extensible Markup Language (XML) format for serialising data so that key information from the auditory display can be shared and interpreted on the many possible devices that users may own. The C# language behind Mhyusics Auditory Display can also be considered platform independent but the application found on the DVD which accompanies this thesis is currently tied to the Windows operating system due to its dependence on the Microsoft .NET framework (Microsoft, 2013a). However, there are open source implementations of this framework, such as Mono (Xamarin, 2013), which will allow the source code to build and run on all the major desktop and mobile platforms. This will potentially enable future iterations of the software to become available to the many devices that the corresponding configuration files can be interpreted on.

4.2 Comparison to other work

One can argue that there are a number of comparable ADs that would be enhanced by the dynamics of a modern day physics engine. Sturm (2001a) proffered many ideas for synthesis from simulation data concerned with particle physics. Be that as it may, he was impeded by the computational complexity required for the many-body simulations which resulted in a large latency period between the simulation and the resultant audio signal. This proves unfortunate because scenarios were presented which demonstrated his system's feasibility as an algorithmic composition tool and he anticipated that musical performance would be possible with the implementation of a real-time system. The emergence of improved technology, alongside the continual development of both game and physics engines, has since provided an opportunity to overcome such limitations and realise many of the ideas laid out in his work.

Hamilton's (2008) investigation was concerned with the data accompanying the wide range of interaction gestures that can be performed by a character in a first-person shooter. He customised a game engine for the purposes of sonification by employing the OSC protocol to route gestural information to objects found in Pure Data (Puckette, 1996). Consequently, this brought about the custom synthesis and control of sounds derived from context sensitive avatar actions, all of which pertain to recognisable physical events. While most of the gestural events corresponded to a single finger press, more complex actions arose from orientation of the avatar, which could determine the initial direction of projectiles spawned by the action of shooting. However, the subsequent behaviour of these discharged objects was determined by the calculation methods provided by the game engine. This resulted in simplified approximations of motion and response in comparison to the more considered interactions offered by a modern physics engine.

The authors of VR-RoBoser (Le Groux et al, 2007) established a context dependent sonification when making a case against predetermined, repetitive audio in a virtual environment. They also chose the gestural medium of an avatar which could be controlled either by a human or computer. This idea again demonstrated that the behaviour of an avatar could generate a unique set of sounds based upon its relationship with its surroundings. In this case, the integration of a modern physics engine would help to create a less static environment while adding another level of variability to the information perceived. Another conceptually similar audiovisual compositional tool chose to populate a three-dimensional scene with more animated entities (Barri, 2009). This example relied on an autonomous camera to act as a roving microphone, creating a sense of motion, or unique journey, when travelling past a collection of audible objects (Figure 18). The absence of a physics engine dictated that objects would remain unresponsive to their surroundings and were instead constrained to a predetermined path in order to procure any motion. Furthermore, the author wrote that 'all the entities generate continuous, monotonous sounds of which the perceived volume only changes as the actor moves closer or further away.' (para. 29). He commented that 'as a consequence the compositions have become quite ambient and drone-like.' (para. 29). The combination of the mapping and messaging systems accommodated by Mhyusics Auditory Display facilitate a more flexible audio correspondence where such rigid definitions do not have to be adhered to. In turn, this bestows the user with a wider range of possible qualities for compositional techniques within the context of a more dynamic simulation.

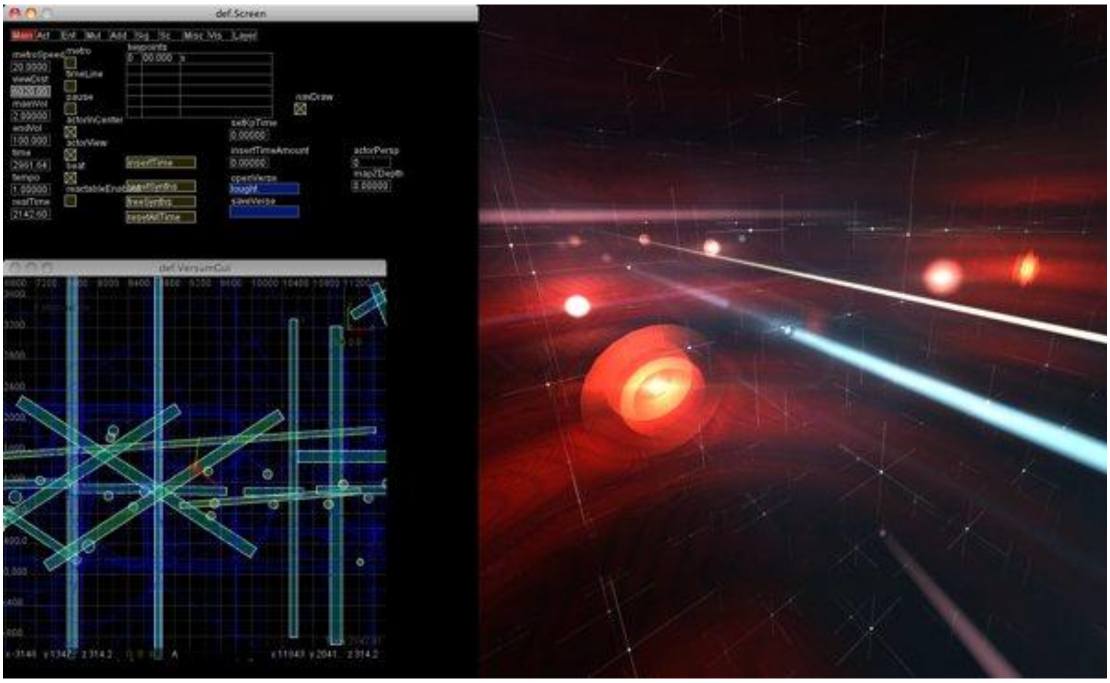


Figure 18: The Versum GUI displays the traversable 3D environment while overlaying a 2D top down representation of the same environment alongside a parameter control window

There are several established precedents which have associated sound in response to more complex simulation systems. The predominant approach in their regard is to accept the physical modelling paradigm as a means of investing simulated objects with the attributes of sound. When establishing the connection between data and audio there was evidence to suggest that this methodology presented a number of drawbacks in comparison to the complete separation of audio attained with the parameter mapping approach. One of the more prominent points in this regard was the inseparable link between the mechanical and acoustic systems. In some cases (Tache & Cadoz, 2009) this necessitated the distinction of macro and micro temporal objects which would hinder audible analysis of data associated with lower frequency bandwidths. On the other hand, one project (Henry, 2004) has demonstrated some of the possibilities afforded when audio is not determined by the resonating structures themselves but by their associated parameters. However, like other physical modelling approaches (Howard & Rimell, 2004; Castagne et al, 2009) it relies on a mass-spring system to govern the actions of the creatable objects. This simulation type has been understood to have a number of shortcomings such as instability (Henry, 2004; Mesit et al, 2007). In particular, it is often difficult to introduce more abstract spring behaviour such as an increase in stiffness that is required for them to simulate rigid bodies. Moreover, creating a simulation of this type requires smaller time steps for integration which, in turn, results in a greater demand on the processor.

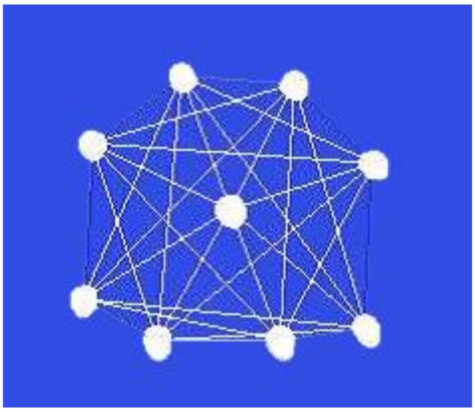


Figure 19: Henry's simulation of a ball comprises a number of point masses connected via springs

One major principle of a mass-spring system is that they simulate objects as the sum of their parts (Figure 19), thus an object's behaviour stems from the individual characteristics of its comprising elements. While this approach has the advantage of being able to simulate soft-body dynamics, or deformable objects, it can pose a considerable problem within the framework of the AD designed as part of this thesis. Castagne et al. (2009) drew attention one particular issue they wrote:

> 'physics mass-interaction networks make it difficult to strictly isolate a sub-network. By nature any part interacts bi-directionally with other parts. While designing, each part regularly requires to be modified, either in its network topology or in the parameters of some modules, in order to be adapted to other interacting parts.' (p. 409)

This not only shows that specific object configuration requires considerable thought and expertise, it also creates a situation where it is harder to compartmentalise the data because it always exists as part of a network. The user is now presented with a more fragmented mental model comprising many integral nodes, each of which represents a unique data set. This infers that a metaphoric approach to data comprehension would be less feasible as sources can be confused. For instance, if we consider the mapping of velocity to sound it becomes harder to define which node's velocity should sonically describe the behaviour of the encompassing object. Further confusion arises should the velocity vector of one node be moving in the opposite direction to another. In this scenario, it is possible that an ascribed sound attribute will not reflect the general visual behaviour. In addition, movement induced through object interaction is less likely to be uniform across all nodes and thus not as predictable.

In comparison, rigid body simulators are less computationally demanding, easier to configure for specific behaviours and create a stable model for data representation. These factors have helped to make it the dominant physics simulation paradigm where it has been subject to extensive development. In turn, this has procured more efficient conditions for real-time interaction, especially when established within the framework of a gaming environment. Unfortunately, game audio is still generally limited to synchronising sample playback with the onset of physical events. As a consequence we are left with a stark contrast between intricate visuals and perfunctory audio, which could instead become enlivened when derived from the same continuous and dynamic interactions. This motivated one author (Menzies, 2008; Menzies, 2010) to select a rigid body simulation as a source to procure interactive sound generation through physical modelling. However, generation of the audio came exclusively from the dynamics of excitation and object resonance as opposed to parameter mapping. This not only restricted the relationship between mechanical and acoustic systems but also ensured that the mapping was less apparent to the user due to the absence of a specific intermediary function. In that sense, users were expected to empirically calibrate parameters in a rather indirect manner in order to change a particular quality of sound, such as an increase in frequency, making it less appropriate for the intentions of AD.

Recent developments have realised cross-platform game prototyping platforms with integrated physics engines (Unity Technologies, 2013). Bukvic and Ji-Sun (2009) noted the potential this software may hold for the sonification of physics engines but chose not to underpin their work with the theoretical studies pertaining to AD design. Dolphin (2009a) gave a more robust account of the possibilities brought about by tailoring interactive simulations for the prospects of composition. One of these projects functioned as a dynamic sonification of a physics engine (Dolphin, 2009b) described by the author as a 'controllable sonic environment intended for "non-experienced" users.' (p. 159). With this particular audience in mind Dolphin (2009a) also stipulated that:

> 'player control of the sound world is symbolic, with no sound or synthesis parameter names included in the virtual space. This allows the virtual space and sound world to be explored by the player without the distraction of specialist terminology that may be intimidating, confusing or 'off-putting' for a user with limited knowledge of sound synthesis terms.' (p. 217)

While he makes it clear that MagNular (Figure 20) was never intended for the purposes of composition or sound design, it was based on our fundamental understanding of gravitational forces as a metaphor for interaction and sonic exploration. This serves to indicate that such a preconfigured environment can offer a more inherent and direct route to the production of sound. Furthermore, it can now be argued that by placing his interaction metaphor within the design framework described throughout this thesis we have the understanding that it can be augmented. Indeed, as previously demonstrated, defining a specific area in which to constrain the gravitational vector field can concede another level of control and produce novel rhythmic possibilities.

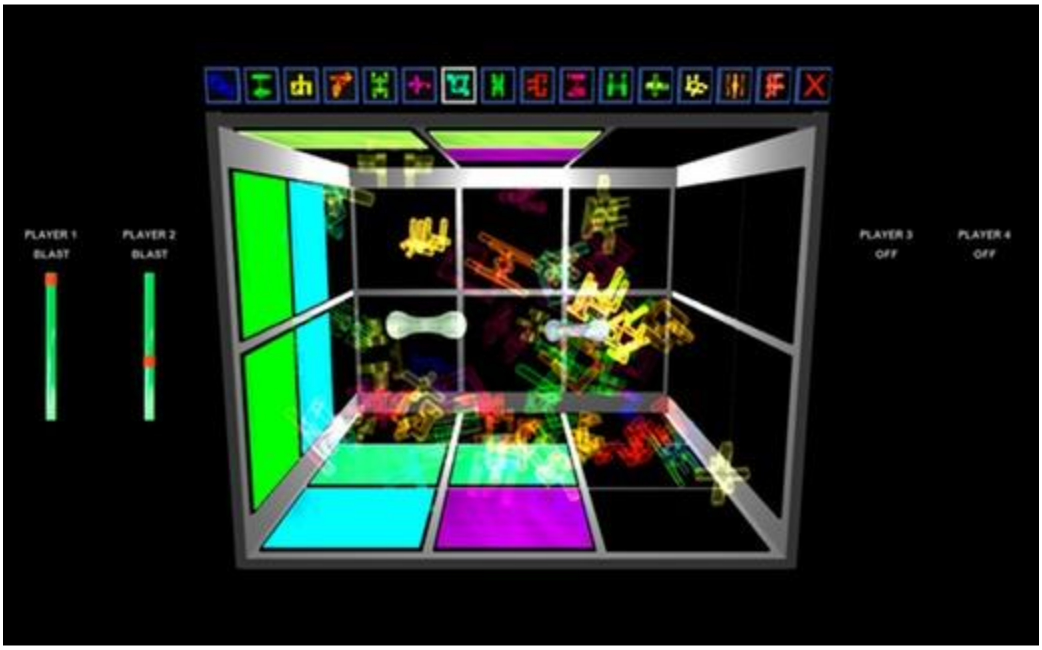


Figure 20: MagNular allows for symbolic control over predetermined sound parameters

In addition to academic research there is evidence that sonifications of simulated physics systems can relate to a more general audience and achieve some mainstream recognition. Electroplankton (Nintendo, 2012) is a videogame that depicts an algorithmic process through the metaphor of leaves and anthropomorphic projectiles. By rotating the leaves the user can determine the trajectories of the projectiles and thus control the period in which a collision event will occur. However, each collision event must be accompanied with the simple playback of an inbuilt tone as there is no concession for external audio or processing tools. Moreover, the simplicity of the overall simulation is reflected through the lack of interaction between the moving objects. Physynth (Simian Squared Ltd, 2012) is another commercially available application that relies on real-time simulation events to procure sound (Figure 21). Unfortunately, it is also compromised by the simplicity of its simulation which restricts all interactive objects to either a dynamic point or kinematic segment, both of which are presented in two dimensions.

Figure 21: Physynth generates sound to coincide with the collision events generated by a simplified simulation comprising dynamic points and kinematic lines

PipeDream (AnimusicLLC, 2008) uses offline analysis of MIDI events to procedurally generate convincing motion algorithms. The result of this is an animation that depicts a complex contraption, assembled from recognisable instrument parts interspersed with rigid body devices (Figure 22). As the animation proceeds many of the rigid body devices spawn convincingly simulated spherical entities which generate a musical tone when coming into contact with the instrument parts. Having been the subject of much praise it has since been replicated both virtually (Advanced Micro Devices,

2013) and physically (Sisu Devices, 2011). While only the physical derivation can be considered a real-time demonstration, all the arrangements never deviate from the original music and instead adhere to a strict set of scripted events. In spite of this fact, given enough time with the tools afforded by the design discussed throughout this chapter, it should be possible to reproduce a more versatile derivation within an AD. When realised in this context it exposes the opportunity to prototype complex mechanical contraptions which are then qualified for data analysis or distinct compositional techniques.

