Ontology of Music Performance Variation

John R. Taylor
Sydney Conservatorium of Music
University of Sydney

A thesis submitted in fulfilment of
requirements for the degree of
PhilosophiæDoctor (PhD)

2014

Declaration

I declare that the research presented here is my own original work and has not been submitted to any other institution for the award of a degree.

Signed:...

Date:...

Abstract

Performance variation in rhythm determines the extent that humans perceive and feel the effect of rhythmic pulsation and music in general. In many cases, these rhythmic variations can be linked to percussive performance. Such percussive performance variations are often absent in current percussive rhythmic models. The purpose of this thesis is to present an interactive computer model, called the PD-103, that simulates the micro-variations in human percussive performance. This thesis makes three main contributions to existing knowledge: firstly, by formalising a new method for modelling percussive performance; secondly, by developing a new compositional software tool called the PD-103 that models human percussive performance, and finally, by creating a portfolio of different musical styles to demonstrate the capabilities of the software. A large database of recorded samples are classified into zones based upon the vibrational characteristics of the instruments, to model timbral variation in human percussive performance. The degree of timbral variation is governed by principles of biomechanics and human percussive performance. A fuzzy logic algorithm is applied to analyse current and first-order sample selection in order to formulate an ontological description of music performance variation. Asynchrony values were extracted from recorded performances of three different performance skill levels to create "timing fingerprints" which characterise unique features to each percussionist. The PD-103 uses real performance timing data to determine asynchrony values for each synthesised note. The spectral content of the sample database forms a three-dimensional loudness/timbre space, intersecting instrumental behaviour with music composition. The reparameterisation of the sample database, following the analysis of loudness, spectral flatness, and spectral centroid, provides an opportunity to explore the timbral variations inherent in percussion instruments, to creatively explore dimensions of timbre. The PD-103 was used to create a music portfolio exploring different rhythmic possibilities with a focus on meso-periodic rhythms common to parts of West Africa, jazz drumming, and electroacoustic music. The portfolio also includes new timbral percussive works based on spectral features and demonstrates the central aim of this thesis, which is the creation of a new compositional software tool that integrates human percussive performance and subsequently extends this model to different genres of music.

KEYWORDS: music, computer, electronic, composition, micro-timbre, micro-timing, performance model, electroacoustic, spectral, percussion, jazz drums, software.

I would like to dedicate this thesis to my father who unexpectedly passed away in October 2012. His support, encouragement, inspiration, and sense of humour will always stay with me.

I would like to dedicate the composition portfolio to Auntie Eileen and Uncle Ted, who passed away in 2010 and 2011. Ee bah gum, I finished it Gelborsh!

Finally, I would like to dedicate the PD-103 software to my sons Max and Sam, for their friendship, understanding, and for keeping me laughing.

Acknowledgements

I would like to acknowledge my friends and family in the U.K., Australia and Spain, particularly my mother, Lesley Taylor, my brother Scott, my partner in crime Terri, and my Grandma June, for their love and support over the past few years. I would like to give a special thanks to my supervisor, Dr Ivan Zavada for his patience, inspiration and friendship, and to my associate supervisor Dr David Kim-Boyle for his support. I would also like to say a special thank you to Dr Jennifer Rowley and Professor Peter Dunbar-Hall for their support and friendship.

I would also like to thank the Apple University Consortium, HCSNet, and the George and Margaret Henderson Music Trust for their financial support and development opportunities.

Other academic staff I would like to thank for their support, guidance, and friendship are (in alphabetical order): Dr John Bassett, Professor Diana Blom, Dr Charles Fairchild, Matt Hitchcock, Professor Keith Howard, Dr David Larkin, Associate Professor Kathryn Marsh, Jacqueline Mees, Dr Helen Mitchell, Mr Daryll Pratt, Professor Anna Reid, Mr Craig Scott, Dr Michael Smetanin, Pierre St Just, Professor Jonathon Stock, Richard Toop, Dr Michael Webb, Gerard Willems.

I would also like to thank a number of general staff, past and present, at the Sydney Conservatorium of Music whose friendship and support have been, and are so very much appreciated (in alphabetical order): Justin Ankus, Stephen Backman, Ross Binfield, Rodney Boatright, Steven Burns, Martin Carroll, Marie Chellos, Elaine Chia, Ivy Chu, Timothy Crowe, David Kinney, Christina Goranitis, Henrietta Holden, Gloria Holland, Andrew

Humphries, Hideki Isoda, Ting Lee, Julian and Alistair Lockyer, Cynthia Marin, Jan Marshall, Guy McEwan, Siobhain O'Leary, Ahiegwu (Heggy) Odeh, Anthea Parker, Gemifa Parra, Cedric Poon, Katherine Rowell, Adrienne Sach, Scott Saunders, Peter and Caro Thomas, Rene Tsiknas, Marianne Uy, Stephen Yates, José and the team of attendants, the cleaning team, the security team, and Dy and all staff in the Jazz Café, for keeping me well fed and appropriately wired on caffeine!

Last but not least, I would like to give a *special* "shout out" to Adam Wilson, Jarrad Salmon, James Vuong, Dominic Blake, Benadict Carey, Aaron Harding, Rob Cornish, Ross Cooper, the Yates family, the Schaafsma family, Eduardo Suavo, Charlie and Kath Cooper, David Edwards, Leo MacPherson, Lucia Harvey, Andrew Robertson and Clemente Yap, John George, Oksana Vanyk, and Ros Hurrell for their support, humour, and friendship.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"Quite early in life I had the urge to make use of pure rhythm..."
– Alexander Tcherepnin, 1962

1.1 Introduction

Since the use of digital technology in creating rhythms, life as a composer of works for percussion has become much easier. No longer do drums need to be disassembled, loaded into a van for transportation to a studio, carried through doorways, and then re-assembled. Instead, different drum sets can simply be stored in a laptop. Paradoxically, digital technology has also made life for composers more difficult as there are now vast libraries of percussive sounds which can entail hours spent trying to find a single drum sound with the "perfect" timbre. The ability to microscopically deconstruct whether the positioning of a single percussive note in a rhythmic sequence is a millisecond too late or too early in order for the drumming to sound "real" also entails hours of tedium for a composer. This time is always spent for the same reason: we want the rhythms and the timbres to sound as though it were performed by a real drummer.

Since so much time is spent making digital drums sound as if they have been played by a real human, electronic composers and designers of electronic percussive programs have moved away from exploring timbral and musical possibilities of drums, in favour of creating software that plays drums out of the box with pre-existing loops. Although many of these software programs contain humanizing functions, the functions do not dynamically adjust themselves to modifications in the rhythmic sequence. With all of the computational power, and digital audio tools available in the digital environment, there is much scope to harness existing knowledge in order to create new percussive music. Unfortunately, this research has so far been limited. The result is a range of commercial models that create recognisable drum patterns and sounds, with limited compositional complexity and a "pop music" sound.

To reverse this trend, this thesis critically investigates human percussive performance, drawing on existing knowledge in the digital environment in order to: create a compositional software tool that can be applied creatively; inspire new musical creation; challenge a listener's understanding of percussive timbre; and push the boundaries of existing percussive models and extend the limits of their creative application. This compositional tool has applications for composers, studios, and sound composers alike. The methodology used to create the compositional tool, can also assist drummers that wish to learn more about drumming technique.

Several models have been developed that aim to create human performance variation

(Bilmes, 1993a; Friberg et al., 2006; Juslin et al., 2002; Waadeland, 2001 and Kirke and Miranda, 2009). However, very few of these models focus on percussion, and none have been used solely for creative purposes for composing new music. Moreover, and more specifically for this thesis, none of these models simulate human performance variation.

So why is there seemingly little focus on modelling human percussive performance variation and then applying this creatively compared to other instruments? Possible reasons could include the popularity of other instruments for example, the piano; compositional difficulty in percussion compared to pitched instruments; difficulties in synthesizing a jazz drum set effectively; or the difficulty in modelling a performance with four limbs. Perhaps it is easier to get a drummer and carry the drum set around, than it is to investigate human percussive performance variation?

An investigation into human percussive performance variation and computer music composition yields a tremendous array of scholarly writings from disciplines as varied as physics, computer sound synthesis, music performance analysis, computer science, and music composition. Within these disciplines, scholars have explored themes such as the instrumental mechanics of pitched and un-pitched percussion instruments and their vibrational characteristics (Fletcher and Rossing, 1998). This comprises the difficulty in accurately representing the micro-timbral variation present in un-pitched membranophones and idiophones in computer sound synthesis (Beauchamp, 2010; Bilbao, 2012; Macon et al., 1998; Masri, 1996); movement, timing and accents and their effects on percussive performance variation (Dahl and Altenmüller, 2008, 2013; Dahl et al., 2010); implementations of humanizing electronic percussive software (Hellmer, 2006); and the lack of creative application of percussive humanizing systems (Kirke and Miranda, 2009).

On the theme of instrumental mechanics of percussion instruments, existing scientific literature comprehensively describe the behaviours of a range of percussion instruments.[1] This research demonstrates the diversity of vibrational characteristics between membranophones and idiophones resulting in micro-timbral variation. Moreover, be-

[1] See Fletcher (1975, 1978, 1999); Fletcher and Rossing (1998); Hall (1991); Hiller et al. (1986); Hoffmann (2000); Kirke and Miranda (2009); Peaden and Worland (2011); Richardson (2010); Rossing (1992, 2000); Rossing et al. (1992); Rossing and Fletcher (1983, 2004); Schwarz (2004); Toulson (2009) and Wilbur (1996). Owing to the comprehensive studies undertaken by Thomas D Rossing and Neville Fletcher, recent literature on this subject is scarce. Examples of recent literature includes Peaden and Worland (2011); Richardson (2010) and Toulson (2009).

tween similar instruments there are significant differences in vibrational characteristics, particularly under different playing conditions.

On the theme of computer sound synthesis of percussion instrument, existing research focussing on the sound synthesis of percussion looks at the ways in which the complex and varied behaviours of these instruments can be simulated computationally.[2] Many different sound synthesis methods exist, each with their own distinct qualities. However, the diversity in vibrational behaviour across percussion instruments, especially considering the behaviours under different playing conditions and subsequent micro-timbral variations, is difficult to synthesize.

On the theme of human percussive performance variation, existing research empirically identifies variations between inter-onset-intervals (IOIs), known as “flutter” (Dahl, 2000). The concept of flutter is presented in many of the commercially available hardware and software products, such as Akai’s *MPC* range, Steinberg’s *Cubase*, Apple’s *Logic Pro* as “humanization” and “quantization” functions. These functions are defined as “random tempo deviations to beats” (Kirke and Miranda, 2009, p. 36). However, these functions are limited in producing both temporal and timbral variations directly related to the physical performance of the instrument. Similarly, of the very few attempts to humanize percussive performance (Bilmes, 1993a; Hellmer, 2006; Waadeland, 2001), none of these systems address micro-timbral variations of a nine-piece jazz drum set from a physical perspective. As a result, many of these models lack the required timbral variety.

On the theme of existing implementations of humanizing percussive software, many existing humanization functions apply temporal deviations and timbral variations without considering the context of a given drum within a broader sequence of drum strikes. That is to say, different combinations of drum strikes require different combinations of human movement. In many examples of humanizing percussive software the biomechanical link between timing and timbre is not made, resulting in timbre and timing variations that are not representative of real human percussive performance. Examples of this include “groove templates” (Kirke and Miranda, 2009; Wright and Berdahl,

[2] See Adrien (1991); Aramaki et al. (2006); Avanzini and Marogna (2010); Bilbao (2009, 2010, 2012); Bilmes (1993a); Cook (1997); Fontana and Rocchesso (1996, 1998); Fourcade and Cadoz (2001); Gouyon et al. (2003); Laird et al. (1998); Lakatos (2000); Legge and Fletcher (1989); Macon et al. (1998); Murphy et al. (2007); Smith (1991); Touze et al. (1998); Trautmann et al. (2001); van den Doel and Pai (2003); Waadeland (2000) and Wright et al. (2008).

2006) and quantization functions. Other performance modelling systems, such as the KTH rule system (Friberg et al., 2006) take some aspects of physical performance under consideration, but these systems focus largely on keyboard-based instruments.

On the theme of creative applications of percussive humanizing systems, attempts for creative application are either very compositionally limited because they model existing compositional styles or genres (Bilmes, 1993a; Gouyon et al., 2003; Waadeland, 2000; Wright et al., 2008), are not musically useful owing to the implementation (Hellmer, 2006), or that the main focus of the application is not music composition (Blaine and Perkis, 2000).

The research presented in this thesis will fill the gap in existing literature by creating a physical and performance context-dependant dynamic data-driven model, using a nine-piece jazz drum set, and applying this to music composition.

One contribution of this research lies in the creation of a large database of samples from a drum set, and the classification of this database based upon the vibrational behaviour and spectral features of the constituent instruments. The samples were recorded at various strike strengths and locations, and were then manually classified into zones based upon location of the strike. Each sample was then analysed to obtain their average loudness, spectral flatness and spectral centroid values.

Another contribution of this research is manifested in the approach to governing timbral variations in the performance model. This research hypothesises that timbral variations in human percussive performance are a result of two key aspects of human movement and biomechanics: firstly the difficulty of moving an arm from one height to another in successive strikes; and secondly, simultaneous bimanual striking at different strike heights. In order to model timbral variations based upon these two aspects of human performance, each drum sample is classified into one of five zones representing different target areas of the drum. The five zones, subdivide the playable surface area of the drum, and are based upon the vibrational behaviour of the instruments. Every combination of drum heights is allocated a difficulty level that results in a greater probability of a given zone being selected. When a zone is selected, a sample is played back from that zone.

A further contribution of this research is linked to the implementation of timing in the performance model. In order to create a data-driven timing model, three drummers of varying skill levels were recorded playing along to a jazz song. The timing information of strike during the performance was then manually verified and extracted. The onset values of the three drummers exist as an array of data which, due to the distribution of values and likelihood of onsets occurring at specific times, determine both positive and negative onset times for the playback of each synthesised note, selected based upon the zone selected.

The final contribution of this research relates to the design of the PD-103 with its MIDI compatibility for live improvisation, and the Rewire capability to integrate the software tool within other commercial compositional tools. Although the software tool can be used with MIDI, it is not designed around the MIDI specification allowing for future revisions of the software outside of the MIDI protocol. The model presented in this research is intended as an example to demonstrate and validate the thesis, which is the creation of new compositions, based upon temporal and timbral performance variations that model human performance variation. Moreover, this system extends the current state of the art with regards to models of percussive performance variation and their use in music composition.

1.2 Methodological Overview

Research areas include, firstly, the vibrational characteristics of drums with specific reference to membranophones and idiophones in a typical jazz drum set. Secondly, digital sound synthesis techniques and their suitability for electronically simulating the instruments in a jazz drum set. Thirdly, the biological and physical aspects of human percussive performance and the effect on instrumental interaction and timbre. Finally, the evolution of acoustic and electronic percussive performance and composition that has led to the development of a compositional framework based upon micro-timbral and micro-temporal performance variations. This multifaceted approach is largely unavoidable owing to the interdisciplinary nature of computer music. This is highlighted by Moore (1990), who states that:

> "Computer music, however, is strongly interdisciplinary. In computer music, therefore, a correct view is one that does justice to several points of view simultaneously" (Moore, 1990, p. 23)

This thesis presents a new methodology for the creation of a compositional tool that simulates variation in human percussive performance using a nine-piece jazz drum set. It combines primary and secondary research as well as employing experimental techniques to extract data relating to musical instruments and musical performance. Empirical data is obtained using a variety of analytical tools and methods, which were interpreted in the context of secondary research presented in the review of literature. This methodology is summarised in Figure 1.1.

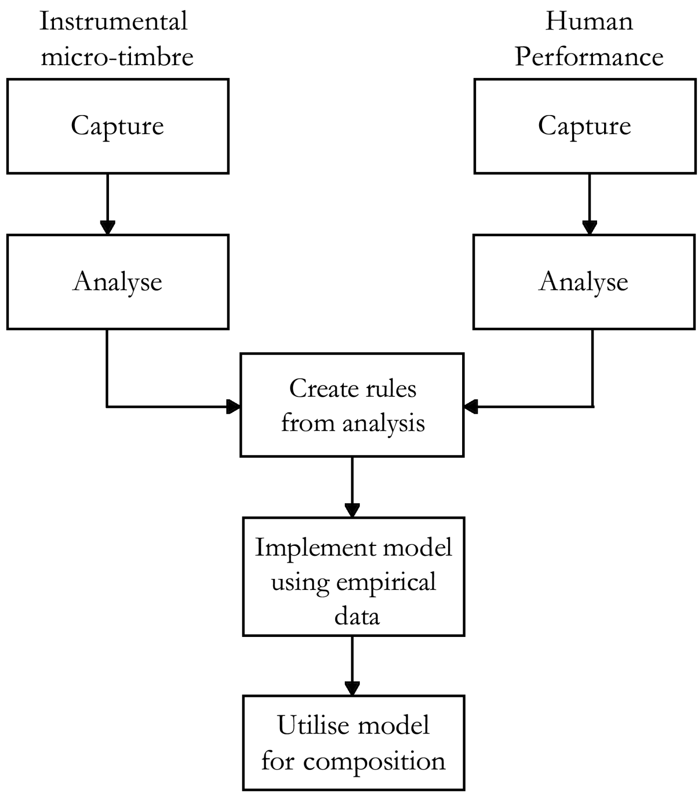


Figure 1.1: Methodological overview of the thesis.

The first part of the research involved capturing, and recording, nine thousand individual drum hits from a nine-instrument jazz drum set, containing both membranophones and idiophones. The result was a database of samples that extensively capture a wide variety of timbral variations for each instrument. This was followed by an analysis of the spectral content of the samples, using Matlab functions contained in MIRToolbox (Lartillot and Toiviainen, 2007; Lartillot et al., 2008). This allowed a parametric remapping of the samples in the database for performance modelling and compositional purposes.

The second part of the research involved capturing, and analysing human percussive performances in order to develop performance rules for the model. To determine the temporal variation inherent in a performance, and to model performer biomechanics, audio, video and accelerometer data based on the performance of three drummers of different skill levels (unskilled, semi-skilled, and skilled) was captured. This data was analysed using Max/MSP and Sonic Visualiser. These results were then used to develop performance rules for the model. The extraction of real timing values was used to represent the timing component of the model.

The third part of the research involved the creation of a software model, the PD-103,[3] which simulates human percussive performance variation. As discussed above, the PD-103 was developed using rules obtained from the analysis of the human performances, and combines the micro-timbral variation in the sample database with real timing values extracted from the performances. These fuzzy logic based performance rules were designed to interactively evaluate the current and first-order context of the instruments selected by the user, with a view to determine, based upon the biomechanical considerations in the performance, an appropriate timbral variation required to simulate both bio- and instrumental mechanics.

The final part of the research involved producing a creative music portfolio that effectively demonstrates the central aim of this thesis: the creation of a new compositional software tool, using a nine-piece jazz drum set, which simulates human percussive performance variation in different genres of music.

1.3 Research Aims

As discussed above, the main aim of this research is to create a compositional software tool that simulates human performance variation in percussion, using a nine-piece jazz drum set, in order to generate new and varied musical works.

Accordingly, this research aims to:

[3] The software is entitled "PD-103" in homage to the theory of "Participatory Discrepancies" or "PDs" (Keil, 1987, p. 275), and because it is version one using three parameters. The theory of PDs and the relationship to African music was an inspiration for interest into the area of percussive performance variation, and composition. This inspiration is reflected in the creation of meso-periodic rhythms in the composition portfolio.

- present an original real-time interactive compositional software model, which generates percussive performance variation by combining data-driven and stochastic approaches;
- demonstrate the creative application of the software model to different and varied genres of music, with specific reference to music with complex rhythms;
- extend existing percussive performance software models by using a nine-piece jazz drum set;
- demonstrate the creation and use of timing profiles extracted from the timing values of real human performances;
- present a new methodology for percussive performance modelling, and propose an original set of performance rules, based upon physiological constraints of human motion and the motor requirements of percussive performance;
- present a new method for classifying multi-sampled percussion based upon instrumental mechanics and the application of fuzzy logic representations of performance for interactive timbre production; and
- propose new timbral and data-driven compositional parameters based upon the instruments in a nine-piece jazz drum set.

1.4 Outline of this Document

As discussed, owing to the interdisciplinary nature of computer music, this research takes a multifaceted approach, taking into consideration: the effect of instrumental mechanics on timbre; synthesis techniques for modelling timbre; human performance; computational representation of human performance; and the subsequent implementation and compositional application of the instrumental and performance model. Subsequently, the thesis begins with a chapter that reviews the limitations of existing percussive performance modelling software.

Chapter Two reviews the literature concerning limitations of existing percussive software systems. Such systems consist of humanizing systems, data-driven systems, interactive systems, and other performance modelling systems. Humanizing systems apply generic timing values to events, while data-driven systems use values of real performances. In each type, the simulation of human percussive performance variation is limited. Limitations in compositional control are discussed in relation to interactive systems, followed by a review of other performance modelling systems applied to other instruments.

Chapter Three presents the theoretical framework of the performance software model, and is divided into three parts. The first part of this this chapter presents a nine-piece jazz drum set, consisting of membranophones (drums with skins) and idiophones (metal plates). Owing to differences in construction, each display unique vibrational characteristics after excitation, caused by differences between the instruments, particularly open and closed bottom drums, and edge clamped and non-edge clamped cymbals. Since a convincing performance model requires a convincing representation of the musical sound, it is necessary to provide a review of literature on the theme of physics of musical instruments (drawn from the work of Fletcher and Rossing (1998)) directly relating to each constituent part of the jazz drum set used within this study. Additionally, this chapter discusses global variables of the jazz drum set, such as their configuration and drum tuning. By investigating the instrumental variables from a physical perspective, major issues concerning the modelling of a musical instrument with such diversity, as well as number of configuration options are highlighted. This is followed by a brief review of the suitability of different sound synthesis techniques with specific emphasis on their ability to simulate the instrumental mechanics of the instruments. An assessment of each technique emphasises some of the advantages and disadvantages of their implementation in the current context, and concludes this section.

The second part of Chapter Three reviews the literature on the theme of human performance. Specifically, it presents a review of the literature concerning the physical factors that can affect human percussive performance, as well as the effect of biomechanics and motor control on timing and timbre. This chapter focuses on the underlying human variables within performance that lead to variations in excitation, altered mechanical and vibrational characteristics of the drums, and subsequent micro-variations in timbre. This discussion is formulated through a re-contextualisation of David Marr's (1982) tri-level analysis of human representative computer systems in the context of drumming development goals. This allows for a biomechanical and kinematic approach to analysing performance, which refines the discussion on the intrinsic human computation and planning of physical control and coordination, including the secondary effects of control and coordination resulting from percussive performance.

The third part of Chapter Three reviews literature on relevant compositional approaches, and presents some compositional applications of the software, which is designed to augment and extend the performance model into new compositional realms. This section begins with a review of relevant literature on the themes of computer-assisted composition, complex rhythms, composing using spectral features, and electroacoustic composition. The emphasis on complex rhythms includes discussion concerning African meso-periodic rhythms, and spectral features employing a data-driven approach to timbral parameterisation as an inherent feature of the performance model.