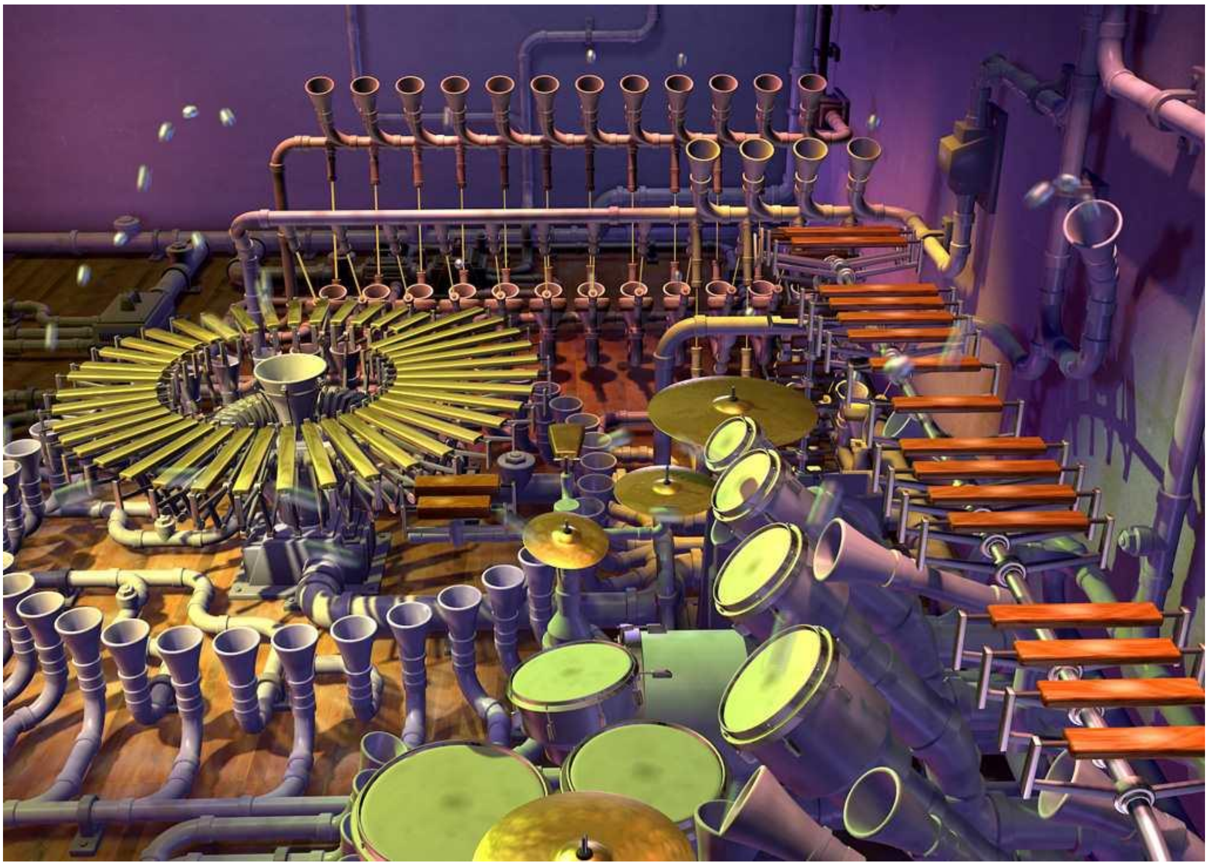


Figure 22: PipeDream synchronises collisions involving animated spherical objects and musical instruments with audio events

CHAPTER 5 - CONCLUSION

5.1 Examining the design

The aim of this thesis has been to broaden the appeal of the sonification process by formulating an interactive toolset which upholds the physics engine as a well-established data set with a cross-cultural understanding. While ongoing technological advances have enabled the development of real-time physics engines across a broad range of devices, the multimodal interactive qualities arising from their integration into the modern day game engine realise a potential framework for interactive sonification. When adapting this framework for the purposes of AD, the simulation of physics creates a gap between the underlying data, its representational modalities and the user. Appropriately, the theoretical framework behind interactive sonification advocates the use of physical causalities for the sake of comprehending these gaps. However, this investigation submits that this comprehension process can also be incrementally structured in the form of mapping, metaphor and magic. In consequence, it demonstrates that while mathematical mapping processes have been understood through metaphors pertaining to physical laws, the extraordinary potential of the computer can interpret and augment these intuitive metaphors for the sake of an improved design.

The rendering of rigid body dynamics in the context of a games engine was found to address two fundamental problems of data exploration (Hunt & Herman, 2004). Here, the data inhabits a familiar space while assuming a polygonal form that offers an intrinsic means of interaction through an avatar. Furthermore, because this form of data exists as a simplified analogy of a physical object, there is evidence to suggest that we all share a basic understanding for the generation and interpretation of its actions (Bruner, 1966; Gibson, 1977; Dourish, 2001). Such evidence has also been reasoned to be transferable to the virtual realm where it can be used for more creative purposes such as interactive scenarios involving digital musical instruments (Cadoz, 2009). From the perspective of interactive sonification, the transferable nature of this evidence serves to indicate that rendering of rigid body dynamics can also be regarded as an intuitive means of interpreting the gap between data and video. In that respect, accepting a conventional mapping approach results in an intermediary model which affords the performance and analysis of familiar tasks involving sound, while its actions serve as a metaphor for the dynamics of the underlying data.

There is evidence from the literature that our experience of physical laws, arising from our everyday interaction with objects, provides us with an inherent body of evidence to interpret the gap between data and audio (Gaver, 1993; Hermann & Ritter, 2004; Jensenius, 2007). The simulation of these same laws relating data and audio has demonstrated that, like our understanding of actions, this reasoning is transferrable to the virtual realm. Indeed, simulated entities incorporating the rigours of physical modelling synthesis have called upon our understanding of their actions and associated sounds for the conception and comprehension of action-based music (Kojs, 2009). The field of sonification also encourages the use of this knowledge when performing interactive tasks with a sonorous model that embodies physical coherences for the sake of data analysis (Hunt & Hermann, 2004). However, unlike the strict ties of MBS, it was found that a PMS approach could interpret the gap between data and sound in a more creative manner. While this introduced the idea of "the mapping problem" (Worall, 2011), a solution embodying a physically informed mapping proposed that we could rely on the same knowledge to inform convincing metaphors in lieu of strict physical laws. Relevant to the context of this design, this criterion was regarded as a compromise between further empirical evidence towards a more conventional solution and musical expression that did not necessitate such a solution. Similar metaphorical reasoning which substantiates this approach has been submitted in place of a natural physical coherence between action and sound as part of an action-sound relationship (Jensenius, 2007). This was argued to be expected as both relationships would coincide when the data is embodied in a model which directly links actions to sound. Be that as it may, it was appreciated that this alternative frame of reference would allow the model to be recognised as a metaphor for emotion and musical activity.

The possibility of bypassing strict physical coherences when creating a relationship between the data and audio meant that there is a need to acknowledge our automatic perception of auditory gestalts. Analogous to visual gestalts, they serve as indicators of prominent data features where inherent connections are yet to be formed and are believed to complement the act of analytical everyday listening (Hermann & Ritter, 2004). They are also congruous with the notion of auditory scene analysis (Bregman, 1990), whose grouping cues for the fusion and segregation of sound streams resemble those classified by the gestalt principles for perceptual organisation. These unique modes of listening characterised the integration of a PMS paradigm which, in comparison to solely relying on a MBS approach, could grant a more explicit level of control over the data transformation process and recognise dynamic data changes in the absence of user interaction. In consequence, sonification would now take place under a variety of voice conditions where the arrangement and classification of information flow could be supported by gestalt laws. It was also possible to introduce two data streaming conditions which provided either continual or static feedback for specific parameters while recognising their applications for both scientific and musical ambition.

While the PMS paradigm can offer some advantages, the model-based representation still plays an important role in the data transformation process as the auxiliary feedback of the visual channel is able to reinforce where the gestalts arise and predicate their emergence. Moreover, by adopting monophonic behaviour each model is guaranteed to only generate a single sound stream at any given moment, which can then be informed by any combination of the underlying data parameters. Direct comparison of data is now feasible since each additional model will generate a single stream that derives from the same pool of parameter types. Consequently, the state of the model can also support the perceptual organisation of these parallel streams by depicting the emergence of grouping cues familiar to auditory scene analysis.

The representation of the underlying data in this combined manner corresponds to a hybrid sonification method known as model-induced parameter mapping. As a method for the interpretation of data through audio it presents a number of advantages over the widely recognised physical modelling approach. It was submitted that the parameter mapping aspect provides a more explicit link between data and sound. This avoids the need to empirically calibrate parameters in a rather indirect manner in order to change a particular quality of sound. It also allows for any combination of the parameters which attribute to a simulated object's behaviour to be aurally examined across a range of synthesisers and synthesis methods. Furthermore, the richer palette of sound implied by such a range can be procured from a greater assortment of gestural interactions as the mechanical movements of the model are not required to occur within a particular frequency band. Despite these benefits for data analysis, there are a number of limitations which come into effect when relying on a metaphoric approach. A many-to-one parameter configuration is not advisable as it requires some form of interpolation between input data values which can lead to confusion over the true target domain. This confusion is also one of the prominent reasons why, unlike physical-modelling, metaphoric parameter mapping is not ideally suited to a mass-spring system. While useful for soft-body dynamics, this system simulates an object as a sum of integral nodes, each of which represents a unique data source. The user is therefore presented with a fragmented mental model where each source acts independently in a manner that does not necessarily align with the perceived behaviour of the encompassing object.

In many respects the infrastructure of the game engine was ideally suited to the data transformation technique chosen by this design. The discrete time-stamped nature of the data parameters, arising from the frame-by-frame update of both the game and physics engines, is perceptually advantageous as it allows for the mapping of data time to listening time and the synchronisation of the audio and visual representations. While this synchronicity has been validated as an important element for optimising the localisation of data (Spence, 2007), the game engine also affords the integration of sound processing libraries. In that regard, it further enhances data navigation by treating each model as a dynamic sound source in a manner that imitates our perception of the real world. It was disclosed that the idea of synchresis (Chion, 1994) also relies on this synchronicity which, when coupled with a data transformation method that is sensitive to gestalt laws, can further assist in creating a seemingly logical connection between video and audio events. The tight temporal connection between data, audio and video also makes it possible for the model to provide intuitive feedback to those outside the control loop. Overcoming one of the shortcomings of PMS, this concurrent arrangement served to dynamically reflect both the state of the sound and the data to an audience that would include the user when delegating control to the computer. The ability to easily manipulate the appearance of each model within this infrastructure was also a crucial factor for the data transformation process. Here, the model could assume more abstract qualities to better afford the expansive sonic palette furnished by the parameter mapping approach and be assigned a particular colour in order to assist grouping cues.

The power of the computer to both manage and present real-time interactive control of the data in terms of qualitative and quantitative representations would respectively support the model-based and parameter mapping elements of the hybrid sonification approach and inform the structure of the GUI. The close ties between these two representations allowed them to work in tandem by dynamically reflecting any changes made to their counterpart. This was regarded as crucial for maintaining increasingly complex levels of control in terms of customising data flow, or energy, in the simulation. In that respect, the most elementary level bypassed the need for text-based operating instructions by using abstractions of a rigid body, realised as modular objects, to characterise particular data processing logics while offering a familiar model-based direct manipulation approach to interaction. At a more involved level, quantitative aspects of each object are exposed to provide a discrete and dynamic account of their underlying parameters and mapping configuration. When expressed in this manner these parameters could be accurately adjusted by the user and orchestrated in more precise and reproducible terms through the scripted actions of the computer. Together, these representations become a language for establishing dynamic data structures, instruments, scores, and algorithms of varying complexity. Furthermore, as it is possible to quantitatively describe each of these activities they can be stored and recalled to the benefit of the user. This led to the implementation of an iterative approach which recognises a discrete stage of planning, practice and familiarisation, where the creation and categorisation of reference points can be used to assist in their analysis, interpretation, modification and development.

When reviewing the many formative decisions that have contributed to this design it is clear that the interpretation of physical causalities pertaining to rigid body dynamics can relate to common knowledge. Indeed, such a pervasive understanding forms the basis of a number of intuitive metaphors which extrapolate from certain regions of our experience and support the sharing of ideas between the disciplines of science and art. While this cross-disciplinary reasoning appeals to a wide audience, the versatility of these metaphors can be appreciated further when they are enhanced by the operational capabilities of the computer. In this context they have been shown to uphold existing strategies while allowing them to be applied in a manner that benefits the purposes of AD. In effect, they substantiate a novel design which, from this examination, has been demonstrated to overcome several limitations of comparable work.

5.2 Future work

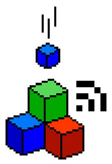
It is hard to imagine that both game and physics engines will not remain relevant in the foreseeable future due to our insatiable demand for immersive entertainment. With the emergence of new technology these engines will continue to evolve and thus further empower others to interpret the framework presented here in new and interesting ways. In a sense this may suggest that what is regarded in this thesis as magical will eventually become ordinary and perhaps invite cynicism. Be that as it may, given the gaming infrastructure of this design it might be possible to appease this cynicism by introducing the notion of a Metaverse (Ondrejka, 2004). In turn, this can encourage collaborative efforts, expedite the number of potential interpretations and increase the possibility of novel and exciting ideas. It is also worth acknowledging other gaming elements, such as objectives with reward structures, which have been argued to motivate participation and engagement in the context of both sonification (Dolphin, 2009a) and music (d'Escriván & Collins, 2011; Liebe, 2013). In addition to these gaming elements, Mhyusics Auditory Display may become more appealing to a mainstream audience by accommodating further concepts in keeping with traditional Western music. For instance, pitch could be quantised to a series of predetermined notes, such as a scale, based on divisions along the range of the mapping function. Likewise, the onset of a voice could be adjusted to match timings derived from common metric levels according to a global tempo. However, in these circumstances the effect should be subtle in order to avoid losing a perceived concurrency between the audio and visual events.

Although the increasing complexity of modern day video games presents a number of technical hurdles that discourage the use of save state systems (Young, 2013), the bespoke nature of the design behind Mhyusics has been able to address these concerns. Moreover, as technology improves it will become possible to enhance the current system by realising a more dynamic iterative mechanism through the implementation of a record mode, in addition to the current pause and play features. This new mode will be capable of maintaining real-time interaction while simultaneously storing the state of the simulation each frame and recording the audio output. When the user decides to terminate the recording process they will obtain an offline time frame which can then be traversed both forwards and backwards in a more fluid manner. During this period various aspects of the AD can be adjusted, such as the mapping configuration or camera parameters, which then initiates a transition to a real-time simulation, or play mode, so their effects can be made apparent. It should be noted that obtaining a smooth transition from recording to real-time will require some careful tuning of the current mapping system as the audio is envelope driven.

It is possible for this framework to be enhanced by the continuing trend towards a more natural means of interaction within virtual environments. Direct manipulation metaphors can be realised by a combination of stereoscopic displays, which provide depth clues for handling simulated objects, and depth sensors, whose greater accuracy enables them to recognise a range of increasingly precise gestures (Hilliges et al, 2012; Weichert et al, 2013). It is also possible to extend the first-person camera approach for data navigation to include head mounted displays which, with the recent addition of low latency head-tracking, provide a more intuitive alternative to a pointing device (García et al, 2013). Many of these developments correlate with a future of sonification that is envisaged by Hunt and Hermann (2011). They proposed that 'with immersive 3D graphics, and fluent interaction using strong physical interaction metaphors, files and folders become virtual graspable units that the user can physically interact with in the information space which interweaves and overlaps with our real-world physical space.' (p. 296). While this thesis has suggested why immersive graphics might not be necessary, their advancement can be considered as a predominant driving force behind the evolution of the game engine, and further evidence as to why these engines will continue to remain relevant in this domain. Likewise, Hunt and Hermann's idea that strong physical interaction metaphors will persist in HCI compliments the many arguments presented throughout this thesis indicating that the physics engine can continue to play a crucial role in the evolution of interactive sonification.

APPENDIX 1 - THE ACCOMPANYING DVD

The accompanying DVD contains the following material which is organised inside folders corresponding with the name of each major bullet point.



- **Demonstrations**
  - A video demonstration which provides a brief introduction to Mhyusics Auditory Display
  - The five video demonstrations that are referred to in the thesis and indicated by the icon located in the left-hand margin of this page
  - The Mhyusics Auditory Display configuration file for each demonstration
- **Deprecated Work**
  - Deprecated versions of Mhyusics Auditory Display, classes, GUI code and SuperCollider patches
- **Mhyusics Auditory Display**
  - The auditory display described in this thesis which enables the user to highlight various features of rigid body dynamics for the purposes of analysis or composition
  - The source code for Mhyusics Auditory Display
- **Optional Software**
  - Audio drivers, MIDI tools, a player for the video demonstrations and a standalone VST host
- **OSC Patches**
  - A Max/MSP patch that will receive OSC messages from Mhyusics Auditory Display and generate sound from the rigid body data (this is the same patch that is used in the video demonstrations)
  - A SuperCollider patch that will send OSC messages to Mhyusics Auditory Display as a method of controlling objects such as the camera
- **User Manuals**
  - The user manuals for Mhyusics Auditory Display and the Voxel Mesh Editor
- **Voxel Mesh Editor**
  - A tool for creating voxel meshes which can then be imported into Mhyusics Auditory Display
  - The source code for the Voxel Mesh Editor

APPENDIX 2 - FORMATIVE PUBLICATIONS

SONIFICATION OF A REAL-TIME PHYSICS SIMULATION WITHIN A VIRTUAL ENVIRONMENT

Rhys Perkins

Anglia Ruskin University,
Department of Music and Performing Arts,
Cambridge, CB1 1PT, UK
rhysperkins@rhysperkins.com

ABSTRACT

There has been an increasing amount of research utilising 3D virtual environments as a core component of interactive sonifications. While showing considerable potential for their ability in producing both real-time visualisation and sound, they often come with constraints as a result of their design decision processes. This paper presents developments of a prototype that has arisen out of my attempts to address some of the issues involved in bringing sonification to a wider audience through a universal metaphor. These new additions allow for an intuitive, elementary introduction into to the world of auditory display, while providing a more flexible and immersive environment for composition and sound design.

1. INTRODUCTION

The emergence of improving technology has provided an opportunity to overcome the limitations of previous work [1] where computational requirements would produce significant latency between the audio and the visual, inhibiting real-time interaction. Dedicated hardware, such as the graphics processing unit (GPU), allow for the sharing of resources and have recently shifted their focus towards more general purpose computing [2] including accelerating physics simulations. Freely distributed 3D physics engines such as Bullet[1] and Open Dynamics Engine,[2] common middleware solutions for modern game engines, are readily available thanks in part to an increasing demand for realism amongst gamers. Their cross-platform approach sees widespread use of long established solutions such as the desktop computer but also extends to mobile devices. As a result this presents an opportunity to revisit previous compositional techniques [1] and reach a wider audience in the process.

Nguyen's approach to TONAL DisCo [3] acknowledges the benefits of using a game engine when combining dynamic visuals with audio but chooses to use a pre-processed sample library over real-time sound generation. The prototype laid out in this paper addresses that limitation by utilising a messaging system that provides a link between the visual and the sound synthesis. Pre-processed sample libraries alongside dynamic visuals have started to emerge in commercial software with applications like PhysSynth [3] for the iPad. However, I would consider this application to be compromised by only simulating basic particles with simple collision detection and response, comprising interactions between points and segments and restricting the process to two dimensions. By sonifying 3D physics engines we can surpass these restrictions to cover more interesting and complex interactions, communicating and exploring a wealth of new dynamic ideas within an auditory display whose interface with the user resembles the physical world.

Sturm [1] laid out some of the benefits that a sonification of particle physics simulation would bring to science, including new ways of understanding physical phenomena, and refers to several artistic merits such as bending of those scientific laws to suit a composer's taste. Pedagogical advantages that Western music composition students might gain from using an audiovisual simulation are also discussed. In particular, he suggests there is an advantage to combining the audio and visual modalities in order to present musical ideas, likening it to listening to a piece whilst reading a score, rather than partaking in one or the other activity separately. Metaphorical correlations between particle physics and sound synthesis have been explored [1][4], serving as a means of providing a bridge between the two aforementioned fields whilst highlighting cultural differences and the problems that might arise from them. For instance Sturm [1] states that composers must possess skills in physics to begin with and only the audience need not be versed. The system described in this paper was designed to cater for all levels, from the well-versed user who comprehends and wishes to explore and extend the open-source scientific algorithms to those that would like to immediately compose and play.

RedUniverse provided a toolkit for sonifications of dynamic systems [5] with the aim of producing a playground for compositional ideas. These systems were also limited to two dimensions and lacked accessible interactivity, requiring a good knowledge of the SuperCollider programming language in order to take full advantage of their potential customisation. What I present here will allow for immediate use and configuration via standard input, such as a keyboard and mouse, but will also cater for further inputs using standard protocols such as MIDI and OSC.

Interactive compositional tool, VR-RoBoser [6] makes a case against predetermined, repetitive soundscapes in a virtual environment by using a context dependent sonification. They present the idea of a user-controlled or autonomous avatar that continuously reacts to its unchanging surroundings in order to overcome this issue. I would argue that the dynamic nature of physics would help create a less static environment. Continuous user interaction would stimulate audible results as the simulated

[1]http://bulletphysics.org/wordpress/
[2]http://www.ode.org/
[3]http://www.physynth.com/

objects react accordingly. Automated movement can be accomplished through inter-object logic allowing the user to pay attention to other aspects of this system, such as the proposed camera control.

In previous work [7] I laid out the design process behind a musical tool employing a 3D space that could be populated with audiovisual objects acting under the laws of physics. Each object exposed fundamental data dimensions that could then be mapped to sound dimensions via the OSC protocol, providing an audible insight into their behaviour within the current environment. A modular approach to the object design meant that a user would construct logic through object interaction and association without the need for coding keywords, operators and the understanding of basic programming paradigms.

The tool comprised an environment where polygonal models, visually representing the underlying simulation data, were introduced alongside a graphical user interface (GUI) that offered the opportunity to control fundamental properties, sonify and compose in real-time. Since designing the initial prototype I have found the need for improvements in a few key areas. This paper will discuss some of the fields I have considered when refining my approach to the sonification of an interactive physics engine. In the next section I will explore the idea of human interaction and how the process came to shape the design of the basic objects. The objects are then discussed along with some thoughts on their potential behaviour and how this can affect the overall output. This is followed by the mapping section which explains why I believe that the same simulated objects are inherently easier to map due to their physical grounding, along with several theories that underpin the conception of the mapping function tool. The camera section describes how the user will view and traverse the environment, how it accentuates the user's experience and why they should be presented with the option for both automatic and manual control configurations. This then leads into the messaging system where the aim is to make the same experience more flexible and personal. The paper concludes by stating some of the advantages the proposed system has over another comparatively close environment [8] along with my thoughts about potential future work and emerging areas of interest.

2. KEY COMPONENTS

2.1. Human Interaction

Investigation [9] has shown that applying human interaction in real-world contexts to sonification can help improve interface design. In this paper, the researchers state that humans are adapted for interaction within their physical environment and making continuous use of all their senses. When we perform an action on an object we expect some kind of reaction and our perception of objects builds up over time through this interactive process. The objects found in this system adhere to the unchanging laws of physics that our neural hardware has been effectively programmed to deal with over many years of evolution. This enables the user to utilise their acquired skills in order to manipulate high-dimensional data via objects with familiar behaviour and response. The authors also argue that one of the main problems in the domain of data exploration is that the data often inhibits a high-dimensional data space that is different from the 3D space we are familiar with. In this prototype the simulation data emulates rigid bodies, and their behaviour in our natural environment, providing familiar grounds for both exploration and interaction.

Interaction with the objects has a direct effect on the procedurally generated simulation in a similar manner to model-based sonification [10] where the user supplies the initial excitation. By grabbing, moving and throwing objects it is feasible to perform a wide range of actions from striking, to more delicate procedures such as plucking; each of which results in changes to the data dimensions. This direct process introduces information manipulation to the average user at a more accessible level when compared to other similar applications [11] since no coding knowledge is required. For example, saved scenarios containing preconfigured entities can be loaded, ensuring that user interaction yields instant audible results.

If the user wishes to create and save their own scenario, or edit existing ones, some basic GUI control knowledge is required. The controls can be toggled at any time and aim to present a more traditional and precise means of modifying the attributes that influence each aspect of the prototype. Presentation of the data in this manner brings its own set of problems in that the potential to overwhelm the viewer with information is increased. When considering high-dimensional spaces one study [12] argues for a mental model simpler than brute-force awareness of every detail in order to avoid cognitive overload. The authors suggest that parameters should be cross-coupled so that the performer naturally thinks of certain parameters as varying together in predefined patterns. The high-dimensional data, encapsulated visually by each model, allows us to intrinsically understand how the parameters vary together. Throwing an object would imply a change in velocity that would be influenced by the mass of the object. Spherical objects are more naturally inclined to roll, providing smoother changes in angular velocity as opposed to the sudden, erratic changes of their square shaped equivalents.

Research into improving sonification tools [4][9] has questioned how information should be distributed to different modalities in order to maintain the best usability. As stated previously, our everyday interactions with physical objects providing a base level for our conceptual understanding of the data dimensions found within. With this in mind I highlighted what I believed were the important elements of the underlying data, choosing to expose those that had a direct impact on the representative model's behaviour. Given the longstanding synergy between humans and physical entities I would suggest that less mental bandwidth is required to comprehend the visual events. Instead, the attentive capacity of the user can focus on the audio, and its governing mapping process, encouraging sonic exploration and creativity.

2.2. Objects

Objects provide a modular approach to the way the user experiences the underlying data. Depending on the object's configurable physical parameters the program will automatically simulate subsequent interactions as the object reacts to its current environment and user intervention. However, it can be argued that there are parameters that have no direct effect on the simulation which are just as important for the user to exploit.

These properties can enhance a user's experience, and encourage them to learn, by creating associations through further visual abstractions that can be audibly reinforced. As one example, the object colour could be changed in order to present object information in a new manner. According to research in Gestalt laws of grouping [13] there is a stronger tendency to group local elements by common colour than by similarity of shape. This would imply that, in some cases, our brains are more receptive to the material that encompasses each shape, rather than the shape itself. Therefore, by involuntarily grouping similar coloured entities, the audience's attention would be drawn to a single contrastingly coloured object, perceiving it as being outside of the group. The performer could then take advantage of this visual phenomenon by using it to introduce a solo theme or demonstrating object-specific sonic behaviour.

In a typical physics simulation most objects will likely come to rest until excitation provides the impetus for movement. If frequent changes in data are desired then further logic can be introduced via object specific context menus (Figure 1). For example the user could define a point where rigid bodies can spawn at regular intervals. Each body created would have a lifespan where the associated object would be automatically removed from the scene after such time had elapsed. Inter-object logic can be extended further by defining connecting mechanical joints and introducing external forces such as gravitational fields. Automated mechanical contraptions would be a logical step in complexity, allowing for the creation of visual algorithms. With the basic building blocks, it should be possible to conceive and construct contraptions in the style of Heath Robinson[1] or Rube Goldberg[2], providing unfamiliarity through the extraordinary.

2.3. Mapping

In this prototype the parametric mapping process grants an insight into the composer's conceptual understanding of the data dimensions. It has been suggested that metaphors help create more intuitive mappings [14] and is well suited to parameter mapping sonification [4]. Whereas the universal laws of physics can represent a predictable visual behaviour by employing a metaphor that fits our everyday observations, the sound representation is more subjective. The mapping of the objects serves to reflect the experiences of the user, making it difficult to produce general metaphors that are valid in any context. What may be coherent and intuitive to one mindset could be judged differently by those from another cultural background. There have been attempts to create online databases [15][16] suggesting mappings based on experimental evidence although it is widely accepted that an affective mapping can't always be predicted [4][14]. A heuristic approach to this area should be adopted to allow for a compositional process that encourages experimentation in order to express creativity where the audience can reflect on the implications of a musician's cultural and physical experiences.

When interviewing scientists, Vogt and Holdrich [4] discovered that strong metaphors emerge from their professional experience. They found that more mapping associations were suggested for the well-known particles and fewer for the rarer proposing that perhaps this arose from fewer encounters, lack of interaction, and therefore less prominent in the mind. They also discovered that everyday properties such as mass were cited more often than abstract ones. This would imply that an object visually described through a recognisable metaphor, encapsulating everyday properties, can be easier to map.

Our experience with physical objects allows us to inherently determine complex data relations. Properties are implied by a rigid body's response to collisions with its surroundings. By referring to the visual behaviour during this event we perhaps reduce the need to refer to the linking of parameters for interpretation. This can be illustrated by focusing on two dimensions, such as mass and velocity, where one could map them to pitch and envelope time, respectively. We understand that an object of greater mass would provide an object of less mass with a higher velocity upon collision. A spectrum of sound can be obtained afterwards where we could assume that objects that have travelled further will differ in pitch, and duration, to those that travelled less distance over the same period of time.

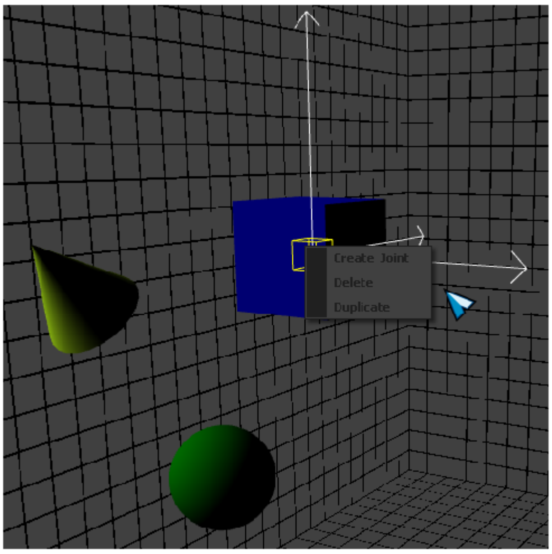


Figure 1: Three objects in the environment. The context menu is displayed over the selected object.

Understanding of the meaning in sonifications depends on the metaphors implied where the choices made during the process are crucial for how a design is understood by its listeners [17]. For instance, the coupling of coloured objects mentioned in section 2.2. Walker [18] states that in order to achieve an effective mapping choice, one must go beyond that of polarity and linear scaling functions while avoiding restrictions placed on the user through bad design [19]. The mapping window controls were devised to encourage flexibility by employing a messaging system, discussed later in this paper, to allow the user to map exposed parameters to potentially any input of a synthesiser. In conjunction with these GUI controls (Figure 2) I created a function editor that serves to display the relationship between the two dimensions. The editor itself contains two permanent breakpoints that define the input domain (x axis) and the output

[1]http://heathrobinson.org/exhibition/index.htm
[2]http://www.rubegoldberg.com/

range (y axis). Further breakpoints can be added and removed in order to construct a bijective mapping curve or polyline. The curvature of the segments, found between each breakpoint, can also be configured in order to account for both linear and non-linear responses.

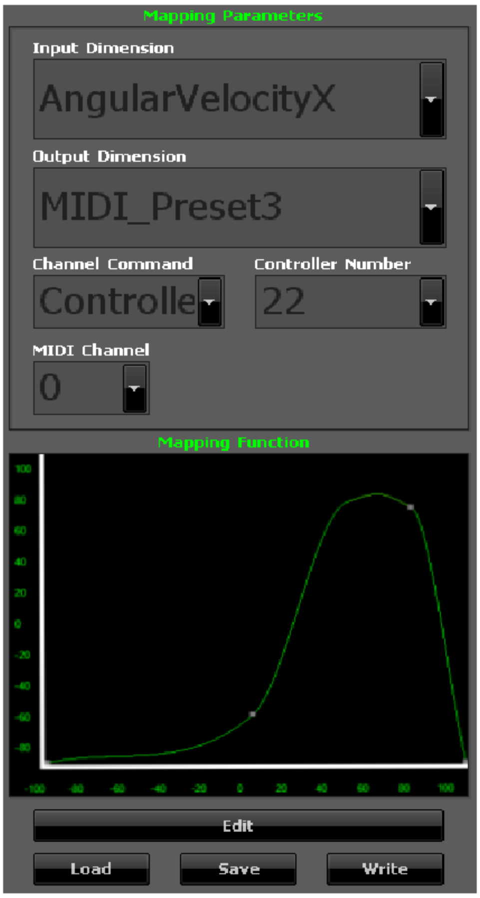


Figure 2: The mapping window along with the available controls for editing the relationship between the input and output dimensions.