Chapter Four describes the methodologies used to collect empirical data from both the instruments and the human percussive performance. The first section further defines the conceptual methodology of the performance model by integrating conceptual discussion from the theoretical framework with empirical outcomes. Parts two to five of this chapter describe the research design methods, and reviews literature relevant to the research designs. In addition, these sections describe the protocols and procedures, and discuss how empirical data were analysed in order to develop a series of rules that could be used in the construction of a software performance model. Part six of this chapter presents an analysis and discussion of the performance data, and describes the performance rules created from a discussion on the finding of the empirical data. The implementation of these rules follows in part seven of this chapter, with particular focus on the amalgamation of the representations of timbre and timing. The computational implementation of parts of the graphical user interface of the PD-103 that assist in achieving the compositional objective is also discussed. The PD-103 software is included on the DVD for Appendix A.

Chapter Five introduces the composition portfolio and presents analytical notes for each work. This chapter comprises of a reflection of the artistic works presented, with a view to evaluating the compositional outcome with reference to the initial motivation and compositional vision. The analytical notes serve to discuss the implementation of the pieces with regard to the described compositional applications, and evaluate the compositional application of the model within the music style. Owing to the subjective nature of artistic evaluation, these analytical notes are for informational purposes, designed to offer an insight into the creative practice with the PD-103. The composition portfolio comprises of the following pieces:

Study No 1: African Meso-Periodic (I) (4:27)

Study No 2: African Meso-Periodic (II) (4:50)
Study No 3: Exploration of Pitch Variation in the Tom-Toms (3:37)
Study No 4: Exploration of Spectral Centroid and Isomorphic Rhythm (5:03)
Study No 5: Quark (1:00)
Study No 6: Demonstration of Improvisational Application (9:18)
Study No 7: Electroacoustic Piece (10:00)
Study No 8: Demonstration of Skill Levels x 6 (2:40)
Study No 9: Demonstration of Parametric Variation (skilled drummer) - 3 x (1:00)
Total Running Time: (43:55)

These compositional studies are included as ".wav" files in Appendix B, and on the Audio CD in Appendix C.

Chapter Six provides an evaluation of the performance model including a reflection on its compositional application. Areas of the performance model under evaluation comprise the data collection methods, the analysis and interpretation of the results, and the subsequent computational implementation of the model. The critical reflection of the composition portfolio combined with the evaluation of the performance model reveals areas for further artistic investigation.

Chapter 2

Limitations of Existing Percussive Software

"I started to be an engineer but I banged me thumb on the first day. I became a drummer because it was the only thing I could do..."
– Ringo Starr

As early as 1972, there were fears by those who did not understand computers that performing musicians would be rendered obsolete (Howe, 1972). In contrast, those who understood computers were describing computer music as being uniform and monotonous. This view is a natural progression of the human psychology towards computers, where "human uniqueness is defined not in terms of strengths but a certain frailty" (Turkle, 1992, p. 69). Humans consider computers to think in a similar way, but to lack "feeling" (Turkle, 1992, p. 72). In computer music, this frailty is apparent in a lack of human performance variation, coupled with lack of musical "feel". Washburne describes feel as the rhythmic positioning of a sound events relating to attacks and releases, and the execution of those sound events relating to "timbre, intonation, inflection, embellishment, and dynamics" (Washburne, 1998, p. 161). In the early days of computer music, creating variations in rhythmic positioning and execution was difficult owing to limited synthesis parameters and computational constraints (Howe, 1972). As technology has progressed, computational constraints have reduced, delimiting the depth and number of simultaneous synthesis parameters.

There have been several recent attempts to humanize computer performance in order to make music using percussion instruments. As well as commercially available *humanizing systems*, other attempts include *data-driven systems* that extract, analyse, and use data from human performances with a view to re-creating them. Attempts also include *interactive systems* whose algorithms modify musical parameters based upon user input. At the same time, other *performance modelling systems* have been used to simulate human expressive performance or to humanize performance on instruments other than percussion. Each of these systems demonstrates a variety of methodological perspectives relevant to modelling percussive performance variation. These will be discussed in the following sections of this chapter.

2.1 Humanizing Systems

Sequencers are the most common programmable devices for composing music electronically. Integrated into many digital audio workstations (DAWs), the sequential nature of the device, and its capability for recording, editing, playback and storage of performances as MIDI information, makes this device particularly popular.[4] The powerful control of MIDI information also allows the integration of external hardware and software, such as drum machines and grooveboxes and particularly percussive step

[4] For an interesting discussion on the workings of a sequencer, see Roads (1996, pp. 675-677).

sequencers.

Commercially available software sequencers allow the recording of live performances from MIDI controllers. In the case of drums, a drum set can be played and recorded as MIDI or as audio. In the absence of a drummer, another option is to manually enter the desired sequence of MIDI information, and to apply a humanization function that renders timing and velocity values to a user-defined selection of events according to a given preset. Software sequencers possessing this functionality include Steinberg's *Cubase*, Apple's *Logic Pro*, Avid's *ProTools*, among others. These presets often contain functions that edit the MIDI note information, for example reducing the velocity of every other note by 50% or adding 20ms of positive asynchrony to every third note. In some cases these humanization functions add "random errors within certain limits" (Kippen and Bel, 1994, p. 82).

Another function that is found in computer-based DAW sequencers is a form of quantization called the groove template. A user selects an existing recording or MIDI sequence, the sequencer analyses material for attack points, and automatically puts the selected pieces into a similar rhythmic pattern as the attack points in the reference material. From a real performance perspective, this reconfigures the syntax of the rhythm, without taking consideration of timbral variations created as a result of that syntax. Thus, the timbral and dynamic level deviations of the events in the selected material do not match the new temporal locations. While this can produce interesting musical results, this function removes context. There are instances when the timbre and dynamic levels of events require separation from the temporal locations, as seen in an experiment by Dahl (2000). Dahl's listening test used a "timing template" to determine whether listeners could identify interleaved accents in drumming based upon the cyclical durational positioning of the accented events devoid of timbre and dynamic level changes. Dahl's findings indicate that the durational differences alone marked accent locations to listeners. This research has important implications for music arranging in sequenced systems with limited timbral variation.

There are many instances of popular music that use single samples as opposed to multi-sampling. One example can be seen in early hip-hop made using hardware samplers of limited memory and sample time. Because of the limited memory, the same sample was often used in whole rhythmic sequences, resulting in no micro-timbral variations between notes. Consequently, the performances of these sequences are perceived to be artificial. Rath and Waltermann (2008) investigated this and found that inherent differences in drums of equally perceived volume, is relevant to the perceived character of a musical phrase. Rath and Waltermann also suggested that using a single sample in a musical phrase can be attributed to the "MIDI studio sound" (Bilbao, 2012; Rath and Waltermann, 2008).

The first programmable digital drum machine, the *LM-1*, was produced by Roger Linn in 1979 (Manning, 2004, p. 306). Since then, drum machines and grooveboxes have used a variety of sequencing formats, including step sequencers. The user is typically able to control the individual musical events and in some cases, like the early Roger Linn drum machines and the Akai *MPC* series, a "swing value" can be applied to a sequence of notes. One of the major drawbacks when using a step sequencer, however, is the ability to create complex polyrhythms in compound time signatures, which are often found in African music. In addition, the modification of individual events to create minor performance imperfections is time consuming. As the system loops through the sequence, these imperfections are also repeated. These swing functions tend to be applied across a sequence of events, for example, 15ms of positive asynchrony to a single instrument, and are generally not performance-context dependant. These asynchronies are described as "swing ratios", the ratio of the duration between long and short patterned eighth notes, (Friberg and Sundstrom, 2002) or "Beat-Upbeat-Ratio" (BUR) which is defined as the duration between downbeat and up-beat eighth notes (Benadon, 2006; Butterfield, 2011).

Despite differences between the swing function and real swing ratios, the swing function introduces a sense of either lateness or earliness to the sequence, sometimes referred to as "feel",[5] by changing the IOIs between other percussive events. Despite the popularity of swing functions in drum machines and groove-boxes, they are far from contextually accurate in realising true human performances. Ideally, a step sequencer should automatically apply a humanizing function (either a timbral or temporal variation, or both) to an event, depending on its current and previous context. Compositionally, a step sequencer should have greater flexibility in creating multiple compound polyrhythmic sequences.

In 2005, Kahl Hellmer developed a VST plugin that simulated timbral variation for

[5] See Stewart's "feel spectrum" as cited in Prögler (1995) for additional asynchrony values and "feels".

MIDI controlled percussive performance (Hellmer, 2005). The aim of his research was to overcome the dullness of typical drum machines, by creating a single plugin using the VST SDK (Steinberg Development Kit). Hellmer's approach used statistical methods to trigger multiple samples of a drum set that would simulate timbral variation. In order to simulate the time deviations, each sample was preloaded with 500ms of silence at the beginning of the file, with the time deviation represented by changing the sample-start position to offset any lateness. By setting the default sample start position to halfway through the silence, Hellmer was able to create an inherent latency of 250ms. Therefore, a sample-start position later than the default 250ms produced an early onset, and a sample-start position earlier than the default 250ms produced a late onset. In order to generate sample-point values, Hellmer used a normal frequency distribution curve, from which individual sample values were drawn. Hellmer noted that as the intention of the research was not to create a unique performance template, using a normal frequency distribution curve was the best approach given that previous research into the playing of a drums found differences in playing across drummers.

However, one problem with using a normal frequency distribution curve to generate timing deviations is the standard deviation. With 95% of the generated values between $+0.6$ and -0.6 (with a maximum of $+1.0$ and -1.0), the lower probability of selecting a higher deviation from the mean value, makes the resultant higher timing deviations more prominent. A large timing deviation could be construed as a playing error rather than an embellishment, particularly if the IOI is relatively low. This can cause timing deviations greater than the IOI resulting in a note that is perceived to be "out of sequence".

Although the insertion of 500ms silence into the header of each sample caters for the simulation of early and late onsets, this approach is not computationally efficient. For every 100 samples, an additional 50 seconds of audio are loaded into the computer's memory buffer. By dedicating computer memory to silence, fewer samples can be loaded into the computer memory, thereby reducing any timbral variability afforded by a multisampling approach. However, because Hellmer only used 81 samples to represent a bass drum, a snare drum, a hi-hat, two tom-toms, a ride cymbal and a crash cymbal, the sonic representation of the instruments and their timbral variation was already relatively low. In his implementation, only six samples sonically represent the complete range of timbres produced by a tom-tom.

Another limitation of Hellmer's implementation was that each instrument was represented by its own plugin and isolated algorithm, rather than being combined, and subject to a single algorithm that accounted for all instruments simultaneously. This resulted in each instrument operating independently of each other, thereby removing the performance context from the timbre and timing deviations. Unfortunately, the plugin was also prone to glitches in the attack phase of the sounds, which, as Hellmer describes, renders the software "useless in any context of music making or audio production" (Hellmer, 2005, p. 19).

In 2006, Hellmer revised the timing deviation algorithm by using timing values extracted from human performance (Hellmer, 2006). The performances were by jazz drummers, although the values used were extrapolated from the analysis of only eight bars of performance. In the revised software, he notes three phenomena related to performance: drift or a drummer's deviation from tempo (Dahl et al., 2000), flutter or the deviations of individual notes from the drift (Dahl, 2000), and swing or temporal shifting of hi-hat strikes (Waadeland, 2001). Drift was implemented as a continuously changing sample-point value that was generated independently from other timing errors and added to the silence offset value at the beginning of each sample. Each new drift value was then calculated from the last value, to a maximum of deviation of −4410 to +4410 samples. These sample point values were generated using a normal frequency distribution curve and scaled so that the maximum deviation values were relative to the current drift value. To simulate flutter, an additional value was added to the current offset position and drift value. The flutter value was user-selectable and multiplied by the standard deviation of the current note. In order to simulate swing, any MIDI notes corresponding to a manually added MIDI note were not emphasised, while other notes were. However, Hellmer identified two combined limitations of his implementation: the four programming pointers (a variable that stores the reference to another variable); and the introductory silence at the beginning of each sample. Where a pointer is reassigned before the end of the introductory silence of the previous value (up to 200ms), the minimum tempo and note value that can be played before mis-timing occurs is eight notes at 150 beats per minute (BPM) (Hellmer, 2006, p. 35).

Another recent attempt to humanize drum patterns involved applying different amounts of velocity to a MIDI drum pattern using fuzzy logic rules.[6] In this system, six fuzzy

[6] For further reading on fuzzy logic, the reader is referred to Bělohlávek and Klir (2011); Chen and Pham (2000); Passino and Yurkovich (1998) and Bergmann (2008).

logic rules were created, that modified the velocity of a bass drum depending on its temporal positioning relative to a downbeat in a rhythmic sequence. O'Sullivan and Boland (2010) note that the velocity patterns of the humanized rhythm were similar to the original, although less pronounced. This was attributed to the limited number of fuzzy logic rules, and fine-tuning of the fuzzy logic implementation. As a humanizing function, this system is limited because the humanization only relates to the velocity of a bass drum. Despite this, O'Sullivan and Boland describe how fuzzy logic has promising applications in the field of percussion humanization, particularly when used in conjunction with other computer techniques.

Computer models of percussive humanization also feature in research related to expressive music performance (Aldridge, 1994; Fletcher and Rossing, 1998; Hiller and Baker, 1964; Kirke and Miranda, 2009; Pinksterboer, 1992; Sullivan, 1990; von Hornbostel and Sachs, 1914; Warren, 1982). Literature on the creation of Computer Systems for Expressive Musical Performance (CSEMPs) has seen a range of approaches and techniques for different purposes (Katayose et al., 2012; Xenakis, 1991). These range from the expressive playback of musical scores and data files, such as Finale and MIDI, to expressive computer-music (Kirke and Miranda, 2009). However, current literature on CSEMPs indicates that there are very few systems that relate specifically to percussive performance.

Notably, one of the more successful percussive CSEMP models was used in Brazilian drumming (Wright and Berdahl, 2006). This system analyses the timbre and temporal location of each note within an audio file and creates a performance template, similar to a groove template, in order to simulate creative performances. Kirke and Miranda (2009, p. 27) note, however, that this system was not designed to generate creative performances. This is important because the creative manipulation of CSEMPs is a significant component of their evaluation (Katayose et al., 2012; Kirke and Miranda, 2009). Despite this, generating creative performances is seldom the primary function of a CSEMP.

2.2 Data-Driven Compositional Systems

As complex computational models of music become more prevalent, researchers in the field of Music Information Retrieval (MIR) have focused on ways of obtaining empirical data from music and musical performance in order to develop a greater understanding of human musical performance and cognition. One example of this is Inverse Performance Theory (Mazzola, 1995, 2011) which, in contrast to modelling a performance process from a score, involves "the reconstruction of the performance process from a given performance" (Mazzola, 2011, p. 227). MIR systems were not initially developed for creative purposes but, given the diverse technical field of computer music, they have been adopted for creative use (Dean, 2009, p. 38). *Mosaicking*, the automatic generation of sound sequences based upon their acoustical properties (Zils and Pachet, 2001), is one example of this. In most cases, the MIR process can be inverted to create models of extracted data. Many of these inverted MIR models are highly specialised and, therefore, compositionally limited.[7]

Another example of a data-driven compositional system is the drum loop system proposed by Riskedal (2002). This system sought to extract each onset within a drum loop, separate the drum sounds, extract the drum sounds frequency boundary (for classification), find a replacement drum sound (based on that classification) from within a database of sounds, and then either rebuild the existing drum loop with the new samples, or create a new drum loop (Riskedal, 2002). However, a reduction in scope during the course of his investigation led him to focus on the extraction of percussive onsets, analysis, and signal separation, resulting in a very limited compositional tool. Nevertheless, Riskedal's work does highlight the potential for compositional tools to be developed through convergent research sub-disciplines in computer music.

The automatic extraction of temporal information from polyphonic audio signals, an automated but non real-time version of Riskedal's example, was later attempted by Uhle and Dittmar (2004) who generated musical scores for un-pitched percussion instruments from an audio signal. The experimental procedure of the investigation, operating temporally, allowed an estimation of metrical structure and tempo (as tatums).[8] The pattern-matching procedure made use of "rules derived from musical knowledge" (Uhle and Dittmar, 2004, p. 6). This included; stylistic determination based upon the IOI of same-instrument events, which can be extended to use "backbeat" to determine bar lines, and metrical bar line start points based upon bass drum position. However, Uhle and Dittmar observe that the methods fail when the source audio material features highly expressive performance (p. 8) and pieces with lots of "fills". Moreover,

[7] For further information regarding the field of MIR, the reader is invited to read Britto Jr, A. D. S., Gouyon, F., and Dixon, S. (Eds.). (2013).

[8] Tatums are defined as a perceived high frequency pulse, and "is the lowest level of the metrical musical hierarchy" (Bilmes, 1993b, p. 21).

a high range of dynamic levels between events for the detection methods negatively affected the overall performance. Another limitation was the use of inferred musical knowledge in the meter estimation process, where the identification of metric position based upon same-instrument IOIs and assumptions of the notion of "backbeat" (p. 6) failed to account for stylistic deviations of these rules. A noted example was the backbeat of the bass drum in Reggae music (p. 8). Moreover, this research, did not stylistically identify the one hundred and sixty-one musical excerpts under analysis. This is particularly important given their acknowledgement of stylistic deviations from inferred knowledge. As a result, the research did not allow sufficient appraisal of the robustness of the methods across different stylistic types and their inherent complexity.

Focussing on re-creating percussive rhythm, Bilmes (1993b) developed a software program called "xited", which contained three perception-based algorithms that "characterise percussive rhythm". The three algorithms were based on metric content, tempo variation, and deviations (Bilmes, 1993b, p. 1), from a Cuban drumming performance and subsequently resynthesized. During the resynthesis process, six different quantization methods were applied to the samples extracted from the original performance and then triggered. Bilmes reported some success in the ability of the algorithms and the effectiveness of the quantization methods. One limitation of this research, however, lies in the re-synthesis of the drum pattern using samples from the original performance.

One contrasting application of MIR can be seen in the development of a method to detect and classify different percussive events in real-time for the purposes of implementing the technique into an interactive musical system (Şimşekli et al., 2011). However, the interactive and real-time nature of the implementation revealed some considerations that served as constraints to the implementation of various techniques. One such technique required the development of a model-based approach, whereby a scaling variable was used as a template to determine the spectral shape of each instrument. In addition, a Hidden Markov Model (HMM) was used to classify the sounds from the audio source. This implementation was then tested using different hand-clapping and Turkish percussive instruments. Şimşekli et al. found that the real-time performance was good, thus lending itself to different interactive applications, although most of the latency was caused by the HMM. They also found that using the technique in relation to other percussive instruments and polyphonic music significantly reduces the real-time performance of the model. This example demonstrates the challenges faced in extending MIR models to composition particularly in real-time percussive humanization.

2.3 Interactive Systems

An interactive music system is broadly defined as a system "whose behaviour changes in response to a musical input" (Rowe, 1993, p. 1). From this starting point, Rowe goes on to describe how interactive music systems draw upon perspectives from several different fields of study, including artificial intelligence. Artificial intelligence is of particular importance because behavioural changes in the system require some form of decision-making, thus inferring a degree of (artificial) intelligence. The levels of interactivity and intelligence or "dimensionality of control" (Pressing, 1990, p. 12) amongst current systems vary significantly and are driven by the intended application of the system: be it composition or live performance.

In his 2009 article entitled "Understanding Interactive Systems", Drummond (2009) writes that the term "interactive composing" was first used by Joel Chadabe in 1981. Chadabe described "interactive composing" as a method of "using performable, real-time computer music systems in composing and live performance" (Drummond, 2009, p. 22). In this article, Drummond discusses the definitions, models, and classification of different interactive systems, and describes a complex interplay between encoded algorithms, mappings and sound generation routines, which evolve with performer interaction.[9]

One of the earliest interactive rhythmic systems, the Rhythmicon, was developed by Henry Cowell and Leon Theremin in the early 1930s (Schedel, 2002). Despite being relatively unsuccessful, Rhythmicon is considered to be an important part of interactive music, despite the lack of definition for interactive music at the time.[10]

Drummond also writes that Chadabe distinguishes non-interactive from interactive musical systems by describing their use as a "design-then-do" procedure (Drummond, 2009, p. 22). Two of Chadabe's compositions, *Solo* (1978) and *Rhythms* (1980), were written using interactive systems. Interestingly, in *Rhythms*, Chadabe said that initially the computer generated the melodies and rhythmic patterns automatically, while he entered commands to modify chords, pitch, melodic and rhythmic variations, and random notes. According to Chadabe, the sounds contained in the piece were "rem-

[9] For excellent discussions on all facets of interactive systems, the reader is directed to Drummond (2009); Paine (2002); Rowe (1993); Winkler (1998).

[10] For further information on Henry Cowell and the Rhythmicon, please refer to Smith (1973) and Schedel (2002).

iniscent of Indonesian, Caribbean and African percussion instruments" (Drummond, 2009, p. 22).

In 1996, Leonard, a composer and designer of interactive computer music systems, wrote an article entitled "*Legacy: San Lazaro. The integration of composition, performance, and computer programming*". In it, he makes specific reference to a piece entitled "*Legacy: San Lazaro*", in which a computer interacts in real-time with an improvised saxophone performance (Leonard, 1996). The composition makes use of both rhythmic and input variation generators. The rhythmic variation generator maps sampled percussion timbres to a polyrhythm's constituent parts, with rhythms imported via standard MIDI files. Markedly, this implementation extended the rhythm in ways that a live percussionist could not. For the input variation generator, the computer listens to the pitch played by the performer, then using a Markov chain, creates a melody and assigns that melody to a polyrhythm. The Markov chain buffer uses a "first in, first out" approach (Leonard, 1996, p. 90) in order to ensure that the latest performance information is represented.

Variations to the rhythm and input generations were designed to coincide with one of three different sections in the piece. Notably, the second section uses the performer input volume to control the panning amount of a simulated organ sound. Additionally, the computer plays a triad of the top pitch of the input note, and two stochastically generated pitches below the performers note. The third section sees changes in pitch input of two or more successive notes alter the rhythmic density of the polyrhythms. Leonard describes how in the final part, he programmed "a call and response between human and machine" (Leonard, 1996, p. 91) in the form of performer MIDI input. Leonard then describes how "this section challenges the performer to adapt to frequent changes in musical directions" (p. 91). Hsu used a similar interactive call and response method applied to the timbre of a saxophone player (Hsu, 2006, 2008). It contained real-time algorithms, which analysed the timbral characteristics of the saxophone input, and then generated the response output from the computer.

Another interactive music system relevant to this research owing to its focus on percussion, is the "Jam-O-Drum" developed by Blaine and Perkis (2000). The Jam-O-Drum is a tactile multi-user device that integrates drum pads and sensors or drum triggers in a tabletop surface. The drum pads are connected to a computer where an algorithm rhythmically quantizes and emphasises the strikes (to account for performer skill level), sends audio feedback to the user based upon these algorithms, and then renders visual projections onto the surface. However, the Jam-O-Drum gradually moved away from musical applications towards more game oriented visual projections, primarily because its focus was interaction design. Like Leonard, Blaine and Perkis consider their call and response interaction system to be the most effective way of encouraging users to interact (Blaine and Perkis, 2000, p. 173). The disadvantages of this system, however, are two-fold. Firstly, the computer often misinterpreted expressive performance. Secondly, the high dynamic range of user input negatively affected the system's performance. Nevertheless, despite these drawbacks, the performer was playing "with" the machine.

By their very nature, interactive systems require the user to defer some form of decision making to a software agent. Moreover, the level of interactivity and decision-making is governed by the type of software agent (Whalley, 2009). In his review of interactive music systems, Whalley describes how such deployments can include multi-agent systems (MAS), distributed artificial intelligence (DAI) (multiple agents with user input) and multi-agent based simulations (MABS). Irrespective of agent type, however, the level of interactivity is largely dependant on the application of the system. Whalley summarises this, within the disciplinary context of interactive music, in Figure 2.1 below (Whalley, 2009, p. 157).