2.4. Camera

Most software synthesisers use flat imagery in order to represent modifiable parameters making use of the two dimensions provided by a computer screen. Their interfaces often restrict the information to a window, presenting the intended audience with a multitude of controls that are difficult to decipher and engage with. The prototype's main interface attempts to address this problem, presenting the information in a more natural three dimensional form. This posed the question of how one might traverse the extra dimension in an immersive and inherent manner using the same display hardware.

Many first person perspective games attempt to provide immersive, high-level simulations of reality. In a typical setup, the player views, navigates and interacts with the game world by operating a camera. Recent studies have compared games in this category to sonification systems [20], where sounds are used to accentuate the player immersion by reacting to their behaviour or to provide sign posts for orienteering. I felt it was appropriate to adopt some of the ideas found in these proven systems by implementing a similar camera system for interactive traversal of the virtual environment.

Camera movement can be kinaesthetically controlled in real-time or automated along a user-defined pathway. Default manual control is that same as that of a typical PC first person shooter setup, utilising the W, A, S and D keys for camera translation, and a mouse for camera rotation. The camera rotation system required slightly different approaches to each mode of operation as manual rotation of a camera with six degrees of freedom would be disorientating with standard mouse control. If the camera pitch was allowed to be greater than 90 degrees in either direction, the mouse controls would be reversed along both the pitch and yaw axes. I therefore decided to emulate more natural head movement by restricting the camera pitch to a ±90 degree range in the same manner that a first person perspective camera does.

Node Property	Description
Position	Location of the node
Rotation/Orientation	Camera's orientation when reaching the node
Speed	The constant speed of the camera until the next node is reached *Time will be recalculated*
Time	The time (seconds) at which the camera arrives at the node *Speed will be recalculated*

Table 1: Configurable properties of an automated camera node.

Automated camera control frees the user from direct control giving them the opportunity to concentrate on other tasks, such as object interaction, and does not require the same restriction for rotation. Camera motion is defined by a series of nodes that comprise a Hermite spline-based path. The properties of each node (Table 1) allow for an increase in the accuracy and response time of the camera when compared to independent user control. Spherical linear interpolation is employed to ensure smooth changes in camera orientation when moving from one node to the next and prevents viewer disorientation. The timing of the camera can add to the overall sense of structure, guiding the viewer to focus on visual snapshots of the environment at designated points in time where precise values, for both speed and time, support various tempi.

By directing the camera, the user can create a sense of motion, guiding the audience through a visual soundscape.

Choices made in constructing the camera's pathway become part of the creative process, enabling the viewer to observe through the cognitive lens of the composer. In this manner, attention can be drawn to specific areas of interest whilst providing an insight into the structures underlying the composition. Sturm [1] touched on this particular benefit of a camera system when he stated 'thus any sonification of a particle system is dependent on the state of the observer; each observer with a unique position and/or velocity will hear the system in a different way – a truly relativistic idea.'

2.5. Messaging System

The messaging system sends and receives OSC and MIDI based messages providing the user with an opportunity to customise the data flow both in and out of the software. Whereas traditional human interface devices, such as a mouse and keyboard, can be used without the need for this system, these two protocols provide a widely accepted standard for interface control, expanding upon the breadth of possible controllers and the levels of immersion they provide. For example, tactile feedback can be introduced by utilising this system. Specific data, sent when an object is grabbed, can be interpreted by the current controller in order to produce corresponding actions such as vibrating.

With the mapping system controlling the outgoing data, I separate the audio processing from the software. Implementing these protocols grants access for communication across a network, extending the reach to computers operating on different platforms, such as OS X or Linux, or mobile platforms such as Android or iOS. This would provide the opportunity to communicate with a vast range of audio applications and synthesis tools, exploiting their existing timbres and increasing potential sonic diversity when compared to an inbuilt synthesis engine.

I have produced a video [21] that demonstrates an example of my prototype connected to sixteen instances of the Alchemy VST plug-in, utilising all available MIDI channels on a loopback network. The table shown below the video contains the mapping details for each object and was automatically generated by the prototype.

3. CONCLUSION

This work sets out to add to the range of tools for experimentation and interaction of a data set using the combination of the visual and aural modalities. It intends to expand on the experience found in Versum [8] by incorporating objects that dynamically respond to real-time interaction, with a unique collision response determined by their configurable properties. Whereas the entities found in the aforementioned software lend themselves to more ambient sounds due to their continuous playback, the objects here also allow for more dynamic sound with full control of ADSR envelopes and sonic response. Furthermore, the flexible mapping system does not constrain the amplitude of objects based solely on their distance from the camera as other relationships can be explored via the messaging system.

Whilst much time has been spent creating the basic tools for both music creation and sonification, I feel that future work should be focused on musicality. Sonifying data into a systematic musical structure to understand patterns and trends in a more traditional sense would have the benefit of making the tools more widely understood by the potential audience.

Nguyen discusses an approach to musicality in the mapping process where he decides to lose resolution of the data in favour for an increase in musicality [3]. He argues that the use of musical structure in sonification has the potential to communicate compound relationships with an increase in clarity that might not be apparent with high resolution data. To integrate this I would suggest changes to the function editor that would accommodate user-defined bands of any width. These regions would then be displayed on top of the mapping transfer function (Figure 3), allowing the user to conceive the varying resolutions. For example, on the y axis, the output dimension of pitch can be constrained to a musical scale (Lydian, Chromatic, etc.) using standard frequency tuning. This idea can be extended into other areas such as rhythm where the triggering of sound can be quantised to match common subdivisions of a bar based on the global tempo assigned. The effect should be subtle as to not lose perceived concurrency between the audio and the visual events.

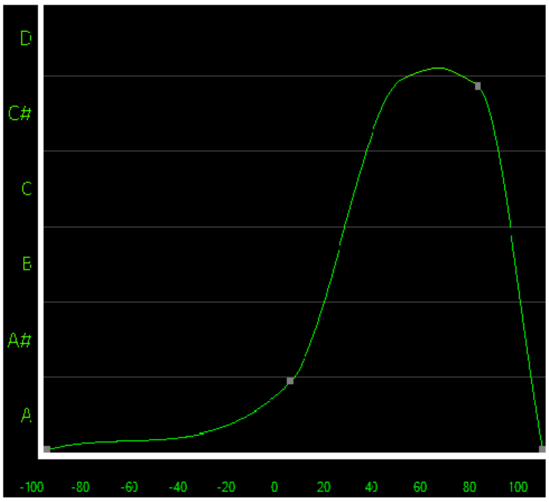


Figure 3: A function with granular regions denoted by the horizontal lines.

Recent developments in human computer interfaces, including the 'See Through 3D Desktop'[1] and the HoloDesk,[2] present more direct ways of interacting with virtual 3D objects. Model-based sonification has been shown to be intuitive by taking important dimensions of sound semantics into account and grounding them in physical sound generating processes in a natural and user-transparent way [22]. By combining projections of the simulated objects with interfaces that emulate a more innate way of interaction, we can extend the model-based method beyond its inherent physical constraints. This would benefit the interface building process as a variety of deformable, polygonal

[1]http://leejinha.com/See-Through-3D-Desktop
[2]http://research.microsoft.com/apps/video/dl.aspx?id=154571

objects could be designed, created and saved in a portable format. It would also encourage the creation of more abstract and imaginative virtual controller shapes whose physical counterparts would be difficult or impossible to implement.

4. REFERENCES

[1] B. L. Sturm, "Synthesis and Algorithmic Composition Techniques Derived from Particle Physics," in *Proceedings of the Eighth Biennial Symposium on Technology and the Arts*, Connecticut College, New London, Connecticut, 2001.

[2] M. Harris. "GPGPU: General-purpose computation on GPUs," in *Game Developers Conference*, San Francisco, California, 2005.

[3] V. Nguyen, "Tonal DisCo: Dissonance and Consonance in a Gaming Engine," in *Proceedings of the International Conference on Auditory Display*, Budapest, Hungary, 2011.

[4] K. Vogt and R. Holdrich, "A Metaphoric Sonification Method – Towards the Acoustic Standard Model of Particle Physics," in *Proceedings of the International Conference on Auditory Display*, Washington, USA, 2010.

[5] M. d'Inverno, F. Olofsson, "RedUniverse – a simple toolkit," in *Live Algorithms for Music Conference*, London, UK, 2006.

[6] S. Le Groux, J. Manzolli, P. Verschure, "VR-RoBoser: Real-Time Adaptive Sonification of Virtual Environments Based on Avatar Behavior," in *Proceedings of the Conference on New Interfaces for Musical Expression,* New York, USA, 2007.

[7] R. Perkins, "Mapping 3D Objects To Synthesised Sound Using A Simulated Physics System," in *Proceedings of the International Computer Music Conference*, Huddersfield, UK, 2011.

[8] T. Barri, "Versum: Audiovisual Composing in 3D," in *Proceedings of the International Conference on Auditory Display*, Copenhagen, Denmark, 2009.

[9] A. Hunt and T. Hermann, "The Importance of Interaction in Sonification," in *Proceedings of the International Conference on Auditory Display,* Sydney, Australia, 2004.

[10] T. Hermann, "Sonification for Exploratory Data Analysis," PhD thesis, Bielefeld University, Bielefeld, 2002.

[11] I. Bukvic and K. Ji-Sun, "µ Max-Unity3D interoperability toolkit," in *Proceedings of the International Computer Music Conference*, Montreal, Canada, 2009.

[12] G. Garnett and C. Goudeseune, "Performance Factors in Control of High-Dimensional Spaces," in *Proceedings of the International Computer Music Conference*, San Francisco, 1999.

[13] P. T. Quinlan and R. N. Wilton, "Grouping by proximity or similarity? Competition between the Gestalt principles in vision," in *Perception 27*, 1998, pp. 417–430.

[14] B. Walker and G. Kramer, "Mappings and Metaphors in Auditory Displays: An Experimental Assessment," in *ACM Transactions on Applied Perception*, vol. 2, issue 4, 2005.

[15] S. Barrass, "EarBenders: Using Stories About Listening to Design Auditory Interfaces," in *Proceedings of the First Asia-Pacific Conference on Human Computer Interaction*, Information Technology Institute, Singapore, 1996.

[16] B. Walker, and D. Lane, "Sonification Mappings Database on the Web," in *Proceedings of the International Conference on Auditory Display*, Espoo, Finland, 2001.

[17] A. de Campo, J. Rohrhuber, T. Bovermann and C. Frauenberger, "Sonification and Auditory Display in SuperCollider," in *The SuperCollider Book,* S. Wilson, D. Cottle, and N. Collins, eds., 2011, Cambridge, UK: The MIT Press, pp. 381-408.

[18] B. Walker, and D. Lane, "Psychophysical Scaling of Sonification Mappings: A Comparison of Visually Impaired and Sighted Listeners," in *Proceedings of the International Conference on Auditory Display*, Espoo, Finland, 2001.

[19] G. Kramer, et al. "Sonification Report: Status of the Field and Research Agenda," in *Report prepared for the National Science Foundation by members of the International Community for Auditory Display*, Santa Fe, NM, 1999.

[20] M. Grimshaw, "Sound and Immersion in the First-Person Shooter," in *Proceedings of the International Conference on Computer Games*, La Rochelle, France, 2007.

[21] R. Perkins, "Video demonstrating MIDI out capabilities of the prototype," Available at www.rhysperkins.com/Sonification/MIDI.htm, Accessed 8 February 2012.

[22] T. Hermann, H. Ritter, "Sound and Meaning in Auditory Data Display," in *Proceedings of the IEEE*, vol. 92, issue 4, 2004.

TOWARDS INTERACTIVE SONIFICATION BASED ON REAL-TIME PHYSICS SIMULATIONS

Rhys Perkins
Anglia Ruskin University
Department of Music and Performing Arts,
Cambridge, CB1 1PT, UK

ABSTRACT

There has been an increasing amount of research utilising 3D virtual environments as a core component of interactive sonifications. While showing considerable potential for their ability to produce both real-time visualisation and sound, they often come with constraints as a result of their design decision processes.

This paper presents developments of a prototype that has arisen out of my attempts to address some of the issues involved in bringing sonification to a wider audience through a universal metaphor. These new additions allow for an intuitive, elementary introduction into to the world of auditory display, while providing a more flexible and immersive environment for composition and sound design.

1. INTRODUCTION

The emergence of improving technology has provided an opportunity to overcome the limitations of previous work [1] where computational requirements would produce significant latency between the audio and the visual, inhibiting real-time interaction. Dedicated hardware, such as the graphics processing unit (GPU), allow for the sharing of resources and have recently shifted their focus towards more general purpose computing [2] including accelerating physics simulations. Freely distributed 3D physics engines such as Bullet and Open Dynamics Engine, common middleware solutions for modern game engines, are readily available and receive regular updates thanks in part to an increasing demand for realism amongst gamers. Their cross-platform approach sees widespread use of long established solutions such as the desktop computer but also extends to mobile devices. As a result this presents us with an opportunity to revisit previous compositional techniques [1] and reach a wider audience in the process.

Physical modelling solutions based on mass-spring systems such as CORDIS-ANIMA [3] allow for computer simulations of physical objects based on the understanding and implementation of their sound production mechanism. Abstractions of this language such as the GENESIS graphical user interface (GUI) [4] have reached the stage where it is possible to interact with the simulation and generate audio in real-time. The underlying technique for these systems relies on the physical vibrating properties of the objects they simulate resulting in the tight coupling of the mechanical and acoustical systems. My approach is to utilise a real-time classical physics simulation of rigid bodies, extract and highlight individual body parameters, and sonify them as streams of data that drive synthesis parameters. Synthesis is not determined by resonating structures themselves but by mapping any combination of the parameters that attribute to the simulated object's behaviour.

Sturm [1] laid out some of the benefits that a sonification of particle physics simulation would bring to science, including new ways of understanding physical phenomena, and refers to several artistic merits such as bending of those scientific laws to suit a composer's taste. Pedagogical advantages that Western music composition students might gain from using an audiovisual simulation are also discussed. In particular, he suggests there is an advantage to combining the audio and visual modalities in order to present musical ideas, likening it to listening to a piece whilst reading a score, rather than partaking in one or the other activity separately. Metaphorical correlations between particle physics and sound synthesis have been explored [1][5], serving as a means of providing a bridge between the two fields whilst highlighting cultural differences and the problems that might arise from them. For instance Sturm [1] states that composers must possess skills in physics to begin with and only the audience need not be versed. The thinking behind the GENESIS environment on the other hand disagrees with this and aims to let a non-physicist musician create musical material through a "physical way of thinking" [4]. It sought an efficient and natural mental model of algorithms where models are more easily internalised as representations of real objects rather than with a more mathematical approach. My software shares this design approach so that a user can construct logic and generate data through dynamic object interaction without the need for coding keywords, operators and the understanding of programming paradigms. However the system also reflects Sturm's thoughts and has been designed to accommodate several user levels, from the well-versed composer, who comprehends and wishes to explore and extend the open-source scientific algorithms, to those that would like to immediately interact and play.

I have discussed [6] developments focusing on the tool as an auditory display prototype, this paper revisits some of the processes behind that tool revealing how certain aspects compare to work within the interactive physical modelling domain.

2. KEY COMPONENTS

2.1. Human Interaction

Investigation by Hunt and Hermann [7] has shown that applying human interaction in real-world contexts to sonification can help improve interface design. They state that humans are adapted for interaction within their physical environment and making continuous use of all their senses. When we perform an action on an object we expect some kind of reaction and our perception of objects builds up over time through this interactive process. The objects found in this system adhere to the unchanging laws of physics that our neural hardware has been effectively programmed to deal with over many years of evolution. This enables the user to utilise their acquired skills in order to manipulate high-dimensional data via objects with predictable behaviour and response. The user can interact through inference relating to expectation due to the behaviour of known physical counterparts [3] reducing the gap between reality and emulation. This is turn means we can naturally envisage several physical gestures to manipulate objects and influence data streams. Interaction with the objects has a direct effect on the procedurally generated simulation in a similar manner to model-based sonification [7] where the user supplies the initial excitation. By grabbing, moving and throwing objects it is feasible to perform a wide range of actions from striking, to more delicate procedures such as plucking; each of which results in changes to the data dimensions. This direct process introduces information manipulation to the average user at a more accessible level when compared to other similar applications [8], since no coding knowledge is required, allowing us to visualise the underlying algorithms with a natural mental model [4].