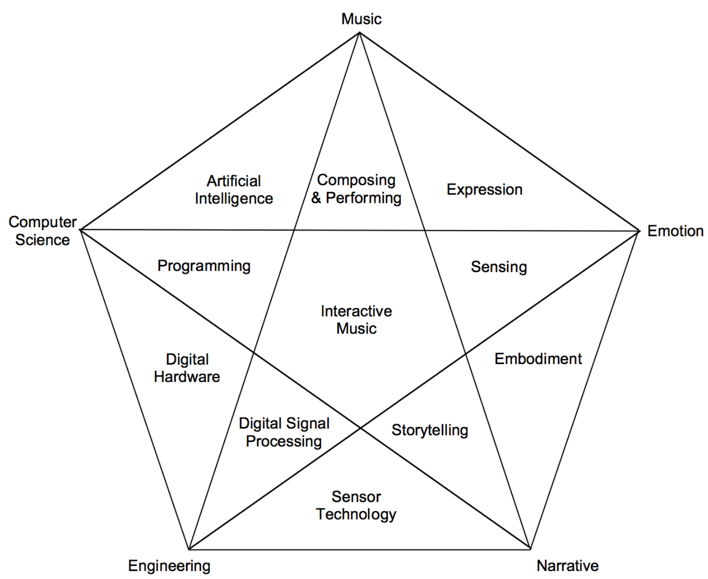


Figure 2.1: The interdisciplinary context of interactive music (Whalley, 2009).

For Whalley, the disciplinary context of interactive music links the disciplines of Music, Computer Science, Music Psychology, and encompasses the sub-disciplines of Artificial Intelligence, Composing and Performing, and Expression. To a certain extent, this paradigm provides a useful foundation for the central thesis of this study. However, in performance, it is difficult to separate expression from performance variation. This is because there is no clear threshold for separating between systematic (expressive action) and random variation (Juslin et al., 2002). It is not the intention to answer this specific question here as it is outside of the scope of this investigation. Rather, this thesis posits that musical events are variations and the listener perceives expression from these variations.

Research that follows a similar contextual paradigm is “Kinetic Engine” (Eigenfeldt, 2006), an interactive performance system aimed at modelling a percussion ensemble. The Kinetic Engine uses four multi-agents, each performing a particular predetermined percussive role, for example bass drum, snare etc., based upon layers of rhythmic organisation. Kinetic Engine also uses constrained random procedures for modelling intelligent behaviour. Eigenfeldt notes that by constraining random procedures, the system produces greater levels of consistency while allowing small variations to occur, because “limiting potential choices does not limit the intelligence of a system” (Eigenfeldt, 2006, p. 2). There are two implications of adding constraints to random procedures in an interactive system. Firstly, because there is greater consistency and stability toward the desired result, the system is perceived to be more skilled in decision-making. This should not be confused with learning systems however, whose function is to independently learn behavioural state changes either with each iteration of a process or each *nth* order process in a chain. Secondly, the system’s freedom in decision-making equates to the degree of control by the composer in the composition process. With fewer constraints the system exerts more influence on the resultant composition. With more constraints, the composer has more influence. This is particularly important in decisions that produce vastly different musical artefacts. Ultimately, this is a choice for the composer. Since the purpose of this thesis is to apply the compositional music software to different rhythms and various musical styles, it is necessary to ensure that the composer, rather than the computer, has more influence on the creative process.

2.4 Other Performance Modelling Systems

Traditionally, music performance theory has been approached either qualitatively (philosophically) or quantitatively (empirically) (Mazzola, 2011; Sethares, 2005, p.31). Unifying these approaches, Mazzola advanced a concept called the "oniontology" of musical performance (p. 21-22). This is a philosophical multi-layered approach that extends the existing dimensions of musical ontology, such as performer reality, communication, and semiotics (Figure 2.2), to include embodiment (Figure 2.3).

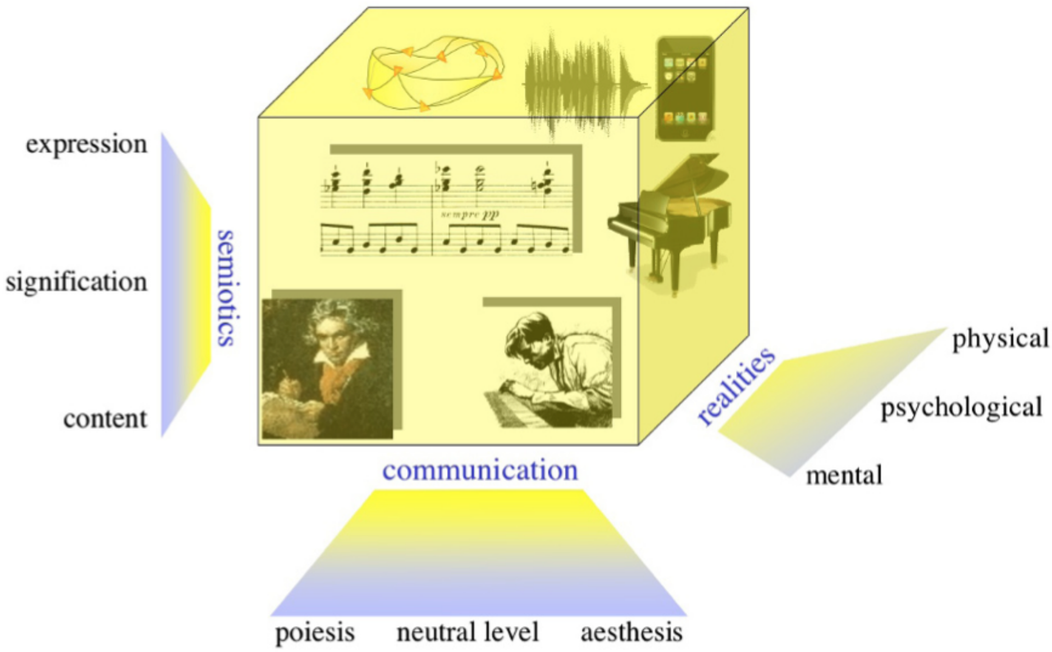


Figure 2.2: The traditional dimensions of musical ontology (Mazzola, 2011).

Although semiotics, communication and realities are all important philosophical aspects of musical performance, it is outside the scope of this investigation to quantify these transformations. For example, the semiotic dimension presents problems in percussion, where the signification, that is the transformation from the score to the content, is subject to many more variables that make interpreting the transformation ambiguous, such as the physics of the drums and the individual expressivity of a performer. In addition, the dimension of communication comprises a discussion on the modelled reference material, or the performance of such material with regards to the performance model. Since this thesis deals with predominantly un-pitched instruments with limited envelope control, there may be difficulties in correctly interpreting communication from any given performance. Similarly, the mental and psychological realities of the performance may be difficult to extract from a percussive performance. In both cases, there is little or no empirical data available from a percussive performance that could lead to the generation of performance rules. The only dimension that facilitates the extraction of human percussive performance variation is physical reality. Placing this musical ontology in context, Mazzola presents two new dimensional layers to musical ontology (facts), gesture and processes. These are shown in Figure 2.3.

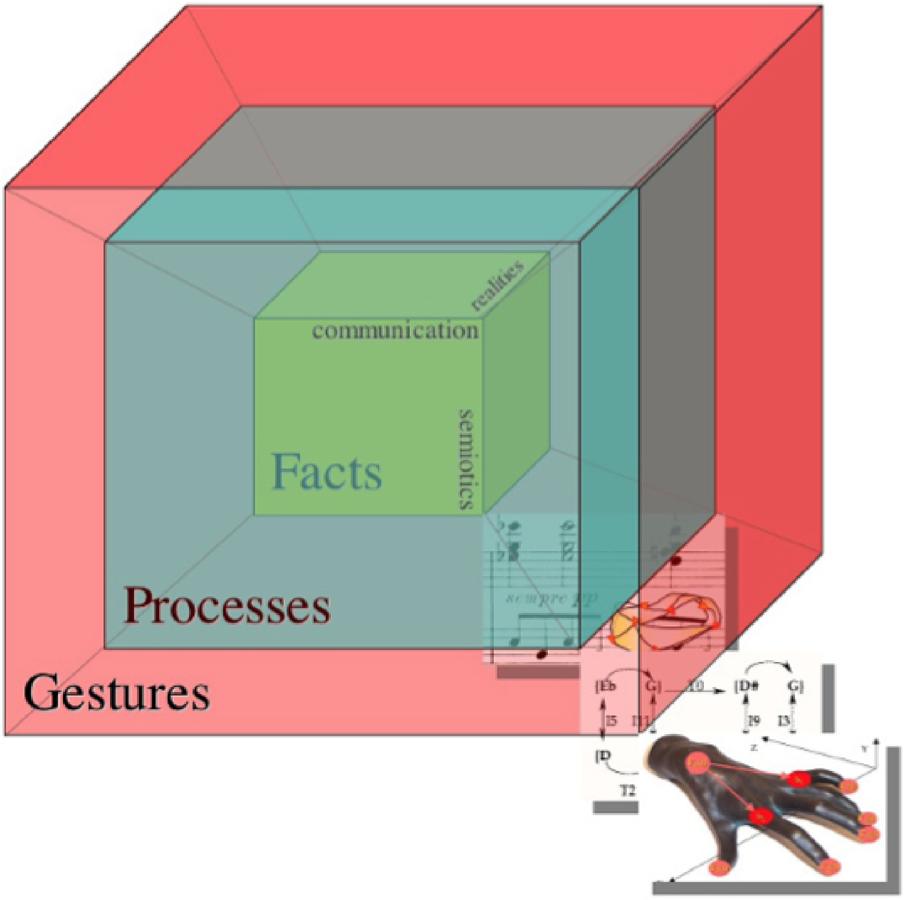


Figure 2.3: The musical “Oniontology” (Mazzola, 2011).

These two new dimensional layers are defined as the gestural schematization of the logic by the composer (gestures) and the logic of the musical construction (processes) (Mazzola and Thalmann, 2011; Roads, 1996). In relating this “oniontology” to this investigation, parallels can be drawn between gestures and performance context, between processes and the control parameterisation, and between facts and the sample database that represents the acoustical output of instrumental physics and performer interaction. Therefore, the most appropriate ontological solution to performance modelling is one that, through logical processes, links the physical aspects of the instrument and the instrumental interaction (facts) to simulate performance; in other words, an inverse model of performance, or a data-driven system.

In summary, Mazzola's work aims to mathematically analyse and reproduce the transformations between these complex oniontological spaces. However, for the purpose of this research, it is impractical to use mathematical modelling to control the sample database. This is because assumptions must be made to address the complexity of physical human jazz drum set performance, as well as the number and relative position of the instruments used. In addition, it is important that the control methods do not affect the computational overhead of the model.

In contrast to Mazzola's work is another quantitative performance modelling approach is called analysis-by-synthesis. Analysis-by-synthesis, which is part of the KTH rule system, was first presented by Anders Friberg at the KTH Royal Institute of Technology in Stockholm, Sweden. This approach has been described in the following way:

> "If expressive actions act upon musical structure in a systematic way, then it should be possible to define these actions in the form of a rule system, much like a "grammar" of performance. Moreover, it should be possible to implement this rule system using computer-based tools to generate "automatic" performances." (Risset and Wessel, 1999, p. 125).

The creation and generation of "grammar", or rules, of performance from the systematic application of expressive acts on musical structure, does not lend itself to drumming. This is because certain rules relevant to the piano may not be musically relevant to un- and quasi-pitched drums, for example, harmonic and melodic rules. In addition, the transformation of a musical score into performance rules is much more difficult in drums due to the nature of the instrumental behaviour and performer interaction. This is demonstrated by several of the rules in the original rule set that cannot easily be applied to percussion (Friberg, 1991; Legge and Fletcher, 1989). Although this will not be discussed in great detail here, suffice it to say that the nomenclature of the rules demonstrates the difficulty in applying such a musical approach to performance containing un-pitched percussion: Double duration; Off-time repetition of tone; Leap distance; Leap tone duration. This is also the case in multiple-parameter rules such as: Melodic Charge; Harmonic Charge (of a chord); Chromatic Charge; High Sharp (intonation rule); Melodic Intonation (intonation rule). Moreover, since it was first described in 1991, the KTH rule system has been continuously augmented with new rules, to include Phrase Arches (Friberg, 1995; Risset and Wessel, 1999, p. 125) and punctuation (Friberg et al., 1998; Pierce, 1992, p. 39) amongst others (Beauchamp, 2010; Friberg, 2006, p. 66).

Looking specifically at the creation of rhythmic performance, Waadeland has proposed an interesting and alternative view to creating expressive performance by linking rhythmic performance with bodily movement:

> "kinesthetic movements of the body [during a performance] are intimately related to *"timbral"* aspects of *rhythmic performance*. Furthermore, from this point of view, *different kinaesthetic movements of the body are expressions of different "rhythmic timbres"*" (Waadeland, 2000, p. 112)

According to Waadeland, rhythmic performance is based upon the frequency modulation of movement. In his research, performed rhythmic movement was considered a series of oscillations, and expressive timing was considered a series of continuous transformations of rhythmic structure. These transformations, generated by rhythmic movement, create the expressive timing. Notably, Waadeland asserts that the rhythmic frequency modulator (RFM) was flexible enough to create rhythmic performances as a standalone compositional software tool and interestingly simulates the compositional effect of phasing used by Steve Reich in his 1971 piece, entitled *Drumming* (Waadeland, 2001). Although Waadeland investigates bodily movement and expressive timing, the relationship is defined strictly in terms of the effects of rhythm on movement, particularly eurhythmics and kinaesthesia. However, in order to simulate a nine-piece jazz drum set using RFM, nine independent modulations are required, with each modulation linked to each other in some way. A problem with such an approach is the large number of links required ($9 \times 9 = 81$), the potential difficulty in defining links between modulations, and the complexity of these links in the context of human movement.

Another model that considers human movement is the GERM model (Beauchamp, 2010, pp. 66-67; Juslin et al., 2002). Using many of the rules from the KTH Rule System, the GERM model is a four-part approach. It consists of: Generative rules or methods, many of which are derived from empirically supported rules of the KTH system; Emotional expression or the application of the results of studies concerning the emotional expression in music; Random variations in performance; and Motion or physiological constraints and instrumental requirements in performance. Although both the GERM and KTH systems were designed to generate automatic piano performances from musical scores,[11] the GERM model has made significant improvements on the KTH model, such as the specification of random variations in performance and human motion. Neither of these systems, however, were designed for composition.

Juslin et al. note that there are two types of performance variation: systematic or intended; and those that are random or "error variances", which are the result of the human motor system (Beauchamp, 2010, p. 69; Juslin et al., 2002). Juslin et al.'s research provides a rationale for including random variations, asserting that it is difficult to distinguish between systematic and random variations. Aesthetically, random variations are an inherent component of human musical performance, which is one of the key tenets of this thesis. For listeners unacculturated to a musical style, it may be difficult to perceive the difference between the two types of variation. This perceptual differentiation is further confounded when deviations occur at temporal perceptual boundaries and, in the case of timbre, create auditory masking. Similarly, phenomena such as temporal drift, indicates that the degree of random variation, may not be entirely random.

The use of empirical data is an important area of discussion in music performance analysis. Analytical models use real performance data such as timing (Beauchamp, 2010, p. 69; Repp, 1992, 1994). One benefit of this is that the extracted data can be quantitatively analysed. This is particularly useful when comparing performance data from different subjects, for example, pedal timing in expressive piano performance (Repp, 1996; Roads, 1996, p. 141), timing deviations (Repp, 1999; Serra et al., 1990, p. 12), and musical synchronization (Repp, 2006; Serra, 1989, p. 86). Some studies using this approach have identified the "swing ratio" in jazz drumming (durational ratio between notes in a eighth-note pattern) (Friberg and Sundstrom, 2002; Serra et al., 1990, p. 12), and temporal coordination in duets (Schögler, 2000; Serra et al., 1990, p. 12). In addition, rules can be generated from identified trends in the data analysis. Despite this, Juslin et al. argue that because of the difficulty in identifying the difference between systematic and random variations in empirical performance data, the result is

[11] There are a number of physiological factors that can affect the generation of percussive performance variation, particularly when compared with musical instruments whose interfaces require less compound physical movements (e.g. a piano). However, as one of the objectives of this investigation is to present a compositional tool that can be used to generate *new* music, the generation of automatic performances from existing notated input, although an interesting avenue of further investigation, is a compositional goal outside the scope of this thesis.

the creation of an algorithm that produces uncorrelated fluctuations, or white noise, in timing variability, which increased with IOI duration. This algorithm was based upon previous studies on finger tapping synchronicity, rather than using actual performance data.

Another important area of research is that of kinematic specification of dynamics (KSD). According to this principle, "movements specify the causal factors of events", (Roads, 1996, pp. 550-552; Runeson and Frykholm, 1983) Interestingly, Juslin et al. make a connection between the KSD principle and music performance in the following way:

> "when we hear a music performance by a human being, there should be kinematic information in the performance pattern that specifies that the performer is human. This information should be missing if the performer is a computer program" (Juslin et al., 2002, p. 84)

This statement has implications philosophically, aesthetically, and practically. From a philosophical perspective, using empirical performance data is one method of applying the KSD principle to a performance model. By extension, the generation of new works using the empirical data re-applies the KSD principle to new contexts. From an aesthetic perspective, the kinematic information in the performance should produce interesting variations, which may be perceived by the listener as either systematic or random. The result of which will yield a perception of real human performance, the extent of which will be based upon the relevance of the kinematic information in the new context.

From a practical point of view, capturing real performance data has the potential to generate significant amounts of information or data sets, which must be used in conjunction with other methods. One such method that works well with large data sets is probability (Ames, 1990) for data selection purposes, where the probabilistic algorithm is representative of rules derived from analysis of the data and the context in which the data was extracted (a synthesis-by-analysis approach). In the case of the fourth component in the GERM model, specifically Motion, real performance data would inherently reflect the motor requirements of a drum set, physiological constitution of the body (Juslin et al., 2002, p. 85), and movement complexity. In the case of timbre, empirical data can inform the timbral selection process and be used creatively for musical purposes.

Chapter 3

Towards a Percussive Performance Model

"The development of motor skill can be traced as the progress from reactive movement to movement fluency, coupled with a flexibility in tailoring action to the details of an infinite variety of contingencies..."
– L. Henry Shaffer, 1982

A review in the previous chapter of existing research on percussive performance models highlights significant gaps in the types of methods employed in order to simulate variations in human percussive performance. In many of the percussion-related methods, timbral variation is limited to rudimentary representations of the instrument which, given the importance of timbral variation to the perception of humanization, is surprising. Also important to the perception of humanization is timing. Percussive performance methods that simulate timing fluctuations, ranging from modulation and stochastic methods to the extraction of real performance timings for re-creation, were also reviewed in the previous chapter. Many of these models do not adequately link timing with timbre and the relationship between their production during a performance. Central to this relationship are the physical constraints of the performer during a performance. While this has been considered in some models (Juslin et al., 2002; Waadeland, 2000), it has not yet been adequately defined in terms of instrumental technique and the biomechanical constraints of playing a jazz drum set.

In order to understand why representations of percussive timbral variation have been limited, this chapter will examine the acoustical and mechanical behaviour of key instruments of a nine-piece jazz drum set and their effect on the production of timbre, particularly small variations in timbre, or "micro-timbre". In order to define the relationship between timing and timbre in a performance context, this chapter will examine the biomechanics of human percussive performance on a nine-piece jazz drum set. These investigations form the foundation of a theoretical framework that considers the importance of instrumental mechanics in the creation of timbral fluctuations in human performance in order to construct a convincing percussive performance model.

A survey of computer systems for expressive music performance was undertaken by Kirke and Miranda (2009). In this review, Kirke and Miranda proposed four terms of reference for their evaluating the expressive performance systems, including "performance creativity". Kirke and Miranda define performance creativity as "the ability of the system to generate novel and original performances, as opposed to simulating previous human strategies" (Kirke and Miranda, 2009, p. 10). In their evaluation, Kirke and Miranda rated the performance creativity of many of these expressive music performance systems to be low, with the only percussive system reviewed, *Drumming*, being rated with the lowest performance creativity, and among the lowest in expressive representation (Kirke and Miranda, 2009, pp. 34-35).

The computer as a compositional tool represents, in many ways, choice. The creation of a compositional software tool is a very personal experience, dependant upon the level of function and compositional control required by the user. This supports the argument that there is no wrong computer-driven compositional approach. This is particularly apparent in the types of compositional systems available, the wide range of levels of interactivity, and the different ways in which to interpret and implement compositional concepts. Human variation in percussive composition has gone full circle. From the inherent variation in the human performance of a score, to the mechanical performance of computers, composers now find themselves looking back at percussive performance in order to create more human-like percussive performances on a computer.

This research will fill the gap in the literature by continuing the explorative work of composers, in order to create a model that is representative of human micro-temporal variations with micro-timbral diversity, is compositionally useable, computationally efficient, and makes a convincing performance model. Moreover, such a model must be able to explore the timbral possibilities of the instruments, and be applied to diverse types of rhythm and music. These factors form the foundation of a compositional framework that allows for the intra- and inter-instrumental exploration of micro-timbral variation in the instruments. This theoretical approach also allows a re-contextualisation of the instrument, in combination with micro-variations in timbre and timing. Such a re-contextualisation further enables an exploration of algorithmic interactions as a compositional element of the performance model.

3.1 Why is Modelling Percussive Timbres so Difficult?

Percussion can be classified in many ways, ranging from their sound production characteristics, their role in musical contexts, "whether or not they convey a definite sense of pitch" (Rossing, 2000, p. 1), or by cultural derivation (Blades, 1992). A typical drum set configuration consists of a bass drum, a snare drum, hi-hat, tom-toms (including floor tom), ride cymbal and crash cymbal (Sweeney, 2004b), although configuration of the individual components (e.g. drum size) and the positional configuration of the set, can be extended by personal preference (Strong, 2006). The standard drum set can be divided into two groups by virtue of their sound generation methods, with the bass drum, snare and tom-toms belonging to the family of membranophones, and the cymbals belonging to the family of idiophones (Rossing, 2000). Regarding membranophones, these three elements of a drum set are similar, insofar as they all have clamped circular edges, whereas the idiophones are unclamped circular plates supported in the middle. The hi-hat differs from the ride and the crash cymbals as it comprises of two opposing circular plates, supported in the middle with the clamping force between the two plate edges and controlled by the performer, in either an "open" position (no contact), "closed" position (forced contact), or a position somewhere in between. This clamping changes the vibrational characteristics of the plates and the resultant sound, and will be looked at together with the vibrational characteristics of membranophones and idiophones described above.

3.1.1 The Vibrational Characteristics of Membranophones

One of the main starting points for the theoretical consideration of membranophones is the membrane. Fixed at the edges, the membrane serves as a primary vibrator which can be described as two-dimensional, where vibrations travel in both radial (concentric) and azimuthal (diametric) directions (Moravcsik, 2001, p. 188) and "the resonant vibrator is the air column inside the drum" (Moravcsik, 2001, p. 191). It is these complex vibrations that produce "inharmonic overtones" that "give percussion instruments their distinctive timbre" (Rossing, 2000, p. 2). In order to analyse these vibrational characteristics, many existing studies describe not only the membrane, but the environmental conditions in which the membrane operates (e.g. an ideal membrane). For example, where the membrane is wholly flexible and vibrating in a vacuum (Rossing, 1992), where the membrane has "zero thickness" and is "perfectly elastic" (Moravcsik, 2001, p. 189; Raichel, 2006, p. 111), has "no stiffness" (Rossing and Fletcher, 2004, p. 70) and is not subject to damping (Raichel, 2006).