Presentation of the data in this manner brings its own set of problems in that there is an increased potential to overwhelm the viewer with information. When considering high-dimensional spaces one study [9] argues for a mental model simpler than brute-force awareness of every detail to avoid cognitive overload. Parameters should be cross-coupled so that the performer naturally thinks of certain parameters as varying together in predefined patterns. The high-dimensional data, encapsulated visually by each model, allows us to intrinsically understand how the parameters vary together. Throwing an object would imply a change in velocity that would be influenced by the mass of the object. Spherical objects are more naturally inclined to roll, providing smoother changes in angular velocity as opposed to the sudden, erratic changes of their square shaped equivalents.

Through the use of certain object types the user can indirectly interact with other entities in the environment. Cadoz [10] creates an object hierarchy using leader and instrumentalist metaphors, alongside basic gestural triggers, to describe a chain of complex multi temporal events without the need for post-treatment. Sturm refers to indirect control when he suggests that with a viscous fluid, any number of mysterious forces, can be applied to a system, creating drag forces and keeping the system under, or out of, control. This example can be considered to lend itself to good instrument design and for performance. Jordà, [11] states that in a good performance, the musician needs to know and trust the instrument, to be able to push it to the extremes and to bring it back and forth out of zones of non-linearity. Simulation algorithms lend themselves to predictability, as previously discussed, but under normal simulated circumstances objects are likely to come to rest, providing stability.

As an example of a combination of the hierarchy and good performance scenarios let us consider two fundamental object behaviours in a classical physics simulation, kinematic and dynamic. Kinematic entities adhere to kinematics branch of mechanics and are not influenced by any forces acting upon them, for instance those modelled by vector fields. This is useful for objects we wish to transform through real-time interaction but should remain in a fixed state immediately after this process. Dynamic entities will respond to all influential forces in real-time and will continue to act even after interaction has finished. The configurable properties are inherent to both types. By combining these behaviours we can group objects and create our own hierarchies. If we constrain a group of dynamic objects inside a hollow kinematic container they will come to rest and remain in an state of equilibrium, or linear zone. By introducing a kinematic object that represents a force field we can now visualise and control another entity that generates an attractive, or repulsive force to excite all dynamic objects within the field's area of influence. The force generator object will not respond to force it represents, instead only responding to user input. By grabbing and moving the object the performer can influence a small number of constrained dynamic objects, brought in and out of rest depending on their containment relationship with the vector field. This demonstrates a basic object hierarchy with the user manipulating a force that can "play" a virtual instrument comprising dynamic objects that also influence one another through real-time collision events.

2.2. Objects

Objects provide a modular approach to the way the user experiences the underlying data bringing more generality to the environment in which they are used. Depending on the object's configurable physical parameters the computer will automatically simulate subsequent interactions as the object reacts to its current environment and user intervention. However, it can be argued that there are parameters that have no direct effect on the simulation which are just as important for the user to exploit. These properties can enhance a users' experience, and encourage them to learn, by creating associations through further visual abstractions that can be audibly reinforced. As one example, the colour of an object could be changed in order to present information in a new manner. According to research in Gestalt laws of grouping [12] there is a stronger tendency to group local elements by common colour than by similarity of shape. This would imply that, in some cases, our brains are more receptive to the material that encompasses each shape, rather than the shape itself. By involuntarily grouping similar coloured entities, the audience's attention would be drawn to a single contrastingly coloured object, perceiving it as being outside of the group. The performer could then take advantage of this visual phenomenon by using it to introduce a solo theme or demonstrating object-specific sonic behaviour.

In a typical physics simulation most objects will likely come to rest until excitation provides the impetus for movement. If frequent changes in data are desired then we can introduce objects that encourage movement without relying on direct user interaction or define a series of events through scripted logic configured via object specific context menus (Figure 1). By generating a script, a user can change any number of object attributes at a specific time allowing for the creation of visual algorithms or a gestural score. Cadoz [10] refers to musical structure and the process of composition, being made of assembled simpler entities that can be put together by a generative process, an automated process or a dynamic gesture. The range of objects provided by the software, along with the various means of interaction, aim to reflect those compositional methods within an automated simulation of mechanical physics.

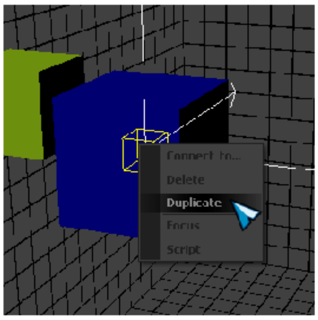


Figure 1: A selected object with its context menu

2.3. Mapping

In the prototype I describe here the parametric mapping process grants an insight into the composer's conceptual understanding of the data dimensions. It has been suggested that metaphors help create more intuitive mappings [13] and are well suited to parameter mapping sonification [5]. When discussing his musical creation "pico.. TERA" [10], Cadoz uses the metaphor of the maraca for the simulation of free particles in a box. This could be seen as a restriction of the tight coupling of the mechanical and acoustic systems only allowing for the generation of noise given the number of vibrating bodies involved. Separation of the two systems provides the means for more freedom to generate alternative acoustic signals but introduces the need for a mapping system to bridge the gap. Mapping of simulation data circumvents the need for restricting vibrating masses to an audible frequency range in order to produce an acoustic signal. Instead, scaling the object attributes means any data value can be made audible or inaudible, regardless of the object's state and behaviour. Considering this alternative approach we are presented with a visual metaphor that fits our everyday observations but the sound representation has become more subjective. That is to say the user must now decide how the data dimensions should be linked to the audio dimensions.

The mapping of the objects serves to reflect the musical decisions made by the user making it difficult to produce general metaphors that are valid in any context. Those from another cultural background could judge differently what may be coherent and intuitive. There have been attempts to create online databases [14] suggesting mappings based on experimental evidence although it is widely accepted that an effective mapping can't always be predicted [5][13]. A heuristic approach to this area should be adopted to allow for a compositional process that encourages experimentation in order to express creativity where the audience can reflect on the implications of a musician's cultural and physical experiences. After interviewing physicists at CERN [5], the authors discovered that strong metaphors emerge from their professional experience. They found that more mapping associations were suggested for the well-known particles and fewer for the rarer particles proposing that perhaps this arose from fewer encounters, lack of interaction, and therefore becoming less prominent in the mind. They also observed that everyday properties such as mass were cited more often than abstract ones. This would imply that an object visually described through a recognisable metaphor, encapsulating everyday properties, can be easier to map. Our experience with physical objects allows us to inherently determine complex data relations. Properties are implied by a rigid body's response to collisions with its surroundings. By referring to the visual behaviour during this event we perhaps reduce the need to refer to the linking of the parameters for interpretation.

Walker [15] states that in order to achieve an effective mapping choice one must go beyond that of polarity and linear scaling functions while avoiding restrictions placed on the user through bad design. The mapping window controls were devised to encourage flexibility by employing a messaging system (based on the OSC and MIDI protocols) to allow the user to map exposed parameters to potentially any input of a synthesiser and increase sonic diversity. In conjunction with these GUI controls (Figure 2) I created a function editor that serves to display the relationship between the two dimensions. The editor itself contains two permanent breakpoints that define the input domain (x axis) and the output range (y axis). Further breakpoints can be added and removed in order to construct a bijective mapping curve or polyline. The curvature of the segments can also be configured in order to account for both a linear and non-linear response.

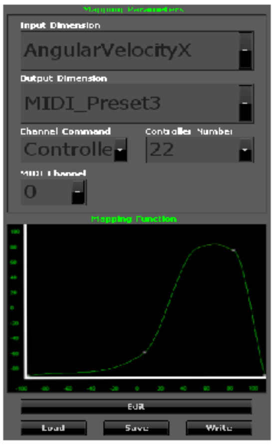


Figure 2: The mapping window

3. CONCLUSION

This work sets out to add to the range of tools for interacting, composing and experimenting with a data set using a combination of the visual and aural modalities. By utilising mainstream open source physics engines we have access to highly optimised algorithms that govern inherently recognisable objects where the data provided by the unchanging laws of classical mechanics serves as common ground for sharing and communicating ideas. Separating the audio from the mechanical system introduces new ways of interpreting data dimensions where, along with a flexible mapping system, we can choose to highlight or ignore individual streams in order to discover new ways of creating music.

4. REFERENCES

[1] Sturm, B. L. "Synthesis and Algorithmic Composition Techniques Derived from Particle Physics", *Proceedings of the Biennial Symposium on Technology and the Arts, Connecticut College*, New London, Connecticut, 2001.

[2] Harris, M. "GPGPU: General-purpose computation on GPUs", *Game Developers Conference*, San Francisco, California, 2005.

[3] Kontogeorgakopoulos, A. and Cadoz, C. "Cordis Anima Physical Modeling and Simulation System Analysis", *Proceedings of the 4th Sound and Music Computing Conference*, Lefkada, Greece, 2007.

[4] Castagne, N., Cadoz, C., Allaoui, A. and Tache, O. "G^3: Genesis Software Environment Update", *Proceedings of the International Computer Music Conference*, Montreal, Canada, 2009.

[5] Vogt, K. and Holdrich, R. "A Metaphoric Sonification Method – Towards the Acoustic Standard Model of Particle Physics", *Proceedings of the International Conference on Auditory Display*, Washington, USA, 2010.

[6] Perkins, R. "Sonification of a Real-Time Physics Simulation Within a Virtual Environment", *Proceedings of the International Conference on Auditory Display*, Atlanta, Georgia, 2012.

[7] Hunt, A and Hermann, T. "The Importance of Interaction in Sonification", *Proceedings of the International Conference on Auditory Display, Sydney*, Australia, 2004.

[8] Bukvic, I. and Ji-Sun, K. "µ Max-Unity3D interoperability toolkit", *Proceedings of the International Computer Music Conference*, Montreal, Canada, 2009.

[9] Garnett, G. and Goudeseune, C. "Performance Factors in Control of High-Dimensional Spaces", *Proceedings of the International Computer Music Conference*, San Francisco, 1999.

[10] Cadoz, C. "The Physical Model as Metaphor for Musical Creation "pico..TERA", a piece entirely generated by physical model", *Proceedings of the International Computer Music Conference*, Gothenburg, Sweden, 2002.

[11] Jordà, S. "Instruments and Players: Some thoughts on digital lutherie", *Journal of New Music Research*, 3, 2004, pp. 321-341.

[12] Quinlan, P. T. and Wilton, R. N. "Grouping by proximity or similarity? Competition between the Gestalt principles in vision", *Perception*, 27, 1998, pp. 417–430.

[13] Walker, B. and Kramer, G. "Mappings and Metaphors in Auditory Displays: An Experimental Assessment", *ACM Transactions on Applied Perception*, vol. 2, issue 4, 2005.

[14] Barrass, S. "EarBenders: Using Stories About Listening to Design Auditory Interfaces", *Proceedings of the Asia-Pacific Conference on Human Computer Interaction*, Information Technology Institute, Singapore, 1996.

[15] Walker, B. and Lane, D. "Psychophysical Scaling of Sonification Mappings: A Comparison of Visually Impaired and Sighted Listeners", *Proceedings of the International Conference on Auditory Display*, Espoo, Finland, 2001.

MAPPING 3D OBJECTS TO SYNTHESISED SOUND USING A SIMULATED PHYSICS SYSTEM

Rhys Perkins
Anglia Ruskin University
Cambridge
rhys.perkins@student.anglia.ac.uk

ABSTRACT

While the sonification of a simulated physics environment has the potential to create meaningful sound, it can be difficult to determine how the data produced should be musically interpreted. What data should be considered significant, and how can it be presented in way that allows music to be produced in a meaningful manner? I propose a prototype graphical user interface (GUI) based environment that provides a consistent approach to mapping the exposed properties of a simulated object to a unit generator (UG). This approach will allow the user to synthesise sound with a unique, customisable system using objects that can emulate, or go beyond, the limitations imposed by their physical counterparts.

1. INTRODUCTION

Freely distributed 3D physics engines such as PhysX [1], Bullet [2] and Havok [3] are commonplace on modern computing devices. Their cross-platform approach sees widespread use that not only utilises long established platforms such as PC and Mac but follows the current trend of technology with software such as Unity 3D [4] providing mobile based solutions for iOS [5] and Android [6].

The purpose of these physics engines is to simulate object collisions and interactions in a computationally simulated environment. The environment itself provides a space to which entities can be added or removed. The associated entities are then updated by the engine, each one undergoing various calculations that cause them to behave in a mathematically determined manner.

Information from these controlled objects can be readily accessed by the user. Objects generate new data each time the engine is updated; if real-time behaviour is required then this occurs at least once per frame. The calculated values obtained from each entity can then be attached to a model which is rendered to provide a visual representation of the underlying data.

In this prototype the visual mapping process has been deliberately simplified as I have chosen to focus on the audio. Here, the model only requires a world transform comprising scale, rotation and translation matrices in order to represent the numerical information in a recognisable manner. Audio mapping does not have to be constrained to such a transform and can be more complex given the increase in potentially usable data. Using the methods laid out in this paper, the interpretation of this data presents an opportunity to explore and experiment with the mapping of the modalities. In turn, this gives rise to a set of metaphors that allow me to compose new music and generate sounds using a dynamic set of rules and algorithms.

2. INTERPRETING THE DATA

With the vast quantity of data available from each object, the potential for information overload is always present. As Cook notes [7], the user must be presented with a limited number of continuously controlled parameters in order to prevent difficulties with live, real-time performance and find a balance between flexibility and usability. With that in mind I felt that I should focus on a small number of properties that had a direct involvement in determining the visual behaviour of the object. This meant that each parameter would provide visual feedback via the model connected to the object, reinforcing the established link between the two.

2.1. Numerical data

All the numerical data exposed, with the exception of mass, can be considered dynamic. This means that it will change each update providing a wealth of information which can be mapped to a UG.

Numerical Property	Range
Angular Velocity	$[(0, 0, 0), (f_{MAX}{}^{1}, f_{MAX}, f_{MAX})]$
Linear Velocity	$[(0, 0, 0), (f_{MAX}, f_{MAX}, f_{MAX})]$
Mass	$[0, f_{MAX}]$
Position/Translation	$((-e_{MAX}{}^{2}, -e_{MAX}, -e_{MAX}), (e_{MAX}, e_{MAX}, e_{MAX}))$
Rotation/Orientation	$((-\pi, \pi], (-\pi, \pi], (-\pi, \pi]))$
Scale	$(0, e_{MAX}]$

Table 1. Key object data along with possible ranges.

[1] The largest possible value of a single-precision floating-point number
[2] Maximum possible extent of the environment

However, before mapping these parameters to a UG, it is important to understand the range of possible values that each exposed property can produce. In the case of the object's rotation, the range is definite. All other properties are subject to the restrictions of the environment where sensible limitations should be imposed. For example it would not be possible to position the object outside of the simulation space and it isn't feasible to scale an object larger than the environment that contains it. The angular and linear velocities should be limited in order to prevent numerical instability in the physics engine and remain visible to the user and their audience. Objects travelling at too greater speed will move faster than the screen is capable of rendering and be difficult to track.

2.2. Non-numerical data

There are two non-numerical data descriptions that should be considered, the object's material type and its shape. Both of these attributes have a profound behavioral effect but unlike their numerical equivalents, they are not linked to the UG. The material type (Table 2) provides a method of determining the object's coefficient of restitution ("bounciness") and its coefficient of friction. It is denoted by the graphical texture applied to the model.

Material Type	Coefficient of Restitution
Glass	0.658
Metal	0.597
Rubber	0.828
Wood	0.603

Table 2. Object material types along with typical restitution coefficients.

The shape type (Table 3) describes the primitive shape of the entity, providing correct inertia tensors for accurate simulation. The model reflects the current shape by changing accordingly.

Shape Type (Primitives)
Capsule
Cone
Cube
Cylinder
Sphere

Table 3. Available primitives to define the object's shape.

3. PROVIDING THE LINK

Having decided what data should be presented, further object description is required to provide a method of controlling a UG. In order to distinguish this abstraction from the aforementioned physics entity and model combination I will refer to this type of object as a sound object.

Each sound object must be capable of sending messages, derived from the data, which can be understood by a UG language. For this prototype I have decided to send the information using the OSC [8] protocol and use the SuperCollider [9] (SC) programming language in order to interpret the data received.

Vallis et al. suggest [10] that transparency is the key to removing potential confusion over how the instrument is being performed where clear, visual clues will aid the performer play/manipulate the instrument faster. With this in mind I felt that I should avoid ambiguity by only allowing one UG instance to be allocated to a sound object at any given time creating a one to one link between the two.

3.1. Sound object types

There are currently three distinct types of sound object:

- Oscillator
- Filter
- Effect

An oscillator object controls a signal generating UG by sending the SC server messages based on its current properties. When this object collides with another object, or its surroundings, it tells the server to create a new instance of a 'Synth' class which is assigned a unique ID. This unique ID is shared between object and class instance allowing messages from the object to be sent to the correct synthesiser voice. Subsequent collisions will retrigger the existing voice, if it is already playing, adhering to the one instance per object policy. Each voice is controlled by an ADSR amplitude envelope that allows the object to keep track of its status over time. While the voice is playing, messages created from the current object data are sent to the SC server to control the UG arguments.

The filter and effect objects modify the signal produced by the oscillator UGs. They do not have any influence over the signal unless they are within a certain radius of an oscillator object. When playing, the oscillator object determines the distance of each signal modifying object in order to create a signal path. This is achieved by comparing the distance from the geometric centre of its current model to the same centre of each influential object's model. The objects are then sorted in ascending order. This calculation is performed each frame and the UGs are dynamically patched according to the signal order provided.

All sound objects have a symbol drawn on top of their material's texture (Figure 1) to represent the type of UG assigned. This allows the performer and their audience to interpret their intended sonic behaviour.

Figure 1. Symbols denoting the type of sound object

3.2. The role of the GUI

The GUI is a combination of 2D windows displayed over the 3D world in which the objects lie. The 3D space is traversed using a free roaming camera controlled by a keyboard and mouse. The mouse can also be used to select, scale, rotate and translate any object added to the environment.

There are two windows available to the user. The first window allows the user to save the current environment and object configurations or load a previously saved arrangement. The second acts as the main window of operation (Figure 2), comprising four tabs:

- Create Object
- Properties
- Sound Control
- Help

When selecting the create object tab the user will be presented with a list of sound objects organised by object type. They can select the object they wish to create and then click anywhere in the environment to create it at a world point relative to the camera position and the location of the mouse. It should be noted that the simulation can be paused/resumed at any time in order to prevent all objects from moving or functioning. This allows objects to be created/removed, setup and saved before playback. All saved setups load in a paused state to encourage slight adjustments between environment takes and rapid musical prototyping.

The properties tab displays all the previously described object data along with the mapping controls, it can be thought of as three separate sections. The first section allows the user to change all the non numerical object data. The second section exposes the various key numerical properties. The user can select/highlight any one of these values by clicking them, new values can also be entered by keyboard and the selected object will update accordingly. All sound objects store a set of mapping instructions for each of their key numerical properties and are assigned to the controls in the third section when a numerical value is selected. This final area acts as an interpreter between modalities by displaying the current mapping configuration for the numerical value highlighted in section two. The input value is restricted to the range defined by the minimum and maximum sliders. Values outside the imposed limits are clamped to the nearest equivalent in range but the object itself is unaffected by the clamped values. The input range can be assigned to a UG parameter by selecting from the elements in the combo box. The range is then scaled to the range denoted by the output min/max sliders before being sent to the UG.

Tools found on the sound control tab allow fine tuning of the selected object's behaviour. For example, users can adjust the collision sensitivity of the oscillator object making it easier/harder to trigger a synthesiser voice, or change the nature of the sound produced by altering the ADSR amplitude envelope. The influential radius of a signal modifying object can also be changed, increasing/decreasing its chance of ascendency over the surrounding signal producers.

The final tab displays instructions for all three of the aforementioned tabs along with navigational controls and details on saving/loading compositions.

4. RELATED WORK

Scaletti [11] describes sonification as the mapping of numerically represented relations in some domain under study to relations in an acoustic domain for the purpose of interpreting, understanding or communicating relations in the domain under study. Given the contrast in this relationship it should be noted that the mathematical data does not necessarily lend itself the audio. To this extent I wish to encourage the composer to interpret the exposed properties in a manner of their choosing and according to their aesthetic taste. Experimentation is the key to this process as small changes to the mapping configuration will often provide diverging sounds. For example an oscillator object with a rubber material type would retain more kinetic energy during a collision compared to an equivalent object with a metal material type. As a result, the rubber oscillator would be more likely to collide again leading to an increase in the density of triggered sounds.

When considering the 3D environment this prototype adopts a similar modular approach to the reacTable [12], allowing a full range of movement and rotation of self contained sound objects. The fundamental difference is that these objects are not completely human controlled and are subject to the laws of the virtual environment. The environment itself can also change, where static model meshes of varying shapes and sizes can be imported to provide a new space for object interaction. As such I have tried to maintain a balance by allowing full movement of objects in both a static and dynamic environment. This encourages either carefully scripted autonomous events or real-time interaction and performance.

The prototype in this paper could be considered similar to the set of tools for sonification and visualisation of dynamic systems found in RedUniverse [14]. I have chosen to build upon some of the ideas found there by moving them into the third dimension and using a dedicated commercial physics engine in the process. I consider the use of the engine a more advantageous approach that grants the opportunity to reuse the sound objects in relatively complex scenarios such as moving liquids, turbulent air streams and mechanical contraptions without the need to code the underlying physics calculations involved in such a process.

Like µ[13] this prototype allows communication between physics engines and musical programming environments. However, I have decided to use unidirectional messaging system in order to provide a solution that can be easily adapted to communicate with other UG based software and keep performance costs low.

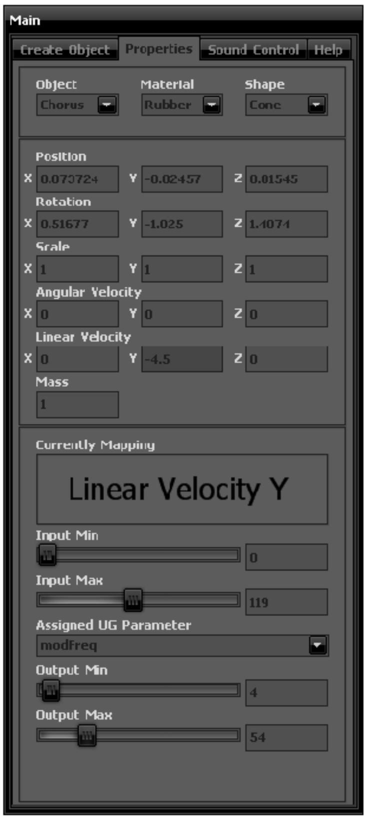


Figure 2. Main GUI window displaying the 'Properties' tab.

5. FURTHER WORK

Although the current implementation is aimed at controlling SC using OSC messages, the prototype's messaging system could be programmed to communicate with other programming languages or digital audio workstations. In the case of the latter, MIDI could be used as an alternative communication protocol where each sound object would be assigned to a separate MIDI channel.

6. REFERENCES

[1] NVIDIA Coorporation "PHYSX" http://www.nvidia.com/object/physx_new.html . (Retreived 7 January 2011). 2011

[2] "Bullet Collision Detection & Physics SDK" http://bulletphysics.com/Bullet/BulletFull/. (Retreived 7 January 2011). 2010

[3] HAVOK "Havok – Havok Physics" http://www.havok.com/index.php?page=havok-physics. (Retreived 7 January 2011). 2011

[4] Unity Technologies "Unity: Game Development Tool" http://unity3d.com. (Retrieved 7 January 2011). 2011

[5] Apple Inc. "iOS Technology Overview – Apple Developer" http://developer.apple.com/technologies/ios/. (Retreived 7 January 2011). 2011

[6] Google Inc. "Android.com" http://www.android.com/. (Retreived 7 January 2011). 2011

[7] Cook, P. "Principles for designing computer music controllers." *Proceedings of the 2001 conference on New Interfaces for Musical Expression*, Singapore, Singapore, 2001.

[8] Wright, M and Freed, A. "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers." *Proceedings of the 1997 International Computer Music Conference*, San Francisco, California, 1997.

[9] McCartney, J. "Rethinking the Computer Music Language: SuperCollider." *Computer Music Journal 26(4):61–68*, 2002.

[10] Vallis, O, Hochenbaum, J and Kapur, A. "A Shift Towards Iterative and Open-Source Design for Musical Interfaces." *Proceedings of the 2010 Conference on New Interfaces for Musical Expression,* Sydney, Australia, 2010.

[11] Scaletti, C. "Sound synthesis algorithms for auditory data representation." *Auditory Display: Sonification, Audificaction, and Auditory Interfaces*, Santa Fe, USA, 1994.

[12] Jordà, S, Geiger, G, Alonso, M and Kaltenbrunner, M. "The reacTable: Exploring the Synergy between Live Music Performance and Tabletop Tangible Interfaces." *Proceedings of the 1st international conference on Tangible and Embedded Interaction*, Funchal, Portugal, 2007.

[13] Bukvic, I. And Ji-Sun K. "µ Max-Unity3D interoperability toolkit"

[14] d'Inverno, M., Olofsson, F. "RedUniverse – a simple toolkit" *Live Algorithms for Music Conference*, London, UK, 2006.

GLOSSARY

auditory display

'auditory display encompasses all aspects of a human-machine interaction system, including the setup, speakers or headphones, modes of interaction with the display system, and any technical solution for the gathering, processing, and computing necessary to obtain sound in response to the data. In contrast, sonification is a core component of an auditory display: the technique of rendering sound in response to data and interactions.' (Hermann et al, 2011a: p. 1)

game engine

extensible software which aims to abstract common game-related tasks such as video rendering, audio, physics and input by offering reusable components.

open source software

'software that can be freely used, changed, and shared (in modified or unmodified form) by anyone.' (Open Source Initiative, 2014: para. 1)

physics engine

extensible software which provides an approximate simulation of physical systems such as rigid body dynamics. They are often utilised as a component of a game engine.

rigid body

a body which does not deform under the influence of external forces.

rigid body dynamics

the study of rigid body motion under the influence of external forces.

sonification

'the use of non-speech audio to convey information. More specifically, sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation.' (Kramer et al, 1999: p. 3)

synchresis

a concatenation of synchronism and synthesis. Used to describe 'the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time.' (Chion, 1994: p. 63)

REFERENCES

Advanced Micro Devices. (2013) *ATI Radeon™ 9700 Real-Time Legacy Demos*. [Online] Available from: http://amddevcentral.com/Resources/archive/legacydemos/pages/ATIRadeon9700Real-TimeDemos.aspx [Accessed: 1st July 2013].

Anacleto, J., Lieberman, H., Tsutsumi, M., Neris, V, Carvalho, A., Espinosa, J., Godoi M. & Zem-Mascarenhas, S. (2006) Can common sense uncover cultural differences in computer applications? *Proceedings of the International Federation for Information Processing World Computer Conference*, pp. 1-10.

AnimusicLLC. (2008) *"Pipe Dream" - Animusic.com*. [Online] Available from: http://www.youtube.com/watch?v=hyCIpKAIFyo [Accessed 10th August 2012].

Apple Inc. (2013) *iOS 6 - The world's most advanced mobile operating system*. [Online] Available from: http://www.apple.com/uk/ios/what-is/ [Accessed 1st May 2013].

Arango, F., Aziz, E., Esche, S. K. & Chassapis, C. (2008) A review of applications of computer games in education and training. *Frontiers in Education Conference*, pp. 1-6.

Arfken, G. (1985) Mapping. In: Arfken, G. (ed.). *Mathematical methods for physicists*. 3rd edition. Orlando, FL: Academic Press.

Ballas, J. A. (1994) Delivery of information through sound. In: Kramer, G. (ed.). *Auditory display: Sonification, Audification, and Auditory Interfaces*. Santa Fe Institute Studies in the Sciences of Complexity. Volume XVIII. Reading, MA: Addison Wesley Publishing Company.

Barrass, S. & Vickers, P. (2011) Sonification design and aesthetics. In: Hermann, T., Hunt, A. & Neuhoff, J. G. (eds.). *The Sonification Handbook*. Ch. 7. Berlin, Germany: Logos Publishing House, pp. 145-171.

Barrass, S. (2009) Developing the practice and theory of stream-based sonification. *Scan Journal of Media Arts Culture*, 6 (2).

Barrass, S. (1998) *Auditory information design*. Ph. D. Thesis. The Australian National University.

Barrass, S. (1996) EarBenders: using stories about listening to design auditory interfaces. *Proceedings of the First Asia-Pacific Conference on Human Computer Interaction*.

Barri, T. (2009) Versum: audiovisual composing in 3D. *Proceedings of the 15th International Conference on Auditory Display*.

Bernays, P. (1935) Platonism in mathematics. Translated from French by C. Parsons. *L'enseignement mathematique*, 34, pp. 52-69.

Berthaut, F., Katayose, H., Wakama, H., Totani, N. & Sato, Y. (2011) First person shooters as collaborative multiprocess instruments. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 44-47.

Blackwell, A. F. (2006) The reification of metaphor as a design tool. *ACM Transactions on Computer-Human Interaction*, 13 (4), pp. 490-530.

Blackwell, A. F. (2001) Pictorial representation and metaphor in visual language design. *Journal of Visual Languages and Computing*, 12 (3), pp. 223-252.

Bly, S. (1982) Presenting information in sound. *Proceedings of the 1982 Conference on Human Factors in Computing Systems*, pp. 371-375.

Boeing, A. & Bräunl, T. (2007) Evaluation of real-time physics simulation systems. *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*, pp. 281-288.

Brazil, E. & Fernström, M. (2011) Navigation of data. In: Hermann, T., Hunt, A. & Neuhoff, J. G. (eds.). *The Sonification Handbook*. Ch. 20. Berlin, Germany: Logos Publishing House, pp. 509-523.

Bregman, A. S. (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: Bradford Books, MIT Press.

Bruner, J. S. (1966) Toward a theory of instruction. Cambridge, MA: The Belknap Press of Harvard University Press.

Bukvic, I. I. & Ji-Sun, K. (2009) μ Max-Unity3D interoperability toolkit. *Proceedings of the International Computer Music Conference*, pp. 375-378.

Cadoz, C. (2009) Supra-instrumental interactions and gestures. *Journal of New Music Research*, 38 (3), pp. 215-230.

Cadoz, C. (2002) The physical model as metaphor for musical creation "pico..TERA", a piece entirely generated by physical model. *Proceedings of the International Computer Music Conference*, pp. 305-312.

Cadoz, C., Luciani, A., Florens, J. L., Castagné, N. (2003) Artistic creation and computer interactive multisensory simulation force feedback gesture transducers. *Proceedings of the 2003 Conference on New Interfaces for Musical Expression*, pp. 235-246.

Cadoz, C., Luciani, A. & Florens, J. L. (1993) Cordis-anima: a modeling and simulation system for sound image synthesis - the general formalism. *Computer Music Journal*, 17 (1), pp. 19-29.

Caoili, E. (2012) *Call of Duty remains UK's top-selling franchise for third year in a row*. [Online] Available from: http://www.gamasutra.com/view/news/39585/Call-of-Dutyremains-UKs-topselling-franchise-for-third-year-in-a-row.php [Accessed 15th June 2013].

Castagne, N., Cadoz, C., Allaoui, A. & Tache, O. (2009) G^3: GENESIS software environment update. *Proceedings of the 2009 International Computer Music Conference*, pp. 407-410.

Chadwick, A., Wingrave, B. W., Varcholik, P. D., Rose, J., Miller, A., Charbonneau, E., Bott, J. & LaViola J. J. (2010) The wiimote and beyond: spatially convenient devices for 3D user interfaces. *IEEE Computer Graphics and Applications*, 30 (2), pp. 71-85.

Chadwick, J. N. and James, D. L. (2011) Animating fire with sound. *Proceedings of ACM Transactions on Graphics*, 30 (4):84, pp. 1-8.

Cherry, E. C. (1953) Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America*, 25 (5), pp. 975-979.

Childs, E. (2007) ICAD 2006: global music - the world by ear. *Computer Music Journal*, 31 (1), pp. 95-96.