In order to find the fundamental frequency (f_0) of an ideal circular membrane, Hall (1991) uses the equation:

$$(f_0) = \frac{0.766}{D} \sqrt{\frac{T}{\sigma}}$$

"where D is the diameter [metres], T is the tension (Nm), and σ the mass per unit area" (Hall, 1991, p. 163). Frequencies for the normal modes of vibration can be calculated using Bessel function-based wave equation theory for an ideal circular membrane:

$$(f_{mn}) = \frac{j_{mn}}{2\pi\alpha} \sqrt{\frac{T}{\sigma}}$$

"where j_{mn} is the nth root of the mth Bessel function, T the membrane tension, α the membrane radius, and σ the mass per unit area" (Dahl, 1997b, p. 59). Fletcher and Rossing (1998) identified fourteen different concentric and diametric modes (m, n), each corresponding to a different relative frequency, where mode (0, 1) represents f_0. The inharmonic overtones produced by these vibrations are represented by the frequencies given in Figure 3.1, which display non-integer relationships with the fundamental in contrast to harmonic tones whose frequencies are integer multiples of the fundamental (Roads, 1996).

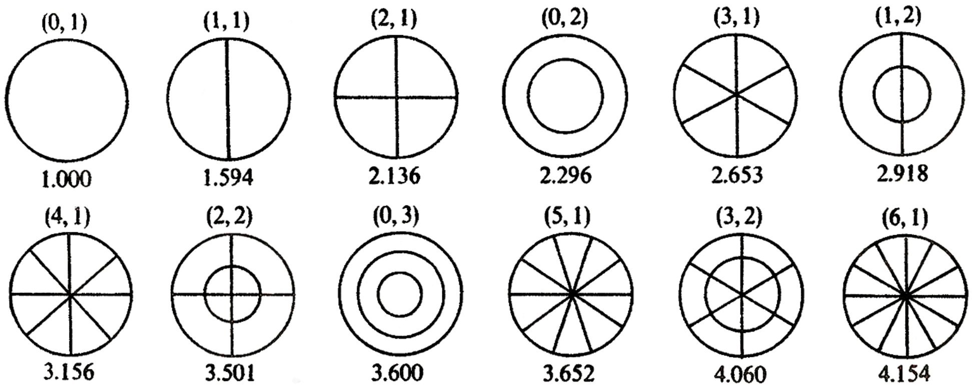


Figure 3.1: The vibrational modes of an ideal membrane, and their relative frequencies (Fletcher and Rossing, 1998, p. 75).

In practice, the frequency modes in a real membrane are different to an ideal membrane (Moravcsik, 2001), because real membranes are subject to air loading, bending stiffness and shear stiffness, in which the latter two raise the modal frequencies while the former lowers them (Fletcher and Rossing, 1998). Furthermore, it is the excitation of a real membrane in different locations that causes different combinations of these effects (Moravcsik, 2001), with the creation and subsequent decays of various modes being unequal. In addition to the inequality of modal decay, different strike locations are more efficient in exciting modes that display a similar vibrational distribution to the strike (Hall, 1991). A strike on a nodal line or point will not excite that mode. Where an impact occurs comfortably within a region of a natural mode, that mode is excited efficiently in the same (positive) direction, see Figure 3.2 (a), (b), whereas an impact that overlaps several modes excites all the modes within that region, and vibrations at different phases can cause cancellation, Figure 3.2 (c).

In the interaction between two membranes (heads) on a single drum (e.g. a snare

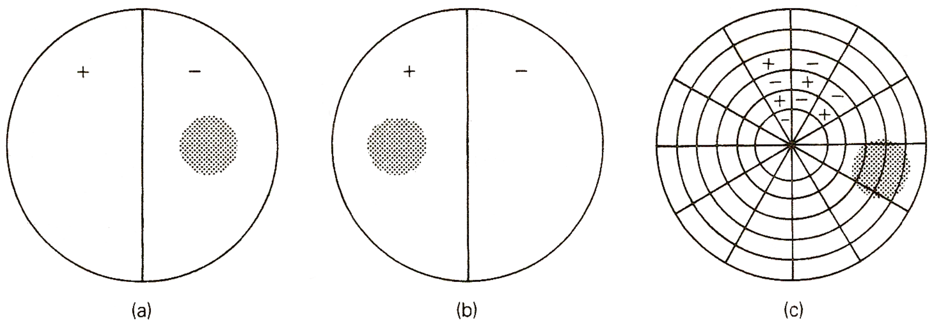


Figure 3.2: Strike locations and regional excitation efficiency (Hall, 1991, p. 75).

drum has a batter head [top membrane] and a resonant head [the bottom or snare head]), a coupling effect arises through either the enclosed air between the membranes, or through the shell (Rossing et al., 1992). The effect of air loading in an enclosed two-head system differs from a single head, in that while the air loading of single head will lower the modal frequencies, enclosed air loading raises the axisymmetric modal frequencies (Fletcher and Rossing, 1998, p. 75). Furthermore, coupled heads can either move in the same direction (in phase) or in opposing directions (out of phase) (Rossing, 1992; Tindale, 2004).

A comprehensive investigation into the properties of the bass drum sound was undertaken by Fletcher (1978), in which strikes of a bass drum were measured in an anechoic chamber, and compared with theoretical (ideal) values. These values were used to digitally synthesise bass drum tones with a subsequent listening test conducted with thirty-one adults to determine the real from synthetic tones. The main finding of this investigation was a general agreement between theoretical and observed frequencies, although some anomalies were reported. Other findings included how, for a hard blow, the frequency decreases over time, that the timbre is affected by strike strength and location, and that the decay rate is dependent on the characteristics of the drum rather than the location or type of strike. The listening test revealed that bass drum sounds played on tape through loudspeakers are different from that of an actual drum, although it is worth noting that this was a technical limitation.

Another aspect of the bass drum sound can be attributed to the moving parts of the bass drum pedal, which occurs in the pre-attack phase of the sound. The release of the mechanism during the decay of the bass drum sound could also contribute to a perceived difference in timbre should the sound be captured by microphones (Huber and Runstein, 2005)

One of the first studies to empirically measure the acoustic properties of a snare drum was Henzie (1960), whose focus was on the relationship between the amplitude and duration of snare drum tones, by using varying drum stick gauges with strokes at varying heights. The main findings of this study, showed that by equating the durations of snare drum tones produced at different heights in inches (e.g. 1, 2, 3, 4. . . 8) to note values at 120BPM, there is a ratio of 4:1 between strike height, and amplitude and duration. Interestingly, this experiment omits strike location as an important variable in either amplitude or durational characteristics. During this investigation, Henzie described how different variables could affect tone production in drums. The first variable described, was that of the drumheads themselves (e.g. size, thickness, tension, and age and condition). This is supported in an investigation on the tonal characteristics of snare drum heads by Lewis and Beckford (2000).

Lewis and Beckford's objective was to assist the practicing drummer in identifying the differences between fourteen different mainstream snare drum batter heads, from manufacturers such as Remo, Evans, etc. The experiment was done using a "gravity actuated "stick machine"" (Lewis and Beckford, 2000, p. 69), and a Fast Fourier Transform (FFT) that analysed the tones of the attack points in a two-dimensional (frequency × intensity) snapshot. Lewis and Beckford also drew attention to the importance of decay in selecting a drumhead and cite computational limitations as the reason for the omission of time-varying spectral analysis from the results. Nevertheless, the results show differences in frequency characteristics between drumheads, both within, and between manufacturers. It is worth noting that in both experiments the excitation location was static, and there was no investigation concerning the effect of using different drumsticks.

Henzie also describes how the construction of the drum (e.g. shell depth, materials and air ventilation) can affect tone production. Where coupling effects are propagated through shell vibrations, Rossing et al. (1992) note that the energy transfer from a snare drum shell to the stand increased the decay rate of the lowest (0, 1) mode, compared to a freely supported snare drum, where the decay rate of this mode was similar to other modes. It was also noted that drum shell mass had an "appreciable effect" on the decay rate, and subsequently the timbre (Rossing et al., 1992, p. 93).

Another variable described by Henzie to affect a snare drum tone was the construction of the snare. The snare is a series of metal wires that are strung across the resonant (snare) head, can be set to different tunings, and are engaged with the resonant head by using a "strainer" mechanism (Tindale, 2004, p. 20). When the batter head is struck, the resonant head moves due to the coupling effect, causing the snare wires to bounce and interact with the resonant head membrane, which produces the "snare" sound. Tindale describes how there is an "activation time" between the initial strike and the snare movement, although he asserts that a lack of investigation relating to elasticity and mass on the snares has failed to provide a mathematical representation of this phenomenon (Tindale, 2004, p. 21).

Citing a study by Bork (1983), where three blow strengths were applied to two snare tensions, (Fletcher and Rossing, 1998, p. 607) describe how a certain amplitude must be reached by the resonant head in order for sound production by the snares. This "critical amplitude" (Fletcher and Rossing, 1998, p. 604) is observed as increasing with snare tension. At a low snare tension, the medium strike strength exceeded critical amplitude, compared to the high-tension snare that required higher strike strength to reach critical amplitude. In addition, a damping effect was observed in low-tension snares.

One method used by percussionists to dampen the sound of a snare drum is the use of tape applied to the batter head, either across the diameter (Koblick, 2007, p. 26), or as a small piece of tape near the rim (Rogers, 2011, p. 7). This was computed by Yu and Wang (2001), with a later investigation into the placing of small strips of tape in arbitrary positions on a circular membrane undertaken by Yu (2004). This later study found that the fundamental frequency decreases as the small strip moves towards the boundary of the membrane.

Tom-toms are available in many different sizes, and have a variety of configurations for incorporation into a standard drum set. Some of the smaller tom-toms (eight to sixteen inches) are usually mounted on the bass drum via metal shafts, while the larger floor tom (typically fourteen to eighteen inches) is supported by legs mounted on the side of the shell (Sweeney, 2004b). Like other membranophones described in this section, the tom-toms are classified as un-pitched, although they do impart a perception of pitch. This sense of pitch is more prominent in tom-toms with only a single head (Fletcher and Rossing, 1998), compared to double headed tom-toms whose tendency is to produce "indefinite pitch" (Berg and Stork, 1982, p. 343). Describing how drum depth is a determining factor in tone quality, Berg and Stork state that double headed tom-toms with greater cylindrical shell lengths have "longer" standing waves, and as a result produces a lower tone than double headed toms with smaller drum depths. It is these differences in tone and pitch conveyance, that distinguish tom-toms from the bass and snare drums, allowing a drum set to have a wider range of timbres.

There are also variations in the tom-tom membrane, with some membranes being manufactured with large round dots in the centre, thus thickening the membrane in the centre. The acoustical effect of these dots is described by Fletcher and Rossing, as "shifting the lowest partials into a more nearly harmonically relationship" (Fletcher and Rossing, 1998, p. 606), resulting in greater perception of pitch. In addition to the change in harmonicity of the membrane, the greater thickness of the dot also increases all modal decay times. The effect of strike force on the spectral characteristics of different sized two-headed tom-toms was measured by Dahl (1997b), who found that an increase in strike strength resulted in an increase in modal excitations, including a change in spectral slope in frequencies above 1kHz, with a typical decrease of 9dB/octave as strike strength increases from soft to hard. Dahl, also notes a frequency shift with stronger strikes, and describes this frequency shift as a characteristic of loud playing, rather than a pitch glide as observed in bass drum tones (Fletcher, 1975).

The decay times of snare drums are affected by their support mechanisms, where decay rates decreased due to vibration transmission through the stand (Rossing et al., 1992, p. 89). The implications of this in relation to the tom-tom are significant, where different drum set configurations allow tom-toms to be mounted either on the floor, the bass drum via an arm mechanism, or via clamp to a cymbal stand, with each of these mounting configurations having different effects on the decay rates of the drums.

In the case of a tom-tom supported by an arm attached to the bass drum, a strong strike will transmit vibrations through the supporting arm mechanism, thus exciting either the shell or the membrane. In some instances, vibrations can cause rattles from lugs (Schroedl, 2003) and squeaks from moving parts, while other moving parts that can cause squeaks are necessary for drum striking (e.g. the bass drum pedal mechanism) (Huber and Runstein, 2005, p. 164), can also overlap the sound during the recoil of the mechanism for the next strike, affecting tone perception. Strong vibrations can also result in the sympathetic production of resonance (Jaffe and Smith, 1983, p. 66) of other supported tom-toms, in the case of a double tom stand on the bass drum, and bass drum, and depending on the construction and consequent sensitivity of the bass drum, both energy transmission and loss through the mechanism can affect the perception of the tone of the primary instrument (in this case the tom-tom). Tom-tom supporting mechanisms also exist for cymbals (e.g. crash cymbal stand), although it is worth noting that cymbals are more sensitive to vibrational transmission, and due to their nonlinear behaviour and higher frequency distribution, the resonant vibrations are more pronounced than the bass drum.

The production of rattles and sympathetic resonance is equivalent to leakage or "bleed" (Koblick, 2007, p. 26) during recording of the drum set, where a microphone on one instrument picks up the signal of another drum in close proximity (Huber and Runstein, 2005, p. 139).

3.1.2 The Vibrational Characteristics of Idiophones

One of the defining features of an idiophone is that the "vibrating material is the same object that is played (free from any applied tension)" (Schloss, 1985, p. 48), which includes xylophones, bells and cymbals etc. (Berg and Stork, 1982), and can be either tuned or un-tuned (Fletcher and Rossing, 1998). This definition is based upon the classification system proposed by von Hornbostel and Sachs (1914, 1961), which presents top-level classifiers based on excitation method, from struck, plucked, friction and blown (Benson, 2007, p. 91; Kartomi, 1990). Idiophones (cymbals) in the drum set relate directly to struck-upon percussion vessels (or sub-class 111.24) (von Hornbostel and Sachs, 1914, in Kartomi, 1990, p. 170). Physically, this subclass of idiophones can be described as circular plates (or mechanically as "a membrane with stiffness" (Fletcher and Rossing, 1998, p. 76)), where the plate is a two-dimensional primary and resonant vibrator, which can also become three-dimensional with the planar deviations of the plate resulting from any striking action (Moravcsik, 2001, p. 188). The shifting of the plate to three-dimensions is only relevant to plates with a free edge (e.g. a ride or crash cymbal), although there are different boundary conditions, such as clamped edges and simply supported edges (e.g. a hi-hat). The vibrational modes of a circular plate with the different boundary conditions are shown in Figure 3.3. Note the fundamental mode (2,0) in the free edge (a) compared to the (0,1) mode of a clamped or simply supported edge (b), and the similarities of the first four modes of (b), in comparison to the vibrational modes of an ideal membrane in Figure 3.1.

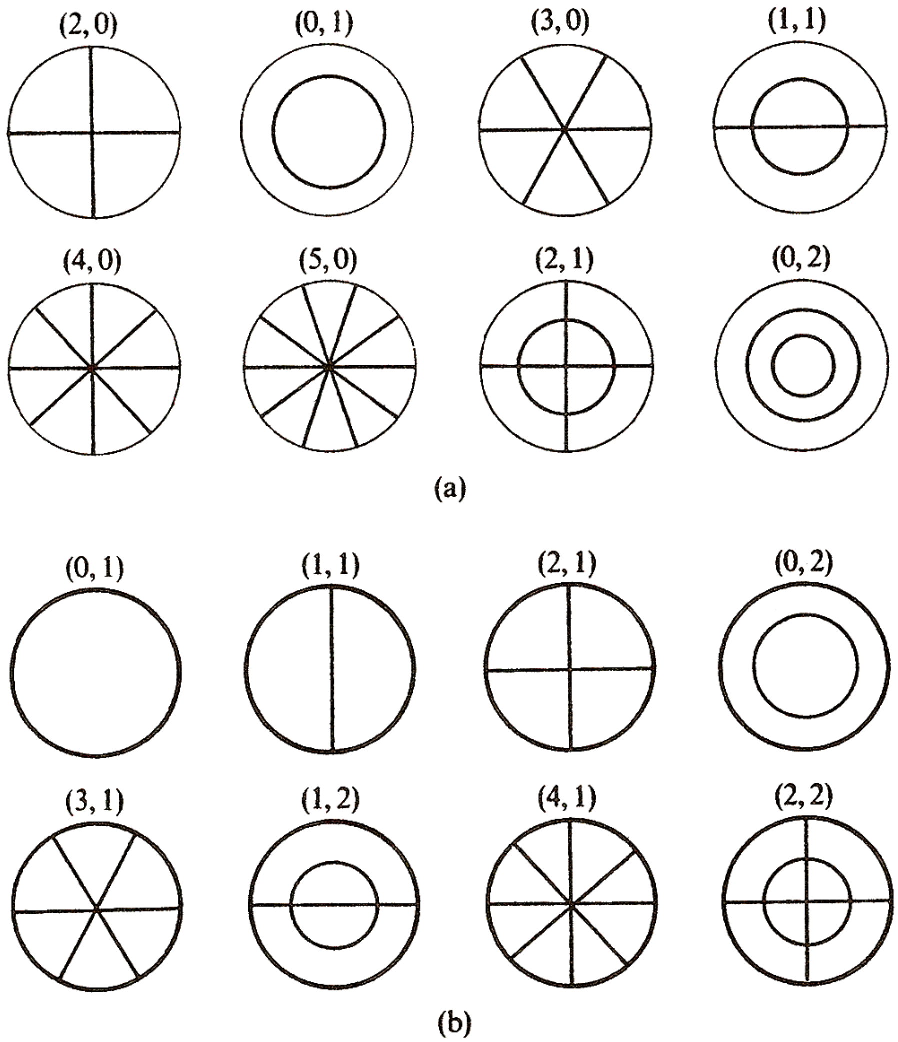


Figure 3.3: The vibrational modes of circular plates (a) free edge (b) clamped/simply supported edge (Fletcher and Rossing, 1998, p. 79).

Although cymbals display vibrational characteristics similar to circular plates, Fletcher and Rossing (1998) describe how there are differences in behaviour between low and high frequency modes, where lower modes are similar to a flat circular plate, and at higher frequencies, the modes merge together and become difficult to identify. Such behaviour can be seen in Figure 3.4, which shows the Chladni patterns of twenty-three vibrational modes of a crash cymbal and the corresponding vibrational modes in a flat circular plate using electronic TV Holography.

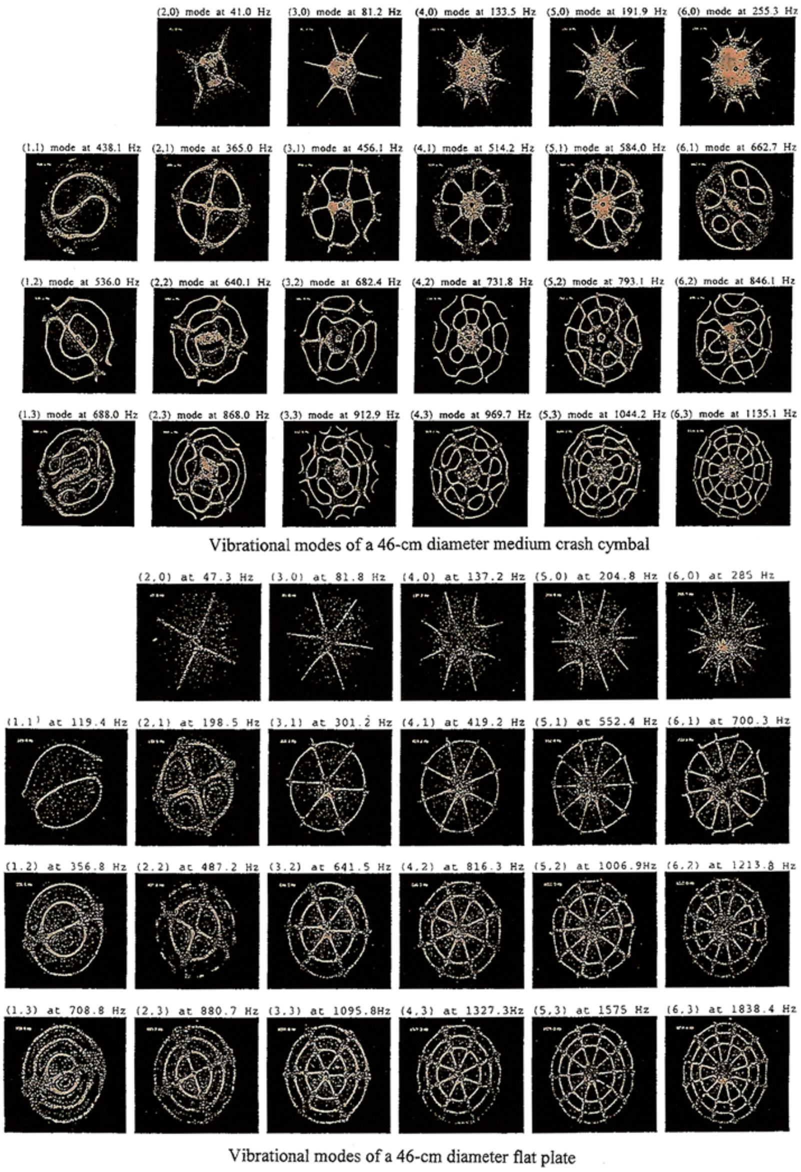


Figure 3.4: The first twenty-three vibrational modes of a cymbal and the corresponding modes on a flat circular plate (Wilbur and Rossing, 1997, as cited in Rossing, 2000, p. 90).

Although ride and crash cymbals possess similar mechanical characteristics, there are some immediate dimensional differences that affect the overall characteristics of the sound, resulting in different uses of these two cymbal types. This can be seen in Table 3.1.

INFLUENCE OF DIMENSIONS ON THE SOUND OF CYMBALS										
Dimens.	**Size**		**Weight**		**Profile**		**Cup**		**Taper**	
Infl. on	small	large	thin	heavy	low	high	small	large	even	gradual
Pitch	high	low	low	high	low	high	-	-	high	low
Volume	soft	loud	soft	loud	-	-	soft	loud	-	-
Response	fast	slow	fast	slow	fast	slow	slow	fast	slow	fast
Decay	short	long	short	long	-	-	short	long	long	short
Overtones	-	-	more	less	more	less	less	more	less	more
Ring	-	-	-	-	ringy	dry	-	-	tight clean	full explosive

Table 3.1: A chart showing the influence of dimensions on the sound of cymbals (based on single parametric changes only). Adapted from Pinksterboer (1992, p. 73)

Typically, the ride cymbal is thicker than other cymbals (Black, 2003), with a smaller "taper" (Pinksterboer, 1992, p. 72) (the thickness change from centre to edge), which produces more sustain, and is used mainly to play ostinato patterns (Black, 2003). In contrast, the crash cymbal is typically smaller and thinner (Black, 2003) with a larger taper (e.g. thinner edges) to produce more sustain, and is used primarily for accentuation and phrasing (Pinksterboer, 1992), thus being played less frequently and at a louder volume compared to the ride cymbal.[12]

Rossing (2000) notes that there are three key features in the sound of a cymbal. The first feature relates to the strike sound where after initial excitation, there is the initial wave propagation in the first millisecond. This is followed by a frequency increase at between 700-1000Hz for approximately 20ms, with the third phase of the sound at around a second afterwards containing frequencies of mostly 3-5kHz. It is this final phase of the sound that provides the "shimmer" effect (Rossing, 2000, p. 92).