Chion, M. (1994) Audio-vision : sound on screen. Translated from French by C. Gorbman. New York: Columbia University Press.

Choi, K. S., Chan, S. H. & Pang, W. M. (2010) Virtual suturing simulation based on commodity physics engine for medical learning. *Journal of Medical Systems*, 36 (3), pp. 1781-1793.

Chowning, J. (2006) Foreword. In: Roads, C. (ed.). *The computer music tutorial*. Cambridge: The MIT Press.

Clark, D. B., Nelson, B., Sengupta, P., D'Angelo, C. M. (2009). Rethinking science learning through digital games and simulations: genres, examples, and evidence. *Proceedings of The National Academies Board on Science Education Workshop on Learning Science: Computer Games, Simulations, and Education*.

Clarke, J. M. (2005) An interactive aural approach to musical analysis. *Proceedings of the International Computer Music Conference*, pp. 85-88.

Climent, R. (2012) Hồ - a sonic expedition to vietnam. *Performed at the 7th Audio Mostly Conference: A Conference on Interaction with Sound*. [Online] Available from: http://www.ionio.gr/~amlists/FALL2012/AM12/Climent.pdf [Accessed 3rd October 2012].

Coleman, G. W., Hand, C., Macaulay, C. & Newell, A. F. (2005) Approaches to auditory interface design - lessons from computer games. *Proceedings of the 11th International Conference on Auditory Display*, pp. 99-104.

Cooley, M. (1998) Sound + image in computer-based design: learning from sound in the arts. *Proceedings of the 1998 International Conference on Auditory Display*.

Coumans, E. (2013) *Game Physics Simulation*. [Online] Available from: http://bulletphysics.org/wordpress/ [Accessed 2nd January 2013].

Creative Labs (2013) *Home - OpenAL*. [Online] Available from: http://connect.creativelabs.com/openal/default.aspx [Accessed 1st May 2013].

Cycling '74. (2012) *A SPOTLIGHT ON PHYSICS IN MAX 6*. [Online] Available from: http://cycling74.com/physics/ [Accessed 1st Dec 2012].

de Campo, A. (2009) *Science by ear. An interdisciplinary approach to sonifying scientific data*. Ph. D. Thesis. University for Music and Dramatic Arts Graz, Austria.

de Campo, A., Frauenberger, C., Höldrich, R. (2004) Designing a generalized sonification environment. *Proceedings of the International Conference on Auditory Display*.

de Paolis, L. T. (2012) Serious game for laparoscopic suturing training. *2012 Sixth International Conference on Complex, Intelligent and Software Intensive Systems*, pp. 481-485.

d'Escriván, J. & Collins, N. (2011) Musical goals, graphical lure and narrative drive: visualaudio in games. *Journal of Visual Culture*, 10 (2), pp. 238-246.

Dietze, L. (2000) Learning is living: acoustic ecology as pedagogical ground - a report on experience. Translated from German by N. Ruebsaat. *The Journal of Acoustic Ecology*, 1 (1), pp. 20-22.

Diniz, N., Demey, M. & Leman, M. (2010) An interactive framework for multilevel sonification. *Proceedings of the 3rd Interactive Sonification Workshop*, pp. 65-68.

d'Inverno, M. & Olofsson, F. (2006) RedUniverse - a simple toolkit. *Live Algorithms for Music Conference*, London, UK.

Dolphin, A. (2009a) Compositional applications of a game engine. Creative practical applications in sound art and music composition. *Games Innovations Conference*, pp. 213-222.

Dolphin, A. (2009b) MagNular: symbolic control of an external sound engine using an animated interface. *Proceedings of the 2009 International Conference on New Interfaces for Musical Expression*, pp. 159-160.

Doornbusch, P. (2002) Composers' views on mapping in algorithmic composition. *Organised Sound*, 7 (2), pp. 145-156.

Dourish, P. (2004) Where the action is: the foundations of embodied interaction. MIT Press, Cambridge, MA, USA.

Evans, K. K. & Treisman, A. (2010) Natural cross-modal mappings between visual and auditory features. *Journal of* Vision, 10 (1):6, pp. 1-12.

Face Punch Studios LTD (2012) *Garry's Mod*. [Online] Available from: garrysmod.com [Accessed 11th November 2012].

Fels, S., Gadd, A. & Mulder, A. (2002) Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, 7 (2), pp. 109-126.

Fiebrink, R., Wang, G. & Cook, P. R. (2007) Don't forget the laptop: using native input capabilities for expressive musical control. *Proceedings of the 7th international conference on New Interfaces for Musical Expression*, pp. 164-167.

Flowers, J. H. (2005) Thirteen years of reflection on auditory graphing: promises, pitfalls, potential new directions. *Proceedings of the 11th meeting of the International Conference on Auditory Display*, pp. 1-4.

Forbus, K. D. (1997) Using qualitative physics to create articulate educational software. *IEEE Expert*, 12 (3), pp. 32-41.

Frauenberger, C. (2009) Auditory display design. An investigation of a design pattern approach. Ph. D. Thesis. Queen Mary University of London.

Fraietta, A. (2008) Open sound control: constraints and limitations. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 163-167.

García, A. S., Olivas, A., Molina, J. P., Martínez, J., González, P. & Martínez, D. (2013) An evaluation of targeting accuracy in immersive first-person shooters comparing different tracking approaches and mapping models. *Journal of Universal Computer Science*, 19 (8), pp. 1086-1104.

Garnett, G. & Goudeseune, C. (1999) Performance factors in control of high-dimensional spaces. *Proceedings of the International Computer Music Conference*, pp. 268-271.

Gaver, W. W. (1994) Using and creating auditory icons. In: Kramer, G. (ed.). *Auditory display: Sonification, Audification, and Auditory Interfaces*. Santa Fe Institute Studies in the Sciences of Complexity. Volume XVIII. Reading, MA: Addison Wesley Publishing Company, pp. 417-446.

Gaver, W. W. (1993) How do we hear in the world? Explorations in ecological acoustics. *Ecological Psychology*, 5 (4), pp. 285-313.

Geijtenbeek, T. and Pronost, N. (2012) Interactive character animation using simulated physics: a state-of-the-art review. *Computer Graphics Forum*, 31 (8), pp. 2492-2515.

Gibson, J. J. (1977) The theory of affordances. In: Shaw, R. & Bransford, J. (eds.). *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Hoboken, NJ: John Wiley & Sons Inc.

Gresham-Lancaster, S. (2012) Relationships of sonification to music and sound art. *AI & Society*, 27 (2), pp. 207-212.

Grimshaw, M. (2007) Sound and immersion in the first-person shooter. *International Conference on Computer Games*, pp. 119-124.

Grond, F. & Berger, J. (2011) Parameter mapping sonification. In: Hermann, T., Hunt, A. & Neuhoff, J. G. (eds.). *The Sonification Handbook*. Ch. 15. Berlin, Germany: Logos Publishing House, pp. 363-397.

Haartsen, J. C. (2000) The bluetooth radio system. *IEEE Personal Communications*, 7 (1), pp. 28-36.

Hamilton, R. (2008) Q3OSC or: how I learned to stop worrying and love the ~~bomb~~ game. *Proceedings of the International Computer Music Conference*.

Harada, T. (2011) A parallel constraint solver for a rigid body simulation. *SIGGRAPH Asia 2011 Sketches*, Article 22, pp. 1-2.

Henry, C. (2004) Physical modeling for pure data (pmpd) and real time interaction with an audio synthesis. *Proceedings of the Sound and Music Computing Conference*, pp. 47-49.

Hermann, T. (2011) Model-based sonification. In: Hermann, T., Hunt, A. & Neuhoff, J. G. (eds.). *The Sonification Handbook*. Ch. 16. Berlin, Germany: Logos Publishing House, pp. 399-427.

Hermann, T. (2008) Taxonomy and Definitions for Sonification and Auditory Display. *Proceedings of the 14th International Conference on Auditory Display*, pp. 1-8.

Hermann, T. (2002) *Sonification for exploratory data analysis*. Ph. D. Thesis. Bielefeld University, Germany.

Hermann, T., Hunt, A. & Neuhoff, J. G. (2011a) Introduction. In: Hermann, T., Hunt, A. & Neuhoff, J. G. (eds.). *The Sonification Handbook*. Ch. 1. Berlin, Germany: Logos Publishing House, pp. 1-6.

Hermann, T., Hunt. A. & Neuhoff, J. G. (eds.) (2011b) *The Sonification Handbook*. Berlin, Germany: Logos Publishing House.

Hermann, T. & Hunt, A. (2005) Guest editors' introduction: An introduction to interactive sonification. *IEEE Multimedia*, 12, (2), pp. 20-24.

Hermann, T. & Hunt, A. (2009) *Proceedings of the Interactive Sonification Workshops*. [Online] Available from: www.interactive-sonification.org/proceedings/ [Accessed 15th July 2013].

Hermann, T. & Ritter, H. (2004) Sound and meaning in auditory data display. *Proceedings of the IEEE*, 92 (4), pp. 730-741.

Hermann, T. & Ritter, H. (1999) Listen to your data: model-based sonification for data analysis. *International Institute for Advanced Studies in System research and cybernetics*, pp. 189-194.

Hilliges, O., Kim, D., Izadi, S., Weiss, M., Wilson, A. D. (2012) HoloDesk: direct 3D interactions with a situated see-through display. *Proceedings of the 2012 ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 2421-2430.

Hirsch, E. S. (1996) *The Block Book*. Third Edition. Washington, DC: National Association for the Education of Young Children.

Howard, D. M. & Rimell, S. (2004) Real-time gesture-controlled physical modelling music synthesis with tactile feedback. *Journal on Applied Signal Processing*, 7, pp. 1001-1006.

Hunt, A. & Hermann, T. (2011) Interactive sonification. In: Hermann, T., Hunt, A. & Neuhoff, J. G. (eds.). *The Sonification Handbook*. Ch. 11. Berlin, Germany: Logos Publishing House, pp. 273-298.

Hunt, A. & Hermann, T. (2004) The importance of interaction in sonification. *Proceedings of the International Conference on Auditory Display*, pp. 1-8.

Hunt, A. & Wanderley, M. M. (2002) Mapping performer parameters to synthesis engines. *Organised Sound*, 7 (2), pp. 97-108.

Hunt, A., Kirk, R., Orton, R. & Merrison, B. (1998) A generic model for compositional approaches to audiovisual media. *Organised Sound*, 3 (3), pp. 199-209.

Hutchison, K. (1993) Is classical mechanics really time-reversible and deterministic? *The British Journal for the Philosophy of Science*, 44 (2), pp. 307-323.

Jacob, R. J. K., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008) Reality-based Interaction: a framework for post-WIMP interfaces. *Proceedings of the ACM CHI Conference on Human Factors in Computing* Systems, pp. 201-210.

Jensenius, A. R. (2007) *Action-sound: developing methods and tools to study music-related body movement.* Ph. D. Thesis. University of Oslo, Norway.

Johnson, M. (1987) *The body in the mind: the bodily basis of meaning, imagination, and reason*. Chicago, IL: University of Chicago Press.

Jordà, S., Geiger, G., Alonso, M. & Kaltenbrunner, M. (2007) The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. *Proceedings of the First International Conference on Tangible and Embedded Interaction*, pp. 139-146.

Jordà, S. (2005) *Digital lutherie: crafting musical computers for new musics' performance and improvisation*. Ph. D. Thesis. Pompeu Fabra University, Spain.

Jordà, S. (2004) Instruments and players: some thoughts on digital lutherie. *Journal of New Music Research*, 33 (3), pp. 321-341.

Joselli, M. & Pagliosa, P. (2008) A new physics engine with automatic process distribution between CPU-GPU. *Proceedings of the 2008 ACM SIGGRAPH symposium on video games*, pp. 149-156.

Kaper, H. G., Wiebel, E. & Tipei, S. (1999) Data sonification and sound visualization. *Computing in Science and Engineering*, 1 (4), pp. 48-58.

Kaper, H. G. & Tipei, S. (1998) Manifold compositions, music visualization, and scientific sonification in an immersive virtual-reality environment. *Proceedings of the International Computer Music Conference*, pp. 339-405.

Keislar, D. (2009) A Historical View of Computer Music Technology. In: Dean, R.T. (ed.). *The Oxford Handbook of COMPUTER MUSIC*. Ch. 2. Oxford: Oxford University Press.

Kiefer, C., Collins, N. & Fitzpatrick, G. (2008) Evaluating the wiimote as a musical controller. *Proceedings of the International Computer Music Conference*, pp. 92-95.

Kojs, J. (2009) The language of action and cyberaction-based music: theory and practice. *Journal of New Music Research*, 38 (3), pp. 285-294.

Kojs, J., Serafin, S. & Chafe, C. (2007) Cyberinstruments via physical modeling synthesis: compositional applications. *Leonardo Music Journal*, 17, pp. 61-66.

Kontogeorgakopoulos, A. & Cadoz, C. (2007) Cordis anima physical modeling and simulation system analysis. *Proceedings of the 4th Sound and Music Computing Conference*, pp. 275-282.

Kramer, G. (1994a) An introduction to auditory display. In: Kramer, G. (ed.). *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Reading MA: Addison-Wesley, pp. 1-77.

Kramer, G. (ed.) (1994b) Auditory display - sonification, audification, and auditory interfaces. Reading, MA: Addison-Wesley.

Kramer, G. (1994c) Some organizing principles for representing data with sound. In: Kramer, G. (ed.). *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Reading MA: Addison-Wesley, pp. 185-222.

Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N., Neuhoff, J., Bargar, R., Barrass, S., Berger, J., Evreinov, G., Fitch, W., Gröhn, M., Handel, S., Kaper, H., Levkowitz, H., Lodha, S., Shinn-Cunningham, B., Simoni, M. & Tipei, S. (1999) *Sonification Report: Status of the Field and Research Agenda.* [Online] Available from: http://sonify.psych.gatech.edu/publications/pdfs/1999-nsf-report.pdf [Accessed 19th September 2011].

Lakoff, G. & Johnson, M. (1980) *Metaphors we live by*. Chicago, IL: The University of Chicago Press.

Le Groux, S., Manzolli, J., Verschure, P. F. M. J. (2007) VR-roboser: real-time adaptive sonification of virtual environments based on avatar behavior. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 371-374.

Levitin, D. J., McAdams, S. & Adams, R. L. (2002) Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound,* 7 (2), pp. 171-189.

Liebe, M. (2013) Interactivity and music in computer games. In: Moormann, P. (ed.). *Music and Game: Perspectives on a Popular Alliance*, pp. 41-62.

Lloyd, D. B., Raghuvanshi, N. & Govindaraju, N. K. (2011) Sound synthesis for impact sounds in video games. *Symposium on Interactive 3D Graphics and Games*, pp. 55-62.

Lokki, T. & Gröhn, M. (2005) Navigation with auditory cues in a virtual environment. *IEEE Multimedia*, 12 (2), pp. 80-86.

Loy, G. (1985) Musicians make a standard: the MIDI phenomenon. *Computer Music Journal*, 9 (4), pp. 8-26.

Ma, X., Fellbaum, C. & Cook, P. R. (2010) Environmental sounds as concept carriers for communication. *Proceedings of the 16th International Conference on Auditory Display*, pp. 353-360.

McCartney, J. (2002) Rethinking the computer music language: supercollider. *Computer Music Journal*, 26 (4), pp. 61-68.

McGookin, D. K. & Brewster, S. A. (2004) Understanding concurrent earcons: applying auditory scene analysis principles to concurrent earcon recognition. *ACM Transactions on Applied Perceptions*, 1 (2), pp. 130-155.

Menzies, D. (2010) Phya and virtual foley, physically motivated environmental sound synthesis for virtual worlds. *EURASIP Journal on Audio, Speech and Music Processing*, pp. 1-11.

Menzies, D. (2008) Virtual intimacy: phya as an instrument. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 71-76.

Mesit, J., Guha, R. & Chaudhry, S. (2007) 3D soft body simulation using mass-spring system with internal pressure force and simplified implicit integration. *Journal of Computers*, 2 (8), pp. 34-43.

Microsoft. (2013a) *.NET Downloads, Developer Resources & Case Studies*. [Online] http://www.microsoft.com/net [Accessed: 4th September 2013].

Microsoft. (2013b) *XNA Framework Class Library*. [Online] http://msdn.microsoft.com/en-us/library/bb203940.aspx [Accessed: 4th September 2013].

Microsoft. (2013c) *Visual Studio - Home*. [Online] Available from: http://www.visualstudio.com/ [Accessed: 4th September 2013].

Millington, I. (2007) *Game Physics: Engine Development*. Series in Interactive 3D Technology. San Francisco, CA: Morgan Kaufmann, p. xx.

Minsky, M., Singh, P. & Sloman, A. (2004) The St. Thomas common sense symposium: designing architectures for human-level intelligence. *AI Magazine*, 25 (2), pp. 113-124.

Moody, N., Fells, N. & Bailey, N. (2006) Motion as the connection between audio and visuals. *Proceedings of the International Computer Music Conference*, pp. 405-408.

Moore, B. C. J. (1997) *An Introduction to the Psychology of Hearing*. 4th Edition. Orlando, FL: Academic Press.

Moore, F. R. (1988) The dysfunctions of MIDI. *Computer Music Journal*, 12 (1), pp. 19-28.

Moss, W., Yeh, H., Hong, J., Lin, M. C. & Manocha, D. (2010) Sounding liquids: automatic sound synthesis from fluid simulation. *ACM Transactions on Graphics*, 29 (3):21 pp. 1-13.

Native Instruments GmbH. (2013) *Komplete : Synths & Samplers : Reaktor 5 | Products*. [Online] http://www.native-instruments.com/#/en/products/producer/reaktor-5/ [Accessed 18th January 2013].

Neidhardt, A. & Rüppel, A. (2012) Multiplayer audio-only game: pong on a massive multichannel loudspeaker system. *Proceedings of the 7th Audio Mostly Conference: A Conference on Interaction with Sound*, pp. 130-134.

Neuhoff, J. G., Wayand, J. & Knight, R. (2002) Pitch change, sonification, and musical expertise: which way is up? *Proceedings of the 2002 International Conference on Auditory Display*, pp. 1-6.

New Interfaces for Musical Expression (2013) *New Interfaces for Musical Expression | Proceedings*. [Online] Available from: www.nime.org/archive/ [Accessed 1st May 2013].

Nguyen, V. X. (2012) Circosonic: a sonification of circos, a circular graph of table data. *Proceedings of the International Conference on Auditory Display*, pp. 105-112.

Nguyen, V. X. (2011) Tonal disco: dissonance and consonance in a gaming engine. *Proceedings of the International Conference on Auditory Display*.

Nickolls, J. & Dally W. J. (2010) The GPU computing era. *IEEE Micro*, 30 (2), pp. 56-69.

Nicolas, C. & Cadoz, C. (2002) Creating music by means of 'physical thinking': the musician oriented genesis environment. *Proceedings of the 5th International Conference on Digital Audio Effects*, pp. 169-174.

Niinimäki, M. & Tahiroğlu, K. (2012) AHNE : a novel interface for spatial interaction. *Extended Abstracts of the 2012 ACM Annual Conference on Human Factors in Computing Systems*, pp. 1031-1034.

Nintendo. (2012) *Nintendo.com - Electroplankton Hanenbow - Game Info*. [Online] Available from: http://www.nintendo.com/games/detail/B4QgT8uckp1VMV3wlDQwPSWeFA1lBbr7 [Accessed 1st December 2012].

Nordby, R. (2013) *BEPUphysics - Home*. [Online] Available from: http://bepuphysics.codeplex.com/ [Accessed: 3rd September 2013].

O'Brien, J. F., Shen, C. & Gatchalian, C. M. (2002) Synthesizing sounds from rigid-body simulations. *ACM SIGGRAPH Symposium on Computer Animation*, pp. 175-181.

Ondrejka, C. (2004) Escaping the gilded cage: user created content and building the metaverse. *New York Law School Law Review*, 49 (1), pp. 81-101.

Open Source Initiative. (2014) *The Open Source Initiative*. [Online] Available from: http://www.opensource.org [Accessed 1st Feb 2014].

Ousterhout, J. K. & Jones, K. (2009) *Tcl and the Tk Toolkit*. Second Edition. Addison-Wesley Publishing Company.

Pereverzev, S. V., Loshak, A., Backhaus, S., Davis, J.C. & Packard, R. E. (1997) Quantum oscillations between two weakly coupled reservoirs of superfluid ^{3}He. *Nature*, 388, pp. 449-451.

Perkins, R. (2012a) Sonification of a real-time physics simulation within a virtual environment. *Proceedings of the International Conference on Auditory Display*, pp. 99-104.

Perkins, R. (2012b) Towards interactive sonification based on real-time physics simulations. *Proceedings of the International Computer Music Conference*, pp. 132-135.

Perkins, R. (2011) Mapping 3D objects to synthesised sound using a simulated physics system. *Proceedings of the International Computer Music Conference*, pp. 292-295.

Pirhonen, A. (2005) To simulate or stimulate? In search of the power of metaphor in design. In: Pirhonen, A., Isomäki, H., Roast, C., Saariluoma, P. (eds.). *Future Interaction Design*, London: Springer Verlag, pp. 105-123.

Puckette, M. (2002) Max at seventeen. *Computer Music Journal*, 26 (4), pp. 31-43.

Puckette, M. (1996) Pure data: another integrated computer music environment. *Proceedings of the Second Intercollege Computer Music Concerts*, pp. 37-41.

Pulkki, V. (1997) Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, 45 (6), pp. 456-466.

Quinlan, P. T. & Wilton, R. N. (1998) Grouping by proximity or similarity? Competition between the gestalt principles in vision. *Perception*, 27, pp. 417-430.

Roads, C. (1996) The musician's interface. In: Roads, C. (ed.). *The Computer Music Tutorial*. The MIT Press: Cambridge, pp. 611-911.

Robertson, P.K. (1990) A methodology for scientific data visualization: choosing representations based on a natural scene paradigm. *Proceedings of the First IEEE Conference on Visualization*, pp. 114-123.

Sanford, L. (2013) *C# MIDI Toolkit - CodeProject*. [Online] Available from: http://www.codeproject.com/Articles/6228/C-MIDI-Toolkit [Accessed: 1st September 2013].

Saue, S. (2000) A model for interaction in exploratory sonification displays. *Proceedings of the International Conference on Auditory Display*, pp. 105-110.

Scaletti, C. (1994) Sound synthesis algorithms for auditory data representation. In: Kramer, G. (ed.). *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Reading MA: Addison-Wesley, pp. 223-252.

Schertenleib, A. & Barrass, S. (2010) A social platform for information sonification: many-ears.com. *Proceedings of the International Conference on Auditory Display*, pp. 295-299.

Schroeder, B., Ainger, M. & Parent, R. (2011) The visible string, a direct audiovisual model. *Proceedings of the International Computer Music Conference*, pp. 328-331.

Serafin, S., Franinović, K., Hermann, T., Guillaume, L., Rinott, M. & Rocchesso, D. (2011) Sonic interaction design. In: Hermann, T., Hunt, A. & Neuhoff, J. G. (eds.). *The Sonification Handbook*. Ch. 5. Logos Publishing House, Berlin, Germany, pp. 87-110.

Shane, T. (2013) *Neoforce Controls - Home*. [Online] Available from: http://neoforce.codeplex.com/ [Accessed: 1st May 2013].

Shneiderman, B. (1983) Direct manipulation: a step beyond programming languages. *IEEE Computer*, 16 (8), pp. 57-69.

Silver, J., Rosenbaum, E. & Shaw, D. (2012) Makey makey: improvising tangible and nature-based user interfaces. *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, pp. 367-370.

Simian Squared Ltd. (2012) *Physynth*. [Online] Available from: http://simiansquared.com/physynth/ [Accessed 1st Dec 2012].

Sisu Devices (2011) *Intel Pipe Dream Demo (IDF 2011)*. [Online] Available from: http://www.sisudevices.com/projects/Intel-Pipe-Dream-Demo-IDF-2011/ [Accessed 1st March 2012].

Smith, R. (2006) *Open Dynamics Engine - home*. [Online] Available from: http://www.ode.org/ [Accessed 2nd January 2013].

Snook, K. & Barri, T. (2010) Versum: data sonification and visualisation in 3D. *Proceedings of the International Conference on Auditory Display*, p. 261.

Spence, C. (2007) Audiovisual multisensory integration. *Acoustical Science and Technology*, 28 (2), pp. 61-70.

Stahl, F. A. (1987) Physics as metaphor and vice versa. *Leonardo*. 20 (1), pp. 57-64.

Steinberg Media Technologies GmbH. (2013) *Technologies*. [Online] Available from: http://www.steinberg.net/en/company/technologies.html [Accessed: 1st September 2013].

Straebel, V. (2010) The sonification metaphor in instrumental music and sonification's romantic implications. *Proceedings of the International Conference on Auditory Display*, pp. 287-294.

Sturm, B. L. (2001a) Composing for an ensemble of atoms: the metamorphosis of scientific experiment into music. *Organised Sound*, 6 (2), pp. 131-145.

Sturm, B. L. (2001b) Synthesis and algorithmic composition techniques derived from particle physics. *Proceedings of the 8th Biennial Arts and Technology Symposium*.

Supper, A. (2012) *Lobbying for the ear: the public fascination with and academic legitimacy of the sonification of scientific data*. Ph. D. Thesis. Maastricht University.

Tache, O. & Cadoz, C. (2009) Organizing mass-interaction physical models: the cordis-anima musical instrumentarium. *Proceedings of the International Computer Music Conference*, pp. 411-414.

Tache, O. (2008) *Conception d'un instrumentarium pour la création musicale à l'aide des modèles physiques CORDIS-ANIMA*. Ph. D. Thesis. Institut National Polytechnique de Grenoble.

Tippach, M. (2013) *ASIO4ALL - Universal ASIO Driver For WDM Audio*. [Online] Available from: http://www.asio4all.com/ [Accessed: 2nd September 2013].

Turing, A. M. (1950) Computing machinery and intelligence. *Mind*, 59 (236), pp. 433-460.

Unity Technologies (2013) *Unity - Game Engine*. [Online] Available from: http://unity3d.com/ [Accessed 3rd January 2013].

van den Doel, K. Kry, P. G. & Pai D. K. (2001) Foleyautomatic: physically-based sound effects for interactive simulation and animation. *Proceedings of the conference for Special Interest Group on Computer Graphics and Interactive Techniques*, pp. 537-544.

Varcholik, P. (2013) *Bespoke.Osc - Bitbucket*. [Online] Available from: https://bitbucket.org/pvarcholik/bespoke.osc [Accessed: 1st September 2013].

Vaughn-Nichols, S. J. (2009) Vendors draw up a new graphics-hardware approach. *Computer*, 42 (5), pp. 11-13.

Vickers, P. (2005) Whither and wherefore the auditory graph? Abstractions and æsthetics in auditory and sonified graphs. *Proceedings of the Eleventh Meeting of the 11th International Conference on Auditory Display*.

Vickers, P. & Hogg, B. (2006) Sonification absraite/sonification concrète: an 'æsthetic perspective space' for classifying auditory displays in the ars musica domain. *Proceedings of the 12th International Conference on Auditory Display*, pp. 210-216.

Vogt, K. (2010) Sonification of simulations in computational physics. Ph. D. Thesis. University of Graz, Austria.

Vogt, K. & Höldrich, R. (2010) A metaphoric sonification method - towards the acoustic standard model of particle physics. *Proceedings of the International Conference on Auditory Display*, pp. 271-277.

Walker, B. N. & Kramer, G. (2005) Mappings and metaphors in auditory displays: an experimental assessment. *ACM Transactions on Applied Perception*, 2 (4), pp. 407-412.

Walker, B. N. & Lane, D. (2001) Sonification mappings database on the web. *Proceedings of the International Conference on Auditory Display*, p. 281.

Walker, B. N. & Nees, M. A. (2011) Theory of sonification. In: Hermann, T., Hunt, A. & Neuhoff, J. G. (eds.). *The Sonification Handbook*. Ch. 2. Berlin, Germany: Logos Publishing House, pp. 9-39.

Walker, B. N. & Nees, M. A. (2005) An agenda for research and development of multimodal graphs. *Proceedings of the International Conference on Auditory Display*, pp. 428-432.

Walker, R. (1987) The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics*, 42 (5), pp. 491-502.

Wang, G., Oh, J., Salazar, S. & Hamilton, R. (2011) World stage: a crowdsourcing paradigm for social/mobile music. *Proceedings of the International Computer Music Conference*, pp. 407-414.

Wang, G., Bryan, N., Oh, J. & Hamilton, R. (2009) Stanford laptop orchestra (SLOrk). *Proceedings of the International Computer Music Conference*, pp. 505-508.

Weichert, F., Bachmann, D., Rudak, B. & Fisseler, D. (2013) Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, 13 (15), pp. 6380-6393.

Weinberg, G. & Thatcher, T. (2006) Interactive Sonification: aesthetics, functionality and performance. *Leonardo Music Journal,* 16, pp. 9-12.

White, G. R., Fitzpatrick, G. & McAllister, G. (2008) Toward accessible 3D virtual environments for the blind and visually impaired. *Proceedings of the international conference on Digital Interactive Media in Entertainment and Arts*, pp. 134-141.

Worrall, D. (2011) A method for developing an improved mapping model for data sonification. *Proceedings of the International Conference on Auditory Display*.

Worrall, D. (2010) Parameter mapping sonic articulation and the perceiving body. *Proceedings of the International Conference on Auditory Display*, pp. 207-214.

Worrall, D. (2009a) An Introduction to Data Sonification. In: Dean, R.T. (ed.). *The Oxford Handbook of COMPUTER MUSIC*. Ch. 16. Oxford: Oxford University Press.

Worrall, D. (2009b) *Sonification and Information Concepts Instruments and Techniques*. Ph. D. Thesis. University of Canberra, Australia.

Wright, M. & Freed, A. (1997) Open sound control: a new protocol for communication with sound synthesizers. *Proceedings of the International Computer Music Conference*, pp. 101-104.

Xamarin. (2013) *Mono*. [Online] Available from: http://mono-project.com/Main_Page [Accessed: 1st September 2013].

Young, S. (2013) *Why we have checkpoint saves | The Escapist*. [Online] Available from: http://www.escapistmagazine.com/articles/view/columns/experienced-points/10301-Why-We-Have-Checkpoint-Saves [Accessed: 23rd August 2013].

BIBLIOGRAPHY

Assarsson, U. & Möller, T. (1999) Optimized view frustum culling algorithms. *Technical Report 99-3*, Sweden: Chalmers University of Technology.

de Campo, A., Rohrhuber, J., Bovermann, T. & Frauenberger, C. (2011) Sonification and Auditory Display in SuperCollider. In: Wilson, S., Cottle, D. and Collins, N. (eds.). *The SuperCollider Book*. Ch. 13. Cambridge: The MIT Press.

DeLoura , M. (ed.) (2000) *Game Programming Gems*. Boston, MA: Charles River Media.

dos Santos, R. P. (2009) Second life physics: virtual, real or surreal? *Journal of Virtual Worlds Research*, 2 (1).

Dunn, F. & Parberry, I. (2002) *3D Math Primer for Graphics and Game Development*. Sudbury, MA: Wordware Publishing.

Eisner, E. W. (1997) The promise and perils of alternative forms of data representation. *Educational Researcher*, 26 (6), pp. 4-10.

Ericson, C. (2005) *Real Time Collision Detection*. The Morgan Kaufmann Series in Interactive 3D Technology. San Francisco, CA: Morgan Kaufmann.

McIntyre, M. E. (1997) Lucidity and science II: from acausality illusions and free will to final theories, mathematics, and music. *Interdisciplinary Science Reviews*, 22 (4), pp. 285-303.

Möring, S. (2012) Tackling the metaphor-simulation dilemma. *Proceedings of DiGRA Nordic 2012 Conference: Local and Global - Games in Culture and Society*.

Myers, B. A. (1998) A brief history of human computer interaction technology. *ACM interactions*, 5 (2), pp. 44-54.

Schneider, P. J. & Eberly, H. D. (2003) *Geometric Tools for Computer Graphics*. San Francisco, CA: Morgan Kaufmann.