Using double-pulsed TV holography, Schedin et al. (1998) captured the wave propagation through a cymbal. A laser pulse excited the cymbal at two points: one millimetre from the edge, and at half the radius. It was found that waves with longer wavelengths were more pronounced, and more reflected from the central dome and edge of

[12] For a comprehensive discussion on the history of and differences between cymbals, the reader is invited to read Pinksterboer (1992).

the cymbal, compared to waves with shorter wavelengths. This transient behaviour will occur irrespective of how the cymbal is excited, due to the nonlinear coupling of the vibrational modes (Fletcher and Rossing, 1998; Touze and Chaigne, 2000). The nonlinear behaviour produced harmonics, then sub-harmonics before becoming chaotic in nature. This can be seen in phase plots for a crash cymbal (Wilbur and Rossing, 1997, as cited in Rossing, 2000, p. 95). Describing how the nonlinear nature of cymbals causes frequential chaos as the initial strike excites a large number of frequencies in a short time, Touze and Chaigne used a Lyapunov exponent in order to successfully quantify the frequency transition from quasi-periodicity to chaos (Touze and Chaigne, 2000).

Other experiments into the vibrational characteristics of cymbals include an investigation into the effect of different sized and positioned holes on commercial cymbals (Peaden and Worland, 2011). This investigation found that the modal frequencies were affected, depending on the size and location of the hole in relation to the modal pattern, where the wave propagation is affected by the hole's lower mass and stiffness.

The hi-hat consists of two cymbals facing each other (Black, 2003, p. 24), typically with a thinner (lighter) cymbal on top, and a thicker (heavier) cymbal on the bottom (Pinksterboer, 1992, p. 71). These cymbals are mounted on a rod through the middle of the cymbals. The top cymbal is mounted to a foot-operated clutch that clamps the cymbals together, by lowering the top cymbal, thus providing an additional method of excitation and changing the boundary conditions of the cymbals to adjust the timbre. The spacing between the cymbals in the open position can be varied by using the clutch, from resting on the lower cymbal to completely devoid of contact altogether. Pinksterboer (1992, p. 88) describes how "a very tight clutch will deaden the sound of the hi-hat", while Black (2003, p. 24) suggests that the optimum space should be between one and two inches to allow for the closing of the hi-hat with the foot.

An open hi-hat displays the same characteristics as a typical cymbal with a free edge, with the exception that the mounting on the rod prevents the third dimensional planar deviation. In the clamped (closed) position the two cymbals have a damping effect on each other, which decreases the overall decay time, with the coupling of the cymbals causing vibrations that are normally only reflected from the edge (in the case of a free edge boundary condition), to be transferred into the edge of the counterpart cymbal. In addition, vibrations are also transmitted through the rod mounting, producing a damping effect. The amount of decay and vibrational transmission through contact is dependent on two variables: the strike strength, and the clamping force between the two cymbals. Both of these variables give rise to significant differences in sound. It is worth noting that there is currently very little literature on the physics of hi-hat cymbals mounted on a stand, under different damping conditions.

3.1.3 Configuration Characteristics

So far, this chapter has been concerned with the mechanics of the individual instruments of a jazz drum set. It has been noted that there are individual micro-variables that can contribute towards timbre and tone production with each of the instruments, although these are predominantly due to the materials and construction of the instrument, and the relationship between strike location and the mechanics of the instrument under ideal conditions. This section will discuss two, more global instrumental variables, both of which affect the resultant sound of a jazz drum set, and are variables that (beyond the strike location) are ultimately controlled by the performer in each instance of performance. These two variables relate to drum tuning, and drum set configuration.

3.1.3.1 Drum Tuning Uniformity

So far the vibrational characteristics of an ideal membrane in relation to the bass drum, snare drum and tom-tom have been discussed, in order to provide an overview of the complexities of tone production in membranophones. This also included a description concerning the differences between a real and ideal membrane. The complexity of these vibrational systems become compounded when considering how an ideal membrane is mathematically modelled using Bessel functions which assume a "uniformly stretched uniform circular membrane" (Bowman, 1958, p. 20) that inherently disregards the notion of membrane tuning dis-uniformity. Such dis-uniformity can occur in new membranes where striking stretches the membrane causing a perceived detuning (Schroedl, 2003). Tuning can be defined as "the process of adjusting a musical instrument such that the tones produced by the instrument obey certain relations" (Christensen and Jakobsson, 2009, p. 5). This definition, although defining relative tuning and aimed at stringed instruments (e.g. violin or guitar), is also relevant to membranophones due to the relationship between modes caused by cross-tensional forces. In addition, the material properties of idiophones cannot be easily adjusted (e.g. tightened)[13] to manipulate tonal production.

[13] This excludes the use of tape and other dampening techniques.

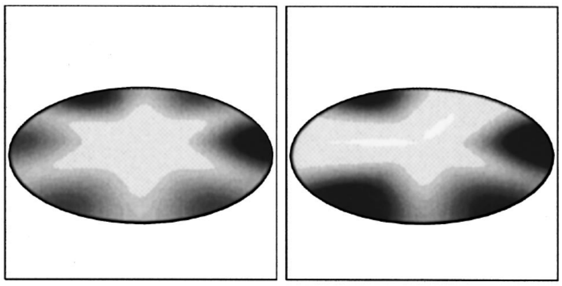
Drum tuning can be done in two ways: cross-tensional and clockwise (Black, 2003, p. 4; Sweeney, 2004b), although only cross-tensional tuning “maintains even tension throughout the tuning process” (Sweeney, 2004b, p. 7).[14] In an investigation into the acoustics of snare drums, Rossing et al. (1992, p. 85) notes the lack of standard practice for tuning snare drum heads, while emphasising how the investigation relied upon achieving the most uniform tension possible. However, drum tuning is not only important when undertaking empirical acoustic investigations, but is also important for live performance and studio-based music production. From a live performance perspective, repeatability of drum setup is important in tour situations, with tuning being important to tone quality and instrumental context in the recording studio (Toulson et al., 2008, p. 2). The investigation by Toulson et al. found that advanced musicians were able to understand and manipulate drum tuning, compared with amateur performers who appreciated the drum tuning, but could not tune their drums. Unsurprisingly, both Toulson et al. and Rossing et al. are in agreement regarding the lack of standard tuning practice, although Toulson et al. focuses on benchmarking different tuning setups for different musical genres, while the focus for Rossing et al. was on experimental validity.

A comprehensive investigation into the axial forces and in-plane displacements of snare drum heads of different materials during a tuning procedure, was done by Antonelli (2010), consisting of various quantitative measurement techniques. Antonelli found that the method of analysis was suitable for latex drumheads, but required further refinement for drumheads made from mylar. An acoustical analysis of the tuning of snare drums was undertaken by Richardson (2010), whose findings showed that detailed and accurate tuning was possible, and that modal frequency ratios that were previously considered fixed could be managed with tuning and damping to create new modal frequency ratios, thus creating a desired tone.

An early mathematical investigation into the vibrations of circular membranes with “non-uniform tensile forces at the edge” was done by Mei (1969, p. 693), who attempted to identify the vibrational behaviour for a non-ideal membrane using a finite element method. In this investigation, Mei found that lower modes were the same as an ideal membrane, although nodal patterns associated with higher modal frequencies became distorted, suggesting that non-uniform tuning can have an effect on the timbre of drum. This is apparent in later research into the simulation of a kettledrum by Rhaouti et al. (1999) where during a comparison of simulations and experiments,

[14] For an overview on drum tuning, the reader is invited to read Schroedl (2002).

Rhaouti et al. states that: "the simulations were used in order to check whether or not this feature [beating] is due to imperfect tuning of the membrane, as it is usually assumed" (Rhaouti et al., 1999, p. 3556). In order to quantify this assumption, Rhaouti et al. simulated a kettledrum at both uniform and non-uniform tension, presenting a comparative example of tension distributions, shown in Figure 3.5. The authors also note that an adjustment in membrane tension within the model could yield similar results to modifications in other parameters of the model.



(a) (b)

Figure 3.5: A comparison of (a) uniform and (b) non-uniform tension distribution. Higher areas of tension are shown in white, lower areas of tension are shown in black (Rhaouti et al., 1999, p. 3556).

In order to assist the approximation of an ideal membrane during the tuning of drums, Worland (2008) analysed the modal patterns of a single-headed drum under non-uniform tension, using electronic speckle pattern interferometry (ESPI) in order to "image the mode shapes on the drumhead and identify corresponding frequencies" (Worland, 2008, p. 5). In one experiment, the tightening of a lug by two turns caused the (1,1) mode to split and curve away from the tuned lug, as the opposing lug retained the original tension (e.g. a two-fold perturbation), thus creating "perpendicular fast and slow axes on the membrane" (Worland, 2008, p. 11). Worland also notes that the modal curvature due to irregular tuning is not "directly related" to the frequency splitting (p. 7), and that higher modes (those outside of the investigation) "can be split by higher order perturbations in the applied tension" (Worland, 2008, p. 11).

Further research by Worland (2010) saw the expansion of this approach to include time-averaged ESPI on the (1,1) mode. Worland then created a generalised model to encompass all other modes, and found agreement with the model in representing frequency splitting in modes from differently applied tension. The principal findings of this investigation found that the splitting of the (1,1) mode via a two-fold perturbation was the "largest contributor to the sound of a drum not being in tune with itself" (Worland, 2010, p. 533).

3.1.3.2 Drum Set Configuration

The typical drum set configuration has constantly evolved since the earliest drum sets in approximately the 1900s (Starr, 2009, p. 263), with changes resulting from either economic drivers, or through stylistic changes in musical tastes over the decades (e.g. Be-bop jazz to Rock) (Aldridge, 1994, pp. 28-30).[15] A typical modern drum set configuration is described in Huber and Runstein (2005, p. 162), Murakami and Miura (2008, p. 450) and Strong (2006, p. 13), and consists of only the basic elements of a drum set, compared to extended configurations that cover "all of the instruments potentially used" (Murakami and Miura, 2008, p. 450; illustrated in Murakami and Miura, 2008, p. 451). Taking these configurations into account, Strong (2006, p. 65) describes left and right handed configurations, where in a right handed position, the hi-hat is struck with the lead (right) hand, and the bass drum with the lead (right) foot. Within different configurations of drum set, there are many ways to arrange the drums, and is usually based upon personal preference. However, the main criteria for positioning the drums are comfort, ease of use, and injury avoidance (Starr, 2009, pp. 12-13), although Black describes how positioning to "minimize reaching, stretching and twisting" is dependant on the performers' "physical size and technical ability", and how correct positioning "will help to assure optimum sound quality and volume while minimizing the possibility of damage to the cymbal" (Black, 2003, p. 22).[16]

3.1.4 Simulating Percussion using Sound Synthesis

Timbral perception plays two very important roles within sound synthesis. Firstly, perception can be causal. Perceptual attributes can be used to assist in new processes for synthesizing sound, or to shape synthesis parameters, and can affect the efficiency of the implementation. Secondly, perception can be affective. Perception can determine the quality of the synthesis by its similarity to the modelled instrument. Owing to the existing body of literature pertaining to perception, synthesis and analysis, this section will provide only a brief overview of the subject, and identify some key considerations

[15] For a comprehensive history of the drum set, the reader is directed to Aldridge (1994).

[16] For further discussion on the proper positioning of a drum set, the reader is invited to read Starr (2009, pp. 12-14).

that must be taken into account when evaluating different sound synthesis approaches.[17]

Chowning (2000) provides a concise history of the convergence and importance of perception within the fields of acoustics and synthesis beginning with the 1960s drive for perceptual studies brought about by computer limitations and high cost (causal), and the need to understand sound from a design perspective (affective). Citing Mathews (1963), Chowning (2000, p. 1) reflects how auditory studies addressed causal motivations more so than affective reasons, largely due to the tangibility of computational constraints. However, the analysis by synthesis approach meant that perceptual models could concentrate on the most perceptually relevant features of the sound, which inherently reduces the computational overhead (Risset and Wessel, 1999). In cases where data reduction is still required to reduce the computational overhead, there is a reciprocal effect on the perception of the sound. Charbonneau (1981) investigated the effect of data reduction along three dimensions: amplitude, frequency and time, and found that data reduction in amplitude values had the most significant effect to the timbre, followed by frequency, and then time.

Using perceptual parameters as synthesis mapping (control) parameters enables improved perception of expressivity in synthesis, and ensures that the parameters are "strong or powerful" parameters that make audible differences to the sound (Jaffe, 1995). Interestingly, Jaffe describes how synthesis techniques with more parameters, tend to have weaker parameters. In other words, changes to the parameter are less audible. Although it is widely accepted that timbre is multidimensional, the parameterisation of perceptual attributes of timbre is not exact. Furthermore, in some synthesis techniques there are more dimensions of timbre than parameters available. The challenge here is to evaluate the "stronger" perceptual dimensions, and prioritise their mappings, for example, spectral centroid and rise time (Lakatos, 2000). However, mappings by perceptual parameter have compositional implications. Unwanted side effects in using perceptual parameters include dimensional incongruences; for example, spectral centroid and inharmonicity. A spectral approach makes for interesting compositional methods. On the one hand, it enables the exploration of an instrument's timbre through a perceptual lens; on the other hand, it challenges the listener's conventions of the instrument, and challenges the listening strategy of the listener (e.g. transitioning

[17] For a more comprehensive discussion, the reader is invited to read Beauchamp (2010); Butler (1992); Cook (2002a); Deutsch (1999); Donnadieu (2010); Hartmann (1997); Howard and Angus (2009); Pierce (1992); Risset and Wessel (1999); Roederer (2008); Sethares (2005) and Warren (1982).

between causal and reduced listening modes, Chion (1983/2009)).

The listening state transition in the perception of an instrument is not limited to the timbre. Correlations have been made between the physical nature of the instrument (materials, dimensions) and perceived nature of the instrument (materials, dimensions) (Aramaki et al., 2006). These studies range from the perception of material (Klatzky et al., 2000), shape and material (Kunkler-Peck and Turvey, 2000), size, shape and material in damped and free vibrating plates (Tucker and Brown, 2002), mallet hardness and applied damping forces (Lutfi and Liu, 2007) to tonal interaction with the environment and the perception of the instrument in that environment (Roederer, 2008). The psychological representation of an instrument includes many variables, and is an extremely complex subject. For this reason, it is more useful here to consider the perceptual implications of the instrument's timbre to be exclusive of environmental factors, and ignore the perception of physical properties. Considering these perceptual implications, a practical approach to evaluating physical and spectral modelling techniques will focus on their computational implementation against the expressive variation of the technique. Such a framework for evaluation is proposed by Jaffe (1995) who describes ten criteria of which, only five are applicable in this context:

1. **Robustness:** e.g. does the synthesis sound similar with expressive variation?
2. **Efficiency:** The efficiency of the algorithm, in terms of:
    - Memory (RAM/Hard Drive Space)
    - Processing (CPU power)
    - Control stream e.g. how are synthesis calculations implemented within the system architecture?
3. **Sparseness of control stream:** e.g. which aspect of the synthesis is doing the most work - the synthesis, or the control data?
4. **Sound class representation:** e.g. whether a sound can be created with a single, hybrid or multiple techniques;
5. **Latency:** e.g. does the computational method prevent immediate playback?

Ideal synthesis techniques have a robust identity, efficient algorithms that are good on memory, CPU, with a sparse control stream, and can represent multiple sound classes, with minimal latency. This section will evaluate these approaches, with a view to selecting the most suitable method for sonically rendering the complete instrument set.

The Karplus-Strong technique maintains a sound's identity with expressive variation for both plucked string and drum timbres. However, in the context of the discussions in Chapter Two, the amount of timbral variation available in this technique is limited due to few control parameters being available. As a result, it is difficult to create an accurate representation of producible timbres caused by different modal excitations. Although it is computationally efficient in its basic form, as demonstrated by the limited system resources of the computer in the initial implementation, there are two factors that could affect the efficiency of this algorithm. With between three and seven active degrees of freedom in cymbal vibrations, a large number of equations are required for physical modelling in order to model the vibrational characteristics of cymbals (Rossing, 2000, p. 96).

As a result, this method will not sufficiently model these active degrees of freedom without significantly changing the method. The number of oscillators and independent random parametric variations required to create an accurate representation of a cymbal sound, degrades the performance from the original real-time implementation and makes the control stream very dense. As a result, the sound class representation of the drum set is a significant problem for using this technique, although an alternative is to combine it with other techniques (e.g. using filtered white noise to simulate a cymbal).

Karplus-Strong Extensions employ additional filters to simulate the damping of a string, particularly the plucked string of the original Karplus-Strong implementation. The Karplus-Strong Extensions fall victim to many of the problems of the original technique, although with the additional filters, simulation of the chaotic nature of cymbals with time-varying filters is possible. However, the number of filters and subsequent use of multiple parameters per filter needed, will not only increase computational overhead, but will make interdependent expressive control difficult.

Digital waveguides are a popular synthesis technique because of their ability to realistically reproduce complex resonators. This is particularly true with one-dimensional models (a plucked string or wind instrument) as the required equations are computationally more efficient than calculating the wave equation (Murphy et al., 2007, p. 59). However, the two-dimensional waveguide mesh typically used to model the wave propagation in membranes, is less efficient due to the number of calculation steps required in the scattering junctions to calculate the wave propagation, particularly when boundary reflections are calculated. This is because numerical errors are more likely to occur as filters simulate losses in wave propagation. Reducing the number of scattering junctions increases the computational efficiency, and in the case of the Digital Waveguide Mesh (DWM) it "reduces the sample rate" (Murphy et al., 2007, p. 59).

Another limitation of digital waveguides is dispersion error. Dispersion error is where a waveguide mesh derives an unexpected mode due to miscalculation of the velocity and direction of the wave propagation (Murphy et al., 2007), which is particularly relevant when considering the effects of drum tuning dis-uniformity on wave propagation and modal excitation. Although existing implementations are robust, there are difficulties in mathematically representing non-ideal drum tunings without creating a dense control stream. Additionally, the real-time implementation is constricted by the resolution of the equations. In the case of a nine-piece jazz drum set, real-time modelling of all instruments via this method is possible but difficult to implement.

Modal synthesis models the modal and harmonic interactions of a system, from the input excitation. As a result, modal synthesis has a very sparse control stream, and is generally robust with expressive variation as it is typically defined by the excitation. However, this method is computationally very expensive when representing instruments with high numbers of modal frequencies, such as membranophones and idiophones. This is exacerbated in the case of modal interactions for non-ideally tuned membranes, where more modal data is required. Expanding this technique to include all nine instruments of the jazz drum set is too computationally inefficient to implement.

The finite-difference time-domain (FDTD) method was noted to accurately reproduce cymbals and other percussion instruments, although the performance of the method for complex vibrations such as cymbals, and higher sample rates, required between 1 to 160 seconds to render (Bilbao, 2010, p. 879). This is too slow for a real-time implementation, particularly where multiple re-rendering was required to simulate different excitation locations.

Additive synthesis is the creation of complex timbres using multiple sine waves from multiple oscillators. Additive synthesis is particularly useful for synthesizing sounds with little noise. As drums, particularly cymbals are very noisy and contain lots of partials, a significant number of oscillators with independent timevarying envelopes will be required (Roads, 1996, p. 141), making the control stream dense and significantly increasing the computational overhead. Similar to the Karplus-Strong (and Extensions) this would require complex control techniques to simulate percussive variation.

Linear predictive coding (LPC) can be applied to percussion by separating the excitation from the resonance, and predict the resonances from the inversely filtered input signal. One of the drawbacks of using LPC for modelling drums is that the analysis component involves source separation and subsequent filtering, and the input data should represent the micro-timbral variations possible on a drum. In the case of Sandler (1990), the input source used was pseudo-random noise; for a nine-piece jazz drum set, this would require a significant amount of input data. This affects not only the computational overhead but makes the control stream sparse. Furthermore, the robustness of the expressive variation is not good with LPC, as noted by Sandler where there is a trade-off between sonic flexibility and precision of recreation. Parametric fine-tuning would be required to ensure that there were significant variations in output, particularly in cases where small micro-timbral deviations are detectable in the input signal, but the resonant effect in the synthesised output is not predicted as expected. Issues arise with latency, where the computational overhead is affected by simultaneous analysis and synthesis of the input data.

In order to create drum sounds using wavetable synthesis, different wavetable techniques are required. This is because drums contain an attack transient and are complex time-varying sounds. As a result, an ADSR envelope must be used to create the attack transients, and wavetable stacking must be done in order to create the complex timbre. Owing to the time-varying nature of cymbal spectra, the stacked wavetables would also need to be crossfaded so that the timbres change over time. In addition, the resultant output of a wavetable is largely dependent on the signal stored in the individual input wavetable, meaning that choice of the input signal is important, and because of the need to store multiple input signals, contributes to an increased computational overhead. In addition to wavetable input selection, the parametric control of an ADSR envelope to each wavetable, and the parameterisation of the wavetable crossfading are also issues for this implementation.

In terms of expressive variation (PCM) sampling synthesis (see section 3.4) is the least robust. Conversely it is also the most accurate and can model any sound class. Samples can be processed in real-time by adjusting gain (amongst other things), but because the sample is fixed, changes in the spectral evolution and timbral variations associated with vibrational changes due to differences in playing do not correspond well to changes applied to the whole sample. To overcome this, multisampling techniques have been used, although these come at a high cost in terms of computer memory and hard drive space. A reduction in sample rate, or in the number of samples in a multi-sampled system, can mitigate the computational overheads although this can affect the quality of the sound. Despite this, the playback of recorded samples is both computationally efficient and there is generally low latency in the playback. The largest problem with the sampling lies in the control parameters, where playback usually consists of “play” and “stop” messages. As a result, the control mapping for sample selection can be complex and highly abstracted. With PCM sampling being the chosen synthesis technique to be used in this investigation, these limitations will need to be addressed.

3.2 Why is Modelling Human Performance on a Jazz Drum Set so Difficult?

In the first part of this chapter, empirical research into the vibrational characteristics of membranophones and idiophones in a typical jazz drum set were reviewed. Differences were highlighted between the instrument groups in terms of modal excitation and decay rates due to material, construction and supporting mechanisms. In addition, differences across the surfaces of the instruments were shown, characterised by the tuning and strike location. Of particular relevance, were differences in modal excitation arising from variations in strike location, which produce micro-timbral variations. Understanding the causes of variations in strike location is important for constructing the model’s timbral representation.

The interaction between the player and instrument is perhaps the most significant variable to affect timbre production. It is the most significant variable in the distinctiveness of performance. This is manifested in different techniques, skill levels, musical knowledge and experiences, and the physical attributes of the performers themselves (e.g. height, body mass, fitness etc.). In fact, the act of musical performance encompasses a variety of different contributory aspects. More specific examples of this include physiological (Fujii et al., 2009; Lee, 2010), cognitive (Dahl and Friberg, 2004; Laukka and Gabrielsson, 2000; Repp, 1999), technical (Dahl et al., 2011) and musical (Repp, 1997), and include both theoretical and empirical perspectives (Shove and

Repp, 1995).[18]

This section will describe a drummer's development goals, and then present an analytical framework in which to understand the underlying causes of performance variation, particularly regarding instrumental interaction. This will involve a summary review of the literature in the discipline of biomechanics, and the subsequent application of these principles in relation to percussive performance, in order to augment the theoretical framework for the underlying performance modelling paradigm. The following sections will deconstruct the principal physical factors that cause imperfections in performance, and their combined effect with instrumentation mechanics on the timbre of the sound. It is not the objective of this investigation to discuss the different options and timbral and acoustical effects of striking implements (see Halmrast et al., 2010, pp. 204-207), nor is it intended to be an exhaustive discussion. There are specific aspects that have been omitted, including the effect of batter head models on timbre (Henzie, 1960; Lewis and Beckford, 2000);[19] the effect of non-uniform tension; potential tonal evolution due to the age (and usage) of the head; tempo (Desain and Honing, 1993); feedback conditions (Brandmeyer et al., 2011; Dahl and Bresin, 2001; Pfordresher and Palmer, 2002) and temporal independence (Goebl, 2011). In addition, aspects such as style and genre which, with their obvious contextual performance differences, will not be discussed in detail.

3.2.1 Drum Rudiments and Development Goals

Drummers develop their technique by learning drum rudiments established by the international drum rudiment committee, part of the Percussive Arts Society (PAS). The rudiments currently consist of 40 techniques (PAS, 2014),[20] and have been derived from different musical styles to form a pedagogical method for learning percussion. This method is designed to provide an "orderly progression for the development of physical control, coordination, and endurance" (see the foreward in Carson and Wanamaker, 1984). Although not explicitly defined, these development goals can be interpreted and summarised as follows:

[18] For an overview, the reader is referred to Gabrielsson (1999, 2003) and Palmer (1997).

[19] Although this is not a real-time timbral phenomena, it certainly has an effect, and can be considered as important as the striking implement, which might change during a performance.

[20] Drum rudiments were initially created by the National Association of Rudimental Drummers (N.A.R.D.) in 1933, and subsequently expanded from 26 rudiments by the PAS. In addition, the 40 rudiments are considered to be a work in progress, to enable future development (Carson and Wanamaker, 1984).

- **Physical control:** refers to the performers' management of stick and instrument interaction, extending to the wrist, hand and including arm control;
- **Coordination:** refers to the strike accuracy and the performers' ability to exert physical control over sequences of strikes in different locations; and,
- **Endurance:** refers to performer attributes, instrumental configuration and the complexity of piece being performed.

Although these development goals can be considered independent of each other, there is a large amount of interdependence between the three. One example of this is where a performer has good stick and instrument management but poor coordination. The result is a drummer that could play the strikes correctly but not necessarily hit the drum in time. Another example is maintaining arm control and coordination of movement in complex sequences, which depending on the endurance levels of the performer, can deteriorate over different periods of time. The relationship between these goals is described in Figure 3.6. Although these goals are fundamental to the development of

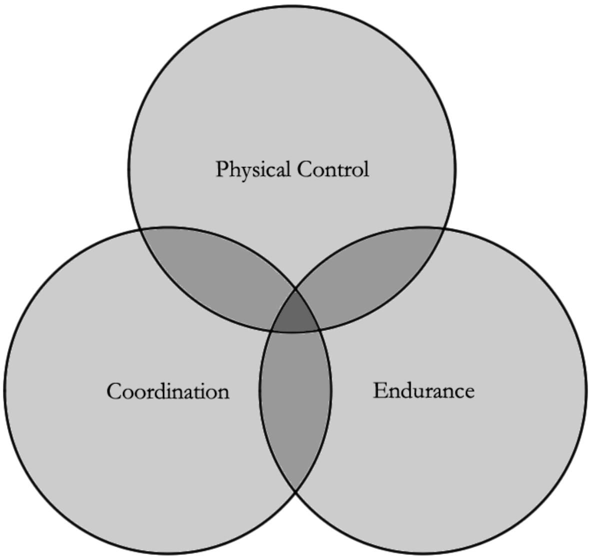


Figure 3.6: The interdependency of the three development goals. Adapted from Carson and Wanamaker (1984).

a percussionist's skill, obtaining an understanding of percussive performance by way of deconstructing principles of human movement from these goals is difficult due to the effect of environmental factors in skilled movements (Dahl, 2005). Such factors could include the effect of temperature and altitude on endurance, auditory feedback on coordination, and stick thickness on physical control, amongst others. As a result, it is both impractical and difficult to account for all of these independent variables. The reduction of independent variables in the analysis of human movement, extending to environmental variables, is not new. In fact, the dimensionality of variables in understanding human movement has been the subject of investigation since Nikolai Bernstein first proposed the theory of the degrees of freedom (DOF) in 1967. He theorised that because there are an almost infinite number of ways a movement could be executed through the large network of muscles, joints and cells in the human body, there are an infinite number of ways that muscles can achieve the different movements.

The control of the nervous system on the musculoskeletal system is highly complex as for any given movement there are a high number of degrees of freedom. This complexity is illustrated during the activation of a single muscular element in either isolation or in any particular sequence from another (Bernstein, 1967). Thus, if the nervous system controls movement by controlling synergistic groups rather than individual muscles and joints, the number of DOF (and therefore the dimensionality of variables) is reduced (Turvey, 1990). Bernstein also suggested that the sensory feedback from the environment interacted with the nervous system to reduce the number of DOF. Turvey substantiates the omission of environmental factors within the context of Carson and Wanamaker's development goals for this framework, by arguing that:

> "if the environment to which the movement system relates is interpreted as just another large set of variables, then the juxtaposition of an animal and its environment would amplify the problem of degrees of freedom" (Turvey, 1990, p. 940).

Including the juxtaposition of environmental factors on percussive performance, would not only concern human movement and the number of degrees of freedom, but would also necessitate extending environmental variables to the vibrational behaviour of each of the nine percussion instruments under investigation. Since the speed of sound increases with air temperature (Fletcher and Rossing, 1998, p. 70) a bigger picture emerges as to the inherent difficulty in adequately applying several environmental factors as variables across the different themes of this thesis. In light of this, Turvey's position will be considered to be the most appropriate view, and consequently environmental factors will be considered outside of the scope of this thesis in relation to instrumental mechanics and human movement.

3.2.2 The Analysis of Human Movement

In 1982, the neuroscientist David Marr presented a tri-level hypothesis by which information processing systems should be analysed. These levels of analysis can be summarised as follows (Marr, 1982, p. 25):

- **Computational level:** what does the system do?
- **Algorithmic/Representational level:** how does the system do what it does?
- **Physical level:** how is the system physically realised?

Marr describes how these three levels of analysis are not intrinsically dependant on one another and how, in some circumstances, analysis can be achieved by using only one or two levels. The choice of analytical level is critical in correctly understanding certain systems. More importantly, Marr continues to describe how the computational level of analysis is particularly important in understanding certain phenomena, particularly where there are significant levels of abstraction between the understanding of a system and the computational representation. Examples of this include a priori understanding of the nature of biological or perceptual processes, prior to computational representation, than by analysing the computational representation of such process in a given computational environment (Marr, 1982, p. 27).

In David Rosenbaum's (2010) book on motor control, he describes how Marr's three analytical levels of information processing systems also represent "the study of human motor control" (Rosenbaum, 2010, p. 4). At the computational level of analysis, Rosenbaum describes how, during physical activity, animals and humans plan their movements using what he describes as "implicit equations". These implicit equations are derived from Marr's computational level, where a system must achieve a function, which in computer terms is often described mathematically. However, for humans and animals this refers to the mental representation of task to be performed. One example of this is the mental representation a rock climber has of a "dyno" (a jump or leap) to the next position. In the context of percussive performance, this can be a mental representation of an impending drum fill and the striking of a sequence of instruments from the current position. Rosenbaum notes that the computational level of analysis does not include the execution of the action, which is unsurprising considering the number of DOF.

In applying Marr's second level, the algorithmic/representational level, Rosenbaum notes that algorithms in a computer system are designed to enable a system to undertake their function, with guaranteed success. In the natural world, movements operate in real-time (analogous to algorithms at runtime) but there is no guarantee of success. For example, a rock-climber might not jump high enough to grab the next hold (and fall to the crash-mat below); the drummer can hit the wrong instrument or strike the shell of the drum by accident. As Rosenbaum points out, each of these real-time movements relies upon a procedure, and the person executing the action will draw upon behaviour and cognition in order to execute and verify the movement, hence Rosenbaum's extension of this term as the "procedural level" (Rosenbaum, 2010, p. 5).

Rosenbaum describes the final level of Marr's analysis, the implementation level, biologically as the physical aspects of the movement. Such physical aspects are described by Rosenbaum as muscle operation, brain activity etc. It is in this level that our rock climber can use their leg muscles to jump, stretch their arms to grab the hold and, finally, use their fingers and forearms to grip and maintain the hold. For the drummer playing a snare drum followed by a ride cymbal with the same hand, muscle operation can include the fingers and hand for gripping the stick, adduction of the lower arm for the strike, followed by a lateral rotation and abduction of the arm to reach cymbal height. These examples are highly simplified, as it is in this analytical level that the DOF problem is encountered.

Rosenbaum's biological adaptation of Marr's tri-level analysis provides a solid approach to understanding the movement process. If this three-stage analysis is undertaken in the context of Carson and Wanamaker's development goals, it is possible to objectively evaluate existing research and literature on human movement, specifically for percussion. Furthermore, the bottom-up nature of the three analytical levels in relation to performing a drumming action, allows for a more comprehensive and structured discussion. This re-contextualisation is described in Table 3.2. Understanding the nature of percussive performance variation requires only the computational level of analysis in order to gain an understanding of the relevance of human performance on timbre and timing, and uncover common aspects of human performance. Although there are other aspects in other levels that contribute to performance variation, the majority of these are outside of the scope of this investigation due to their highly individual and highly subjective nature, and the difficulty in adequately proving these. These parts of the framework are highlighted in light grey.

Level/Goal	Physical	Coordination	Endurance
Computational	Planning the control of: (a) The physical movement; (b) Instrumental interaction;	Planning coordinated movements: (a) Coordinating simultaneous multiple physical events; (b) Multiple instrumental interactions.	Planning movement for improving endurance: (a) Economy of movement.
Procedural	The behavioural and cognitive aspects of carrying out a physical movement, relating to: (a) Timbre; (b) Timing.	The behavioural and cognitive mechanisms for: (a) Measuring current position; (b) Verifying next movement; (c) Anticipating next timbre/timing.	The behavioural and cognitive aspects of improving endurance: (a) Performance psychology.
Implementation	The physical aspects of carrying out a movement: (a) Muscle activity; (b) Brain function.	The physical aspects of coordinating multiple instruments: (a) Inter-limb coordination; (b) Muscle activity; (c) Brain function.	Physical ways of improving endurance: (a) Training; (b) Warm up protocols; (c) Performer impairments.

Table 3.2: Three analytical levels applied to a drummer's development goals. Adapted from Carson and Wanamaker (1984) and Rosenbaum (2010).

The first aspect of this framework outside of the scope of this investigation is the behavioural and cognitive aspect of carrying out a physical movement (physical/procedural). This is because behaviour and cognition are highly individual, and also highly dependent on the context of the performance (e.g. genre). An important cognitive element of this analytical level and context comprises sensorimotor synchronization (SMS), which is the rhythmic coordination of an action with a regular external event (Repp, 2005). As a result, the computational representation of SMS would be difficult to realise, and the empirical testing required for such a model is outside the scope of this investigation.[21]

Another area of the framework outside of investigative scope is the physical aspect of carrying out a movement (physical/implementation), particularly regarding muscle and brain activity. This particular area presents two separate problems. In terms of muscle activity, the most significant modelling challenge lies with the DOF problem and determining which classifiers and representative organisational systems of muscle activation to model. One such solution would be to use a single DOF as a representative for all similar movements in the model. In the case of a drummer, there would be more than one DOF that would need to be modelled to cover all limbs. In addition, determining the most appropriate DOF for the movement, and making such assumptions, will produce some theoretical shortcomings (particularly for neurophysiologists). Modelling muscle activations also presents problems in relation to the relationship between abstracted models of muscle movement and timbre production - a problem that is also found in modelling brain activity.[22]

The behavioural and cognitive mechanisms associated with performance feedback (coordination/procedural) encompass a range of methods of feedback acquisition. These include auditory, visual, tactile, haptic and kinaesthetic, or combinations of one or more. Each of these individual types of feedback has different effects on cognitive and behavioural mechanisms, and varies depending on the performance conditions. With so many combinations of feedback conditions and environmental variables, finding an appropriate representative model is difficult. Additionally, modelling specific effects of

[21] For further reading on this subject, the reader is invited to read Fujii et al. (2010); Hove et al. (2007); Repp (2005, 2006); Wing et al. (1989) and Wing and Kristofferson (1973a,b).

[22] For further reading on muscle activation and brain activity during performance, the reader is invited to read Fujii et al. (2009); Fujii and Moritani (2012a,b); Gabrielsson (2003) and Todorov and Jordan (2002).

certain feedback conditions would have limited practical application.[23]

It was noted above that modelling muscle activity was difficult given the DOF problem, the high level of abstraction from timbre production and timing of both muscle activity and brain function, and the selection of suitable organisational systems for modelling control and muscle activation. This problem is compounded when considering inter-limb coordination as a physical aspect of coordinating multiple instruments (coordination/implementation), particularly in complex tasks such as rhythm production. In creating complex rhythms bimanually, task difficulty between the hands (which include cooperative and disjointed tasks), together with the dexterity levels and handedness of the individual, will affect the brain's organisational control of the two hands. In the case of drumming, it is more likely to include leg control for operating the bass drum and hi-hat. This would result in a highly complex study with many variables that may or may not be relevant to performance modeling. Consequently this is outside the scope of this research.[24]

Endurance is unique to individuals and can be increased with correct training. However, during performance, endurance can be affected by an individual's level of physical exertion, which can be mitigated by employing techniques that increase their economy of movement. The behavioural and cognitive aspects of improving levels of endurance fit firmly in the realms of performance psychology, which are difficult to represent in a computational model of a performance. Similarly, the modelling of training and warm up protocols is also outside of the scope of this investigation, as they do not bring any direct benefit to the modelled system. Similarly, no benefit would be gained by modelling a performer with an impairment because, like modelling a drummer with low levels of endurance, the system would be designed with a level of performer obsolescence, resulting in poor playing after a period of time. Therefore, computational, procedural, and implementation levels of analysis relating to endurance are outside the scope of this investigation.[25]

[23] For further information on aspects of performance feedback, the reader is directed to Brandmeyer et al. (2011); Dahl and Bresin (2001); Fujii et al. (2010); Gabrielsson (2003); Petrini et al. (2009); Pfordresher and Palmer (2002) and Pfordresher and Benitez (2007).

[24] For further reading on this subject, see Bernstein (1967); Calvin et al. (2010); Iannarilli et al. (2013) and Kelso et al. (1979).

[25] However, for further reading the reader is directed to Abernethy et al. (2005); Gabrielsson (1999, 2003) and Shaffer (1989).

It is worth noting that although some aspects of the framework are specifically noted as being outside the scope of investigation, there are overlaps between some of the variables mentioned and aspects of performance that will be discussed in the following sections. Their inclusion within the discussion serves to highlight the complexity of percussive performance, and demonstrates the wide reaching implications and importance of the discussion.

3.2.3 Controlling Instrumental Interaction

Why is physical control so important? Striking an object with another object has two repercussions. Firstly, as the struck object produces sound, vibration in the stick travels through the fingers to the hand. In some instances, and depending on the force of the strike and the materials involved, this can extend into the arm. In severe cases this can cause discomfort (e.g. striking with extreme force a large mass of solid metal using a metal bar). Secondly, striking an object can cause the striking tool to be deflected away from the surface and, depending on the elasticity, the level of deflection will be either minimal (e.g. a hard metal struck surface) or more significant (e.g. a struck membrane under tension). Since playing the drums requires striking many objects consisting of different materials at different strengths, the amount of vibration experienced in the player's body differs between instruments. In addition, each instrument will deflect the striking implement differently depending on the elasticity of the struck surface, the angle of the initial strike, and strike force. It was noted earlier in the chapter, that strike location played a significant part in modal excitation and the timbre of the drum. Since playing the drums often requires multiple strikes, it is important for timbral consistency that the drummer maintains physical control of the striking implement across a diversity of potential strike interactions.

The first point of interaction between a drummer and the drum lies with the stick contact with the surface of the drum. Earlier in this chapter the effect of stick contact location was described in relation to modal excitation and timbre, but the contact time between the striking implement and the surface was not discussed. Billon et al. found, during finger tapping exercises of accented beats, that finger contact time was greater than non-accented beats (Billon et al., 1996). An investigation into stick contact times on a tom-tom by Dahl (1997b) found contact times to decrease with strike force. These stick contact times were measured electrically by using adhesive copper foil on both the surface of the drum and the stick, for different strength strokes (*pp*, *mp*, *mf*, and *ff*), in the centre of the membrane (Dahl, 1997b, pp. 64-65). The results showed that contact time decreased in a non-linear manner with increased strike strength, ranging from 8ms to 5.5ms for the four dynamic levels of strikes in the experiment. Dahl ruled out the surface material and vibrational reflections from the rim of the drum as contributing to this behaviour by performing similar strikes on a softer surface (a carpet) and obtaining similar results, and by measuring the reflected waves of the drum head on the stick with an accelerometer. Dahl found that the reflected waves were not strong enough to influence the stick motion, but did affect the stick's bending mode at around 475Hz.

These findings present an interesting paradox in accented playing and contact times. In some kinematic analysis of percussionists, Dahl found that interleaved drumming accents were played with increased stroke height (Dahl, 2004). With a correlation between higher preparatory strike heights and striking velocity, including higher dynamic levels (Dahl, 2004, p. 768), Dahl found that accented strikes tended to have lower stick contact times. In a direct comparison between tapping with a finger and drumstick on a force transducer, Fujii and Oda (2009) found that there was little difference between tap speed and peak force variability between the finger and stick, in ten-second tapping bursts between seventeen drummers. However the authors found that tapping with a stick produced shorter contact times, with a larger peak force and greater stability in the intertap interval than finger tapping. The authors concluded that the stick "allows drummers to play drums powerfully and stably" (Fujii and Oda, 2009, p. 969). The authors also noted a difference in tap rate and stability between the left and right hand, with the left hand being generally the weaker of the two in right-handed drummers.

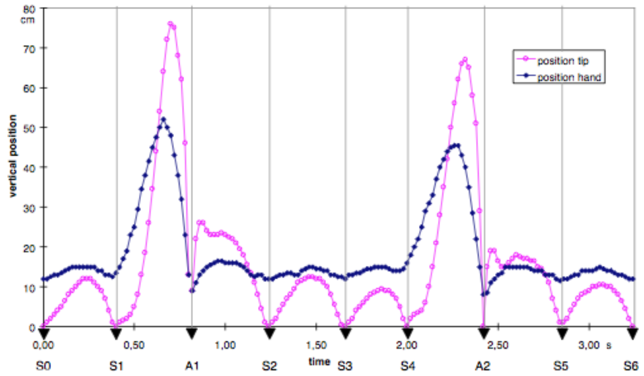
Beyond the practical aspect of force and stability, the player can dampen the vibration of the instrument by forcing extended contact with the drum head, thereby adjusting the timbre of the strike. Additionally, due to the small contact times with the drum head, such actions must be preparatory and integrated into the strike (Dahl, 2005, p. 19). The difficulty faced by the player, particularly with higher striking velocities, is the deceleration of the drumstick when it makes contact with the membrane, and the rejection of the drumstick when the membrane accelerates it in the opposite direction (Wagner, 2006, pp. 20-23), or rebound. In the case of a damping effect, the player must exert an opposing force greater than the accelerating force of the membrane, and in the case of a non-damping strike, the player must cease the downward force on the membrane to ensure no further stick contact is made once the initial opposing acceleration has subsided (e.g. a stroke that can rebound freely). In either case, the stroke is largely determined by the player's grip on the stick.

The effect of stick grip on the sound characteristics of a drum was investigated by Dahl and Altenmüller (2008, 2013) who measured contact force, duration, and pre- and post- stick velocity for two different types of grip: a normal grip where the stick was allowed to rebound freely, and a controlled grip where the player was asked to stop the stick as close to the membrane after the strike as possible. The authors measured the movement of the stick, index finger knuckle (Metacarpophalangeal, or MPC joint),[26] and wrist for both grip types and found that more energy was transmitted to the drumhead in the controlled stroke, with higher peak force and lower contact durations. In addition, the constraining actions of the wrist and MPC joint in the controlled stroke produced a lower post-strike velocity. In order to identify the effect of these grips on the sound of the drum, listening tests were carried out and the normal stroke was considered to have a more full timbre compared to the timbre of the controlled stroke. The authors note that this is due to the longer contact durations dampening some modes of the drum but, more interestingly, they "appeared to have affected both the effective mass and possibly also the stick modes" (Dahl and Altenmüller, 2008, p. 1494).

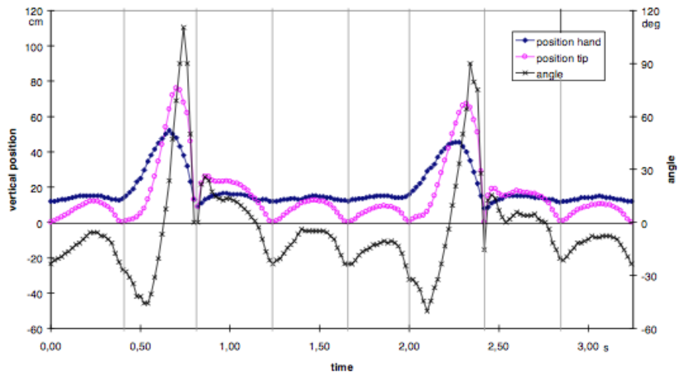
Exclusive of the timbral variations created by the instrumental mechanics, producing an accent or a desired timbre requires preparation on the part of the performer. The performer must be able to adjust (loosen or tighten) their grip or adjust the looseness of their lower arm (the wrist and MPC joint) to change the interaction of the stick and the drumhead, thus producing variations in timbre. One of the key drivers for grip modification in drumming is to control the amount of rebound. However, as Dahl and Altenmüller notes, a grip adjustment for controlling rebound should theoretically be done post-strike because the implicit equation in the preparation for physical movement affects the sound production of the stroke.

Stick rebounds can have both a positive and negative effect on drumming depending on the required loudness of the subsequent stroke (Dahl, 2003, p. 11). Furthermore, stick rebound is determined largely by the player's grip, with looser grips allowing more rebound. In a pilot study into drumming sequences with interleaved accents, Dahl (1997a) found that the stick tip position immediately following an accented strike was heavily influenced by the rebound. This was shown by the height of the tip and the upward velocity immediately after the strike (shown in Figure 3.7 (c)). In addition,

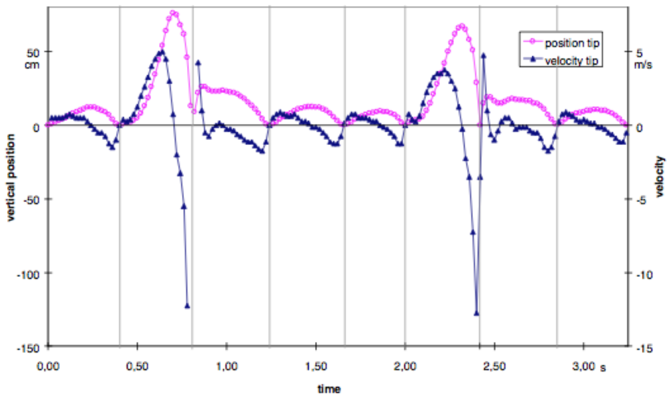
[26] An approach similar to Wing et al. (1989).



(a). Vertical position of the hand and tip of the drumstick during the drumming sequence. The hits are indicated by arrows and marked by S (soft) and A (accented).



(b). Vertical position of the hand and the tip of the drumstick, and the angle α between the drumstick and the horizontal plane.



(c). Vertical position and velocity of the tip of the drumstick.

Figure 3.7: The vertical position of a hand and tip of the drumstick of the drumming sequence, in relation to (a) the accented sequence; (b) the horizontal plane; and (c) velocity of the tip of the drumstick (Dahl, 1997a, p. 4).

the horizontal angle of the drumstick is at its second highest point immediately after the accented strikes (Figure 3.7 (b)). Dahl notes that the tip height of the rebound following an accent is “above the optimal starting position for the following soft blow” (Dahl, 1997a, p. 5), the relative starting position of the tip for the other unaccented strikes in the sequence (S1, S3, S4, and S6).

The players’ dampening of the rebound is described in Figure 3.7 (a), where the tip position fluctuates after the accented strikes. This fluctuation is a result of the player exerting force on the stick, in opposition to the upward acceleration, and then reacting to a minute overcompensation before returning to the typical stick motion (albeit with a higher starting position for an unaccented strike). Despite the increased (sub-optimal) stick height, the IOIs of the unaccented strikes (S2 and S5) appears unaffected, owing to an increased strike velocity to counteract the height, even though planning of the pre-accented strike included playing S4 early (Figure 3.7 (a)).

A later study by Dahl presented data addressing this issue, in which larger ranges of IOIs occurred in sequences played at softer dynamic levels and at slower tempi (Dahl, 2000, p. 229). In this later study, Dahl concluded that the 68% drop in IOI range from *ff* at 200BPM to *pp* at 160BPM was the result of weak rebounds from the softer strikes that, in turn, “makes the playing more difficult to control” (Dahl, 2000, p. 232). Generally, notes can be accented using either higher dynamic level or prolonged note durations. The former method for accentuation requires higher preparatory movement. Despite this, Dahl observed that movement increase did not necessarily equate to a delay in the accented note. Instead some of the unaccented strokes following an accent were delayed, although this delay was not consistent. This lack of consistency is suggestive of rebound control stemming from the greater accented stroke preparatory movement that, when combined with the difficulty in controlling weak rebounds, could account for the “short term variations between adjacent IOIs” or “flutter” (Dahl, 2000, p. 228). This flutter ranged between 2-8% of the relative tempo of the subjects and was more noticeable at slower tempos.

In contrast to weaker rebounds, stronger rebounds are more conducive to player control. Furthermore, players can exploit the upward acceleration of stronger rebounds to achieve greater preparatory height with less effort (Dahl, 2000, p. 232). Ignoring stroke height apex control, the exploitation of rebounds requires significant planning and has some far-reaching implications on a player’s drum set performance. A drum set consists of more than one drum and often features drums on different dimensional planes. For example, tom-toms (particularly those mounted on a bass drum) can be adjusted to any degree of angle from the horizontal plane depending on preference.

Additionally, cymbals tend to be angled to avoid weakening the edge of the cymbal, chewing up the drumstick, and to allow the player to strike the bell. From a practical viewpoint, these instruments can also be positioned at any angle on the vertical plane relative to another instrument, depending on personal preference. Thus, the angles of deflection of the rebound can be more or less complementary to a subsequent stroke on another drum depending on their relative positioning. Furthermore, the angle of deflection on the first drum is also dependent on the initial stroke angle. This is determined largely by the positioning and deflection angle of precedent strokes (if any), planar positioning of the drum relative to the player, and player posture. This is described in Figure 3.8. For clarity, the strike locations have been placed off centre, although the principle applies to a centrally struck drum.

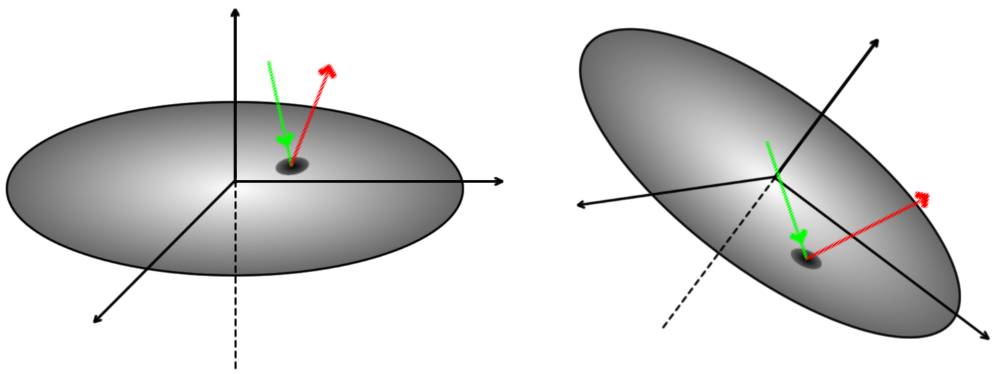


Figure 3.8: An illustration showing a potential deflection angle on a horizontal drum (left) and an angled drum (right). Arrows indicate strokes (green) and the rebounds (red).

The horizontal drum on the left is a simplified example of a typical rebound of a drum in a low horizontal plane, similar to the position of a snare drum, where the downward stroke tends to be more vertical due to the player's superior position. The stroke angle is assumed to be closer to the player and the rebound angle away from the player and follows the motion trajectory of participant S2 in Dahl (2000, p. 227). The angled drum on the right is an example of a drum closer to the vertical plane, and positioned to the right of the player. This example shows the effect of a single strike in a drum fill, where the previous instrument was positioned to the left of the player, and the rebound points towards the next drum target. In this example, the deflection angle is complementary to the following stroke.

However, a deflection is reliant on a number of variables. One of these is shown in Figure 3.9, a theoretical ideal of the angle of deflection, relative to the strike trajectory. Figure 3.9 (b) shows how, during a narrow strike angle (the angle between the strike trajectory and the membrane), coupled with momentum, contact between the drum and stick can be forcefully increased. In this example the angle of deflection becomes wider, potentially reducing the complementarity to the subsequent stroke. The prolonged interaction of the stick and drumhead can also affect the timbre of the drum.

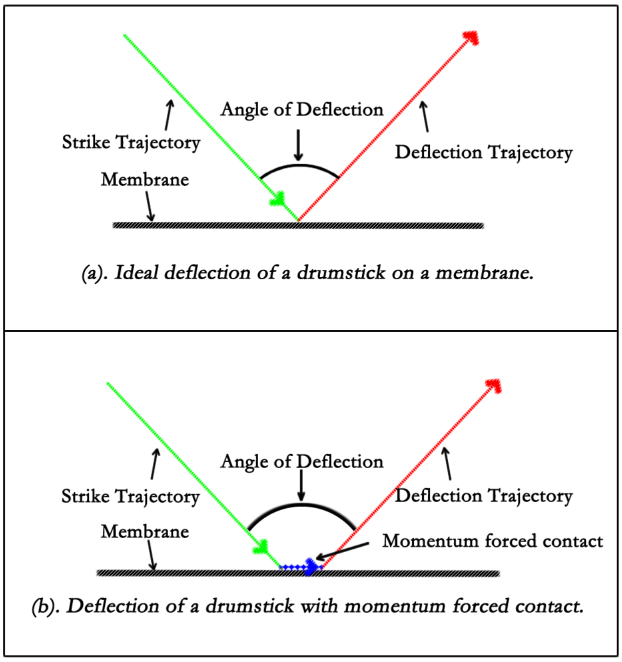


Figure 3.9: An ideal deflection of a drumstick on a membrane (a), and the deflection of a drumstick with an acute strike trajectory and momentum forced contact (b).

There are many variables that can affect the rebound angle and velocity, several of which have been discussed. However, it has also been noted by Wagner (2006) that the rebound speed is also dependent on the tension of the membrane. Wagner's experiments on force, contact time, and acceleration of a drumstick at different membrane tensions, demonstrate that the increase in stiffness affects the speed of the transversal wave propagation and internal reflection from the rim, together with a decrease in contact time as fewer vibrational modes are excited due to a reduced force pulse (Wagner, 2006). The results of Wagner's experiments are shown in Figure 3.10.

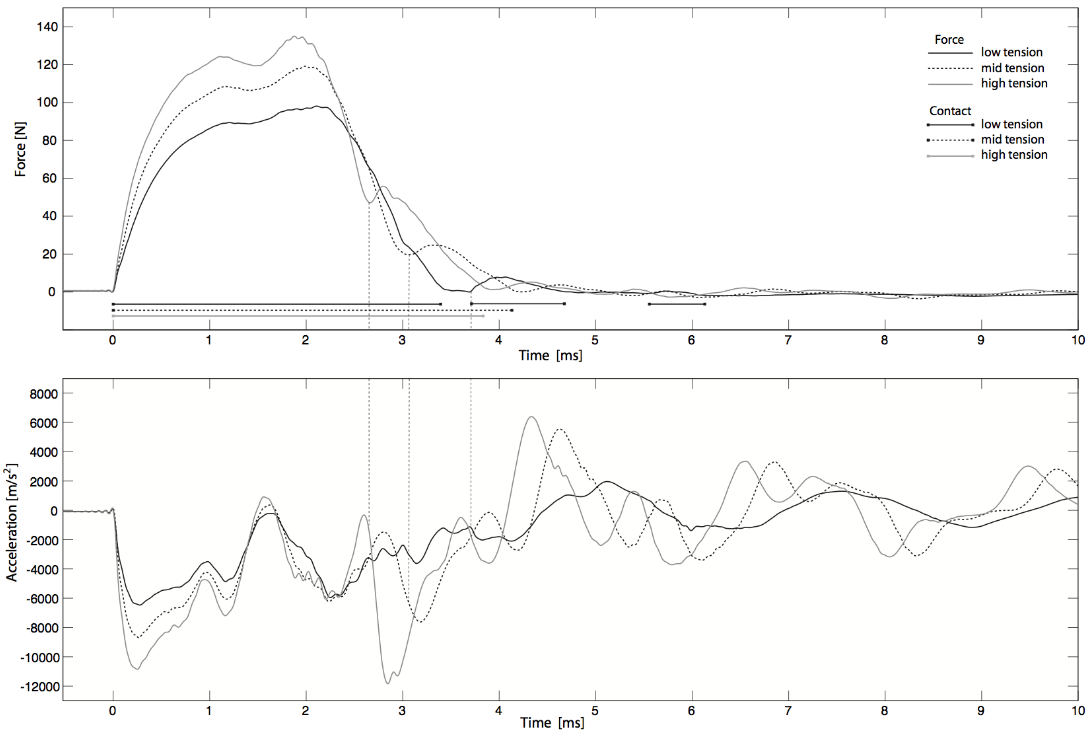


Figure 3.10: Two graphs showing force, contact time and acceleration of a f drum stroke in the centre of a single drumhead at different tensions. The vertical lines denote the first arrival of the reflected travelling wave (Wagner, 2006, p. 33).

It is worth noting that drums come in different sizes with different tensions, so a drum set contains variations in rebound behaviour. While Dahl's experiments used two-headed drums, the investigations focussed on player movement. Wagner's experiments concentrated on the interaction between the stick and drum, but used single-headed drums. There is currently no detailed literature investigating stick and cymbal interaction although, given the pivotal movement of a ride or crash cymbal on a stand, it can be assumed that there would be limited interaction with the stick as the cymbal moves away from the stick with the downward force. In the case of a rebound that is in opposing trajectory to the subsequent strike, Dahl's controlled striking experiment demonstrates that the player compensates for this prior to the stroke, using finger, wrist, and arm muscles to counteract the acceleration.

Rebounds also occur in the foot pedals associated with the bass drum and the hi-hat.

In the case of a bass drum, the static strike location of the beater on the membrane ensures that rebounds operate consistently in the same way between strikes. The bass drum pedal mechanism also amplifies the rebound, as the weight of the beater on the swing arm, and torque from the movement of the foot pedal, tends to push the beater to its default "open" position. As a result, pedal control is important and muscle use is more constant. In the case of the hi-hat, the opposing force to the foot is in the counterweight of the upper hi-hat cymbal, thus, releasing the pedal opens the hi-hat. Because opening the hi-hat generally occurs less than striking the bass drum, constant pressure is applied to the pedal with pressure release opening the hi-hat. The ratio of pedal to hi-hat movement can be adjusted but, in general, controlling this is much easier than controlling a bass drum.

Rebounds do not always have an inherently positive effect on playing. Weak rebounds are difficult to control, and opposing rebound trajectories can require either more effort to control or quick reflexes for immediate control. For immediate control, the player must anticipate the rebound and/or modify the preparatory movement of the stroke.

In Dahl (2000) the comparison between the motion of a drumstick tip and the hand, during an accented stroke, showed that the hand moved upwards before the tip of the drumstick. Moreover, this occurred while the stick was still in contact with the surface of the membrane (Dahl, 2000, p. 232). The maximum upward velocity of the hand was 2m/s with a height of 50cm above the membrane compared with 4m/s for the tip and a height of 70cm (Dahl, 2000, p. 227). Dahl describes how the differential in upward velocity and height means that it is:

> "Not until the stick has passed its upper turning point an actual force delivery may be applied by the wrist (or fingers) to increase the speed. The result is a "whiplash" of the tip of the drumstick but the motion of the hand is smooth, resembling a fishtail-gesture. This characteristic fishtail motion of the hand in the preparation and delivery of the accented blow is certainly used in other ways in drumming, like reaching a position on another drum far away in ample time. By starting the movement before the last note is finished the player gains time and thereby effort. While the hand and fingers still control the last stages of the present tone the lower and upper arm have already started to move in position for the next" (Dahl, 1997a, p. 5).

The fishtail movement described by Dahl is characteristic of several of the findings by

Kelso et al. (1991) in which the sequence of strikes resembles prone in-phase movements of the forearm (where similar muscles simultaneously contract) to a metronome. Such prone in-phase movements were found to produce a more curvilinear trajectory than anti-phase movements (Kelso et al., 1991, p. 440), and has implications in drumming performance. Firstly, Kelso et al. found that at a certain metronome frequency, around 1.5Hz, or 90BPM, a participant starting in an anti-phase manner spontaneously switched to an in-phase movement to keep in time with an increase in metronomic frequency. Kelso et al. note that, prior to the automatic switch to an in-phase movement, the velocity (which was generally more stable with in-phase movements compared to anti-phase movements) became unstable with a sharp decrease of instability observed shortly after the phase switch. They also observed that, conversely, a participant starting in an in-phase position does not switch to an anti-phase position with an increase in frequency. This one-way automatic phase transition suggests that the heterologous nature of the muscle activity in the anti-phase movement is less economical, resulting in a decrease in consistency of velocity and ultimately comfort. Additionally, the prone in-phase hand positions coupled with the velocity stability allows drummers greater control of the stick to hand interaction.

On the subject of movement analysis, two distinguishing features exist between Kelso et al. and Dahl. These relate to the lowest part of the movement (the stick and drum interaction) and highest part of the movement (the fishtail motion at the top of the stroke height). When a player starts a movement early (e.g. from the moment of impact) it causes the player to be ahead of time, therefore reducing the need to reactively make inefficient movements. It also allows the player to take advantage of the existing lower and upper arm movement, which is important in instances where the subsequent stroke requires bodily rotation to achieve optimal positioning for the next preparatory movement. For example, this is particularly relevant in sequences involving drums positioned at distance from one another. The greater the distance between the drums, the further the body must move. At higher tempos, it becomes increasingly difficult to make the movement in the required time.

The second component of the hand movement is the fishtail motion at the upper turning point of the tip. It is this upper turning point that is subject to the least amount of force, and so is easier to influence. In contrast to the bottom of the strike, there is no rebounding force, so the player employs a fishtail motion of the hand to cause the tip to change direction quickly, and to move with greater acceleration. Because playing the drums is an ongoing time-sensitive activity, both of these two components must be employed if the player wishes to increase their efficiency through improved economy of movement. The overriding goals of these components draw parallels to Shaffer's description of the motor geometry in piano performance:

> "Getting the fingers to the right locations on an instrument is important but only part of the motor task in playing. The performer can learn to shape the trajectories of movement so as to achieve timing of rhythm and variation of dynamic and tone quality with an economy of motor effort" (Shaffer, 1989, p. 383).

It is evident from both Dahl and Shaffer's description of performance that drumstick management, particularly the control of rebound, and control of the stick at the height of strike motions, are all technical elements of playing the drums that contribute towards accuracy in timbre production and timing control. Although Dahl describes variations in the overall motion between the participants (especially at different skill levels), the curvilinear trajectory follows the findings by Kelso et al. (1991).

3.2.4 Bodily Coordination

The previous section discussed aspects of physical control that can create variations in performance. Specifically, the way a performer interacts with the instrument, and how this interaction can be controlled, through implicit calculations and preparatory movements. Of particular relevance was how the interaction with the drum has much wider affects than localised variations in timbre and timing. One example was seen in the rebound of the drumstick and it's effect on the trajectory of the subsequent strike.

Playing the drums requires both bilateral movement (both pairs of limbs moving in unison) and unilateral movement (one limb moving at a time). Aruin and Latash (1995) investigated the effect of opposing bilateral fast movements in the shoulders (with and without load) of subjects standing on a force platform. They found that anticipatory postural muscle adjustments (APAs) in the trunk and leg muscles were made by the subjects, increasing to a maximum when arms were moved in a forward or backward motion, and decreasing to no APAs when moving the arm along the coronal plane. Furthermore, there were no significant differences in muscle adjustment as a result of additional load on the arms (Aruin and Latash, 1995). These APAs were evident by changes in the subjects' anterior, posterior, and vertical centres of pressure and gravity on the force plate, prior to the movement (Aruin and Latash, 1995, p. 326).

In the case of jazz drumming, it is quite common for the drummer to be in a seated position with much of the player's weight supported by the seat. This relegates the effect of the leg muscles, in the redistribution of centres of force and gravity for an APA, and is further complicated by the legs' use in applying independent pressure to the hi-hat and bass drum pedals. Consequently, upper body stabilization is carried out by the trunk, specifically the Erector spinae (ES) and Rectus abdominis (RA), irrespective of the types levels of support in the legs (Aruin and Shiratori, 2003). These findings were supported by Santos and Aruin (2008) who also found that the lateral muscles contributed to upright posture control in feedforward movements, akin to feed-forward movements in drumming, with the level of muscle activation being directionally specific. For a drummer with both legs in a fixed position for operating the hi-hat and bass drum, directional posture control is of great importance, particularly in controlling movements requiring axial rotation of the upper body.

Thus, APAs in "compound" multijoint movements, especially those involving changes in direction (Holmes, 1939, pp. 17-19), in this instance, bilateral fast movements of shoulders coupled with point-to-point axial rotation, are critical in maintaining postural stability. However, when playing the drums the player does not only play bilateral movements. In many cases, arm movements are unilateral, are not directly opposing, and are executed at different strengths and speeds depending on the relative location between subsequent drums. Furthermore, with different maximum arm heights, relative to the horizontal plane, and different maximum distances in arm reach required from the centre of the torso between strikes, postural control and stability also affects movement on the vertical (sagittal) plane, where a "hunched over" position is not conducive to playing strikes at greater heights. With this in mind, it is easy to imagine the variations in the centres of pressure and gravity of a player during the course of a percussive performance. In fact, Alén (1995) suggested similar links between movement and performance variations. In his analysis of the Cuban toque macota, Alén described how the large size of a Cuban bulá drum may have affected the performer's stabilization, requiring torso movements that could have contributed towards timing deviations.

Although there are vast differences between the jazz drum set and the bulá, it is conceivable that Alén's links also apply to playing the jazz drum set. One theoretical view is that a performer mitigates these effects by maintaining a postural equilibrium, with extreme changes in postural stability countered by APAs stemming from performance planning and musical “read-ahead”, both of which can be linked to performance skill, having repercussions on musical gesture as a learned deviation.[27]

Therefore, it can be concluded that one general rule of performance variation is that the greater the distance and angle of movement (relative to the torso) prior to the strike, the greater the inequality between the opposing reach angle and distance of the other hand, and the synchrony/asynchrony of the arm movements, the more complex the biomechanical and neurophysiological process and the increased likelihood of performance variation.

The trajectory of a drum strike is important in drumming, to such an extent that drum strike trajectory has been used as an important component in compositional specification (e.g. Stockhausen’s composition *Zyklus*). Previous discussion has described how rebound control can be used to affect the trajectory of the subsequent strike in a sequence of percussive hits. Between rebounds, the player must move the stick from one strike location to another at a speed sufficient enough to maintain correct timing. The success of which is largely dependent on trajectory, defined by Abend et al. (1982) as:

> “the path taken by the hand as it moves to a new position and the speed of the hand as it moves along the path” (Abend et al., 1982, p. 331).

In their study of hand trajectory to target, Abend et al. found the majority of subjects when asked to move their hand deliberately to a target with no instruction, opted for a straight line. With the shortest distance between two points being a straight line, one would expect movements with straight trajectories to have a shorter duration than curved trajectories to the same target. Although this was found to be true, movement duration is also dependent on speed, which Abend et al. found to be more irregular during curved trajectories. However, in cases where the average speed was low, even straight trajectories showed irregular speed patterns, suggesting greater difficulty controlling the movement. In a performance context, a lower movement speed and, therefore, a lower strike velocity, will produce weaker rebounds. Thus, the interaction with the instrument, in terms of rebound control, and the movement between the

[27] Musical gesture is closely associated with the behavioural and cognitive aspects of carrying out a physical movement (physical/procedural analytical level). Consequently, gesture is outside of the scope of this investigation. For further information on this topic, the reader is directed to Godøy and Leman (2010).

strikes is harder for the player to control.

Regarding the irregular speed profiles of the curved trajectories in Abend et al., it was noted previously that the movement of a drumstick during a strike has curvilinear resemblances due to the phasing of muscle movements (Kelso et al., 1991). However, a connection between the two cannot drawn as there were differences in planar movement. The participants in Abend et al., for example, operated on a horizontal plane, compared to sagittal movements in Kelso et al., and compared to both sagittal and horizontal movements in Dahl (2000). Despite this, there was a correlation in the increased irregularity in hand speed, relative to the anti-phase of the angular velocity of the shoulder and elbow; a joint-focussed dichotomy with parallels to Kelso et al.'s muscle synergies.

Drumming invariably uses multiple joints, each with different torques applied from the muscles which, in a multijoint movement, extends to the interaction of other joints and torques in the movement. In the case of multijoint movement, each joint will be subject to different velocity interactions at different points in the movement. Where a trajectory is changed mid-air and not using a rebound (e.g. at a higher preparatory stick height, as in Dahl et al., 2011), the joint torques will change depending on the new trajectory. Such a movement is subject to interactional forces during the planning and control of the movement, such as Coriolis, reaction, and centripetal torques (Abend et al., 1982, p. 331), although the effect of these forces change dynamically over the movement. Hollerbach and Flash (1982) observed such behaviour in relation to a curved trajectory:

> "The velocity interaction torques in fact completely dominate the dynamics at the movement midpoint because the inertial torques go through zero as the movement switches from acceleration to deceleration and the arm is moving the fastest at this point" (Hollerbach and Flash, 1982, p. 76).

In the case of a single stroke, as measured in Dahl et al. (2011), the midpoint would be the arc at the peak of the preparatory movement. In some instances, a change in trajectory at this point would have three benefits. Firstly, this enables a greater preparatory stroke height for the next strike. Secondly, the greater height enables higher maximum acceleration and downward velocity. Thirdly, as a point with the least amount of inertial torque, the player can prepare for the joint torque of the next movement. Such torque control can mitigate timing variation.

In terms of accuracy, it has been found that the trajectory of aimed movement can be learned. These learned trajectory movements were demonstrated by Georgopoulos et al. (1981) during a study of aimed movements in Rhesus monkeys. They found that practice over a period of time reduced the mean variability of the trajectory towards a target, together with improved accuracy, irrespective of target location. The implication here is that a human drummer is likely to do the same using the drums as targets. However, as previously noted, drumming requires bilateral and unilateral arm movement, and humans can be either left handed or right handed. Each of these have been demonstrated to be a contributing factor towards target accuracy (Garry and Franks, 2000), with increases in reaction time for bilateral strikes with targeting aimed by the weaker hand compared to unilaterally mirrored targeting. The effects of this can be minimised through drum set configuration, with little impact on multijoint bilateral movement.

Although there are several factors that can affect trajectory and control during percussive performance, the most significant factor occurs during multijoint movement, where joint torques impact not only the choice of trajectory, but also the control and speed of the movement. In the case of drumming, sequences involving multijoint movements can often include multiple simultaneous planes of motion and axes of rotation. Such an action is illustrated in Figure 3.11, where a drummer's movement is described between changes of strike location, from a strike on a snare drum to a strike on a crash cymbal.

In this example, a movement of the right hand from the starting position (snare drum) to the crash cymbal, there is abduction and extension of the right shoulder on the frontal plane with a posterior axis of external rotation. There is also an elbow and wrist extension on the sagittal plane with a lateral axis of rotation. Assuming no movement to the left arm, then there is also a vertical axis of rotation of the trunk on the horizontal plane in order for the drummer to position the body for reaching the new target. Kinetically, each of these axes of rotation and movement in this multijoint sequence contain torque forces that affect the movement.

If the drummer in the illustration had not included a strike at the crash cymbal, but a repeat strike to the snare drum, there would have been minimal changes to the existing patterns of joint torque and muscle activation. Additionally, another drum located at the same height as the snare drum, but closer to the crash cymbal, would cause the

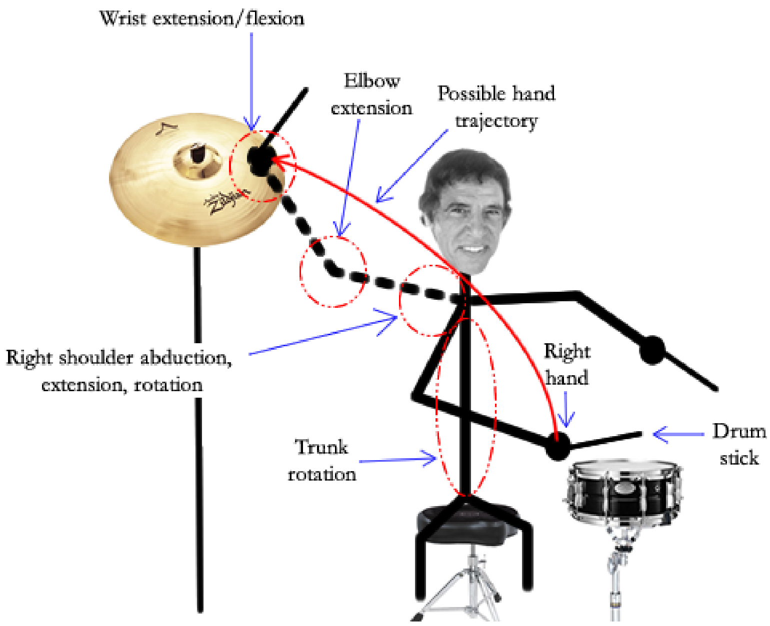


Figure 3.11: An illustration showing the typical movements associated with a change of strike location with a drummer moving from a snare drum strike to a crash cymbal strike (Sources: instrument pictures: Musician's Friend Inc., 2013; Buddy Rich photograph: JazzCorner.com, 2012).

drummer to make a trunk and shoulder rotation. However, because the drums are at a similar height, there would be less movement over the three planes. Therefore, movements spanning multiple planes of motion and axes of rotation are most likely to affect the movement of a drummer and, subsequently, the timbre and timing variations. Multijoint movements such as those in Figure 3.11 are considerably difficult to model, due to there being 17 DOF; 9 kinematic net moments and 8 dynamic with optimised muscle forces, in movements of the shoulder, elbow, and wrist (Chadwick and van der Helm, 2003, p. 15).

3.3 Compositional Application

This investigation has two objectives: to develop a percussive performance model; and to create a compositional tool that can be applied to a variety of different musical genres. So far, this chapter has presented a theoretical framework for the implementation of the performance model. The methodologies and theoretical foundations of the performance model should be considered an intrinsic part of the construct of the compositional concept. Additionally, the technological implementation of the model's processes must be considered an intrinsic part of the compositional process, in order to conjugate the theoretical and conceptual frameworks of performance and composition. The compositional application of the software tool is critical in evaluating the efficacy of the model. As Jordanous (2011) points out, it is important to identify the creative focus or "domain" (Colton, 2008; Ritchie, 2007) in which the system will be applied, and then identify the components that are relevant to that domain. The domain here is musical composition, which encompasses a vast amount of literature, approaches, and techniques and, therefore, is too broad to define the relevant components. Instead, a compositional objective must be defined. In addition, further sub-domains defined in order to identify the components of each. In the context of compositional computer systems, Pearce and Wiggins (2001) identifies two types:

> "A general distinction can be made between those systems which are designed to compose within a particular genre of music or in the style of a particular composer and those which designed to allow the generation of new styles (essentially an artistic pursuit)" (Pearce and Wiggins, 2001, p. 25).

These two types of systems are descriptive of "empirical style modelling" and "active style synthesis" (Ames, 1992, p. 55; Pearce and Wiggins, 2001, p. 3). Although the empirical data in a performance model may contain stylistic traits from the initial performance, it is not the intention here to empirically model either the genre or the compositional attributes of the music from which the performance data was captured. Existing evaluation of empirical style modelling makes use of quantitative methods based upon audience judgement ratings and controlled experiments (Eigenfeldt et al., 2012; Katayose et al., 2012). However, the generation of new musical styles is more suited to qualitative methods, that is, the analysis of the pieces.

For a percussive composer, the creation of new music is of paramount concern. To limit the compositional tool to a particular genre of music would inhibit creative possibility. Moreover, such a tool would benefit only those in a particular genre. Furthermore, the artistic imitation of a piece in the style of a particular composer would be counterproductive. Although composers may exhibit some form of compositional influence, stylistic copying to the degree implied by Pearce and Wiggins (2001) is contradictory to the point of being a composer of new music, even if the goal is objective evaluation of a system. Therefore, in preserving maximum compositional application and integrity, the purpose of the system presented in this thesis will not only allow the generation of new styles, but will be applied to existing styles in order to demonstrate the versatility of the compositional tool. Consequently, this thesis will take a combined quantitative and qualitative approach to performance modelling and the generation of new music. The successful application of this tool to different genres of music will be relevant to composers, studios, and sound designers. These styles, discussed in the next sections, consist of jazz drumming; the complex meso-periodic rhythms of Africa; live improvisation; spectral-based composition (feature-based parameters); and electroacoustic composition.

Consequently, for the purposes of this research, an evaluation of the compositional application of the software is limited to its application across different and varied musical genres. A qualitative evaluation of the composition portfolio, and the suitability of the compositional tool to those music types, will be presented in later chapters.

3.3.1 Computer-Assisted Composition

Composers and musicians have always had different views on the role of computers in music making. This includes, but is not limited to, the creation of musical instruments themselves (Mathews, 1963); tools for augmenting existing instruments (Maki-Patola et al., 2006); to one-stop compositional environments (Decker et al., 1986); computational representations of pre-existing ideas (Xenakis, 1992); and new idea formalization or representations of music afforded only by computer technology (Zavada, 2008). The latter two examples, from Xenakis and Zavada, are examples of how computers have changed the way composers think about making music.

This change in thinking is best described by Laske (1989) where composers have changed from model-based thinking (a mental representation of a subject, and the accumulation of awareness and abstraction of existing musical works to create a new composition) to rule-based thinking (the awareness, including analysis, of compositional processes). Rule-based thinking, Laske argues, involves three stages: interpretive (analysing the computergenerated musical structure); design-based (specifying the relationship between abstract rules); and improvisational (real-time exploration of the composition). Although not implicitly discussed, these three stages of rule-based thinking are evident in many computer compositions and compositional systems, and demonstrate the change in compositional thinking brought about by the computer. One common theme of many computer composition systems is the use of stochastic operations to control timbre and timing and, in some cases, structure.

One of the first major computer compositions was Hiller and Isaacson's work *Illiac Suite for String Quartet* (1956) (Hiller and Isaacson, 1957; Kirke and Miranda, 2009; Miranda, 2001). This piece was borne from experimentation to assess the suitability of different programming methods on computer composition using random selection algorithms constrained by rules and Markov chains to generate pitches and rhythms. This work resulted in a composing computer system called MUSICOMP[28] and a subsequent composition called the *Computer Cantata* (Ames, 1987; Hiller and Baker, 1964). The *Computer Cantata* used the general methods used in the *Illiac Suite*, with the addition of serial methods drawn from Pierre Boulez's *Structures* (1952) (Ames, 1987, p. 171), in order to test MUSICOMP's efficiency (Hiller and Baker, 1964). The piece employed multiple instruments, with two sections (the *Prologue* and *Epilogue*) written solely for two pitched percussion instruments (glockenspiel and xylophone), eight un-pitched percussion instruments (snare, tambourine, castanets, cymbal, tabor, maracas, bass drum and tam-tam), and noise.

Both the *Prologue* and *Epilogue* of the cantata use density of attacks and dynamic level over time, combined with probability distributions, to create the structure. The general logic for these sections is described by Hiller and Baker[29] and is summarised in a representative flow diagram in Figure 3.12. The density of attacks was derived from a weighted probability distribution of twenty-two different event durations, divided into five classes based upon their relationship to common time measures, with Class I more highly weighted than the others. As noted by Ames, "the laws of probability distributions only apply to large populations" (Ames, 1987, p. 84), thus, the distributions only become meaningful as more samples are taken. Ames describes fallacies in Hiller and Baker's original implementation, where decisions rejected by the computer could be re-evaluated in view of the random nature of the decision-making process. If a solution was not found after a number of attempts, the program would fail (Ames, 1987, pp. 93-94).

Interestingly, there are some synergies between the approach that Hiller and Baker took implementing *Computer Cantata*, and the raw/organised timbre dichotomy de-

[28] An acronym for MUsic Simulator Interpreter for COMpositional Procedures.

[29] For a more comprehensive description of the organisational principles, orchestrations, and rules, please refer to Ames (1987, p. 170) and Hiller and Baker (1964, pp. 66-67).

scribed by Boulez, where structural articulation is determined by characteristics of the music, and the timbre (raw timbre; as defined by roll, tremolo and natural decay in this context) determined by the localised structure of the density and duration of attack points (organised timbre). The incorporation of probability distributions within the algorithm as a basis for parametric selection are akin to performer freedom, although the extent of this is depends on the number of set operations to validate the complete distribution. Even though the distributions are weighted, they remove implicit instruction and introduce pseudo-random variations to the composition.

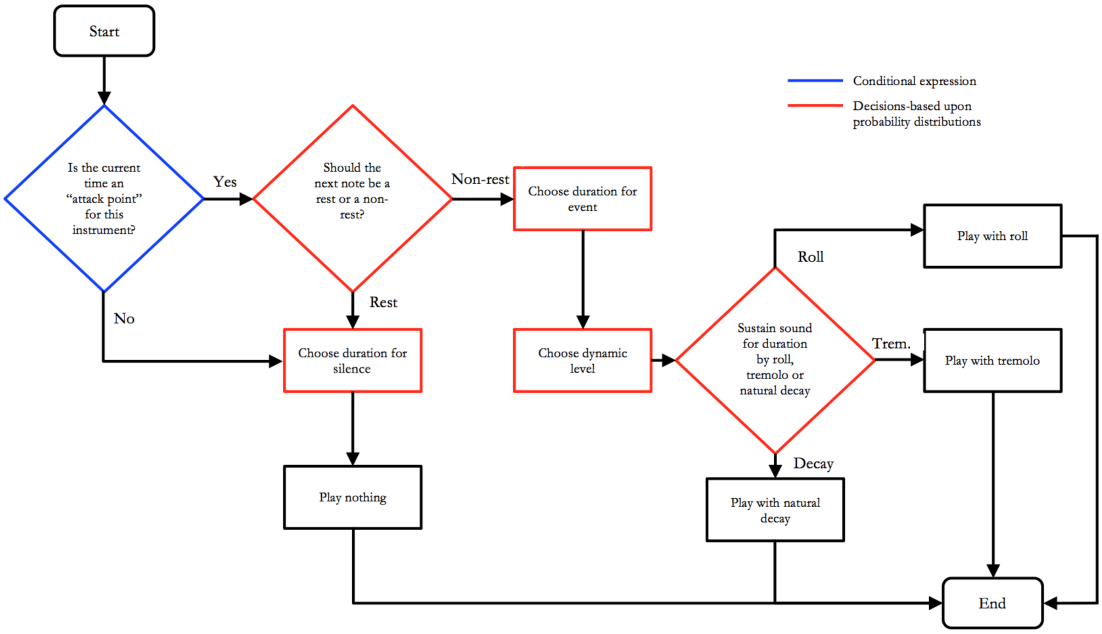


Figure 3.12: An overview of the computer programming logic to *Prologue to Strophe I* and *Epilogue to Strophe V*. Adapted from Hiller and Baker (1964).

The experiments conducted by Hiller and Baker were designed to test the suitability of a computer for musical composition, rather than represent an existing compositional approach. It is, however, interesting to note some conceptual parallels and dissimilarities between the two. Given the lack of precedent for such an approach, the creative output is unconventional, underpinning its distinctiveness. In this case, the typical human performance characteristics do not apply. By extension, the authors were constrained by technology, which limited the timbral variation of the percussion, thus reducing even further the perception of human performance. Despite these differences, one could argue that the programmed procedures are conceptually equivalent to compositional specification, and that the probability distributions are equivalent to random variables in human performance. However, given that the computer lacks prior knowledge of human performance, the probability distributions are simply to add aesthetic variation.

Another piece that was written entirely by computer was *GENDY3* in 1991, by Xenakis. This piece was based around the stochastic processes and the theory of probability that Xenakis had introduced into his compositions between 1953-1955 (Xenakis, 1992),[30] an approach that was not computationally realised until *Achorripsis* (1956-57). From the 1960s, Xenakis began using the computational power of the computer to improve the stochastic operation of his works (Serra, 1993, p. 237), where the computer was used to generate the stochastic elements, which were subsequently notated by hand.

As an aside, in 1978 British mathematician John Myhill began experimenting in computational representations of this method, and proposed an implementation designed to simplify and improve it (Myhill, 1979b). During this research, Myhill began experimenting with periodic and aperiodic rhythm, by combing a deterministic background with a stochastic foreground. Myhill notes that an example of this formalism can be seen in his percussive piece *Dialectic* (Myhill, 1979a).[31]

However, it wasn't until the 1970s that Xenakis extended this technique to stochastic synthesis program called *GENDYN*, where the stochastic operations controlled the synthesis and timbre. In 1991, he linked the stochastic elements of the computer operations with microstructure and macrostructure through the use of probability distributions, thus creating "dynamic stochastic synthesis" (Hoffmann, 2000; Serra, 1993, p. 236). For increased functionality, the user could choose a microstructural preset based upon stochastic laws (Xenakis, 1991, p. 518). Xenakis' further enhancements to the stochastic algorithm led to another work entitled *S.709* (1994) (Hoffmann, 2000). Xenakis' GENDYN program was then extended by Hoffmann to include the mathematical concept of "random walks" to create structural and timbral deviations.

Because Xenakis' random numbers are drawn from probability distributions in the same way as Hiller and Baker's, the application of stochastic operations to timbre pro-

[30] Most notably *Metastasis* and *Pithopraka*.

[31] The only published work of John Myhill is *Toy Harmonium* (1986), performed at the Japan Expo 1985, and later released as "Computer Music Retrospective" by Lejaren Hiller on WERGO (Hiller et al., 1986).

duces interesting variations, which is particularly suited to the abstract nature of the synthesis. Had the synthesis paradigm been any different, or had the resultant sound been an approximation of a real instrument, then Xenakis would have needed to increase the level of constraint of the random nature of the stochastic operations, in order to make the resultant timbral variations consistent with aesthetic expectations of the synthesised instrument.[32] One way to achieve this is by applying probability distributions to various abstracted parts of the instrument. This can be done by grouping parameters with similar characteristics and, in the case of a real instrument, grouping the sounds by virtue of the excitation location.

Using stochastic operations for manipulating timing is more complex, as the timing variations need to be constrained in order to preserve relevance to the performance context. This is where methods employing techniques such as Gaussian distributions fail to provide contextually relevant timing variations (see Hellmer, 2006). In order for stochastic timing methods to work, a large sample set must be taken into account, and minimum/maximum deviation times need to be constrained within a current metrical level.

3.3.2 Live Improvisation

Tipei (1975) developed a computer program called the MP1 for music composition. Tipei's approach embedded general musical assumptions into the MP1 to assist the composition process through "restricted tasks of local scope and consequence" (Tipei, 1987, p. 49). Like Hiller and Baker and Xenakis, Tipei also employed stochastic distribution in order to simulate random occurrences, although he describes how the MP1's randomness and probability indicates an "absence of form" (Tipei, 1989, p. 193). One major difference with the MP1 was that it was designed for both musicians and non-musicians, and had a base rule set that could be augmented by anyone wishing to expand the program to fit additional compositional needs. In addition, the musical patterns could be pre-programmed using pre-prepared data cards that were read at runtime. This allowed the composer to retain greater control over the compositional details, rather than let the computer have complete control and determine the aesthetics. Nevertheless, the MP1 still used some stochastic processes and Tipei adopted the view that it is better to combine the organisational and creative skill of humans towards musical structure, with the capabilities for randomness and transitory random-smoothness of computers (Tipei, 1989, p. 191). For this reason, the MP1 contained

[32] This assumes that the algorithmic architecture remains closely related.

presets where:

> "any continuous succession of sounds or of values for the same sound parameter can be described either as a Markov chain, as a stochastic distribution or as a random occurrence" (Tipei, 1975, p. 1).

In order to create more coherent compositional form, Tipei restricted the possibilities resulting from the stochastic level by using various techniques. The first of these techniques, "sieves" (Tipei, 1989, p. 192), was a concept first used by Xenakis (Dean, 2009, p. 121; Xenakis and Brown, 1989). Sieves are sets of restrictive selection and parametric correlation functions, consisting of logical expressions capable of probabilistic weighting. Secondly, Markov chains were used to calculate probabilities between parametric intervals over time and generate more coherent sequences of values (e.g. pitch interval distances). Finally, relying on pattern input from the user, a subroutine called "UPDATE" (Tipei, 1987, p. 54) was used in order to cause the computer to faithfully reproduce the pattern. For this to occur, the MP1 evaluated the pattern for reproducibility given the current sieve constraints and, if possible, reproduce the pattern with 100% accuracy. In the event that sieve constraints inhibited suitable playback, the subroutine reverted to the default pattern algorithm, ignoring the input user pattern.

Unfortunately, no percussive compositions were made using this program, although subsequent revisions of the MP1 led to the composition and score for *Maiden Voyages* (1987).[33] Fundamentally, Tipei took a different conceptual approach to machine composition than Xenakis (in the case of his early computer compositions that were notated by hand), insofar as the attribution of compositional responsibility and local choice is not yielded completely to the computer.[34] In each case, the creator of the program inputs operations, rules, and constraints by which the computer works but, in the case of Xenakis, the computer was ultimately responsible for the output. In Tipei's case, at runtime the computer allowed human intervention to create the desired outcome (a support system). In both cases, the computer generated notational material rather than the sound itself, leaving performance-related (timbre and timing) deviations in

[33] Tipei (1984). *Maiden Voyages* for trumpet, piano, three slide projectors and two-channel tape, on the 25th Anniversary of the Experimental Music Studios album. *School of Music, University of Illinois at Urbana-Champaign.*

[34] A sentiment echoed by Horacio Vaggione in an interview with Budón (2000, p. 12).

the domain of the performer.[35]

These examples of computer-assisted composition have dealt with multiple algorithms combined into a single computer program as either a support system or notational generator. Since one outcome of this thesis is to apply the compositional software tool to different rhythmic types and different genres of music, it is important that the composer retains control of the software, rather than deferring the compositional control to the algorithm. This involves ensuring that the computer operations are linked to specific humanization functions, in an effort to allow maximum creative application of the software tool. This has two benefits. Firstly, the composer can obtain inspiration from the interaction and musical input, which may be more difficult in a system completely controlled by computer. Secondly, the level of interaction with the composer may differ, from entering a pattern at runtime and allowing the computer to apply the humanization, to complete control of the output by the composer. In order to achieve this, the humanization algorithm can be bypassed to allow the composer total control of the live output.

3.3.3 Complex Rhythms

In order for the human mind to make sense of any musical input, it must first be organised into some form of temporal order. One of the first ways a listener will attempt to create order is to establish the pulse. A pulse is described as a series of "regularly recurring, precisely equivalent stimuli" (Cooper and Meyer, 1963, p. 3) that serves to demarcate equal points in time. In some music, such as the bass drum in electronic dance music, pulse is evident. In other music, such as Aphex Twin's *Vord Hosbn* (James, 2001), the pulse is less evident. Pulses can be objective (occurring with sounds) or subjective (implied by a mental representation of pulse), and are necessary for the perception of meter. Pulse also acts as a supporting mechanism for the perception of rhythm (Cooper and Meyer, 1963). The objective and subjective nature of pulse is characterised by three pulse listening models presented by Danielsen (2010):

1. **Metronome model**
   This model assumes that there is only one dominant or correct placement of the internal beat, and that the beats are equally spaced.

2. **Local time shift model**

[35] For a comprehensive discussion on machine musicianship, the reader is invited to read Rowe (2001).

This model takes a view that variations in interval duration of the internal beat differ from the global tempo; the listener shifts the expectation and focus of the pulse to a different temporal location. This may be temporary or permanent and most often occurs upon syncopation. In this model, pulse is a dynamic feature of rhythm (this model is adapted from Honing, 2001).

3. **Beat bin model**
In this model, each beat is described as having a rhythmic tolerance - the temporal distance from the beat, to which rhythmic details (temporal deviations from a prescribed beat) appear in time.

Danielsen presents these three models in relation to the musical example *Left and Right* (D'Angelo, 1999). While these models can be applied successfully to this example, each of them appears to suit different types of rhythm. For example, the metronome model appears the most suitable model for describing divisive rhythms, like that of electronic dance music. This is largely because divisive rhythms are metrically regular. In contrast, the local time shift model assumes internal beat differences from global tempo caused by dynamic pulses. The metrical regularity of divisive rhythms does not produce dynamic pulse locations relative to the meter, making this model less favourable to this rhythmic type. Instead, this is more suitable to additive rhythms, which feature irregular durational groups and patterns. The beat bin model can be applied equally to both additive and divisive rhythms although, in the case of divisive rhythms, the metronome model is most widely assumed. Therefore, the beat bin model captures a higher complexity of rhythm than would typically be evident in a divisive rhythm.

In his 1985 doctoral thesis concerning the automatic transcription of expressive music, Schloss (1985) proposed a global theory of rhythm. This theory is illustrated in Figure 3.13. In this system, Schloss proposes three additional types of rhythm: two hybrid rhythmic types, denoted by H_1 (Indian music) and H_2 (Latin American music); and meso-periodic rhythm. In this thesis, particular emphasis will be placed on composing with meso-periodic rhythms. Commonly found in African music, examples occur in the music of the Shona people of Zimbabwe and of the Ewe people of Ghana (Schloss, 1985, pp. 41-42). The relationship between African music and the meso-period, together with the explanation of these rhythms, is described by Schloss:

> "Music involves periodicity at many levels, from the signal itself to high-level structure. In the case of African music, there is a very fertile middle-ground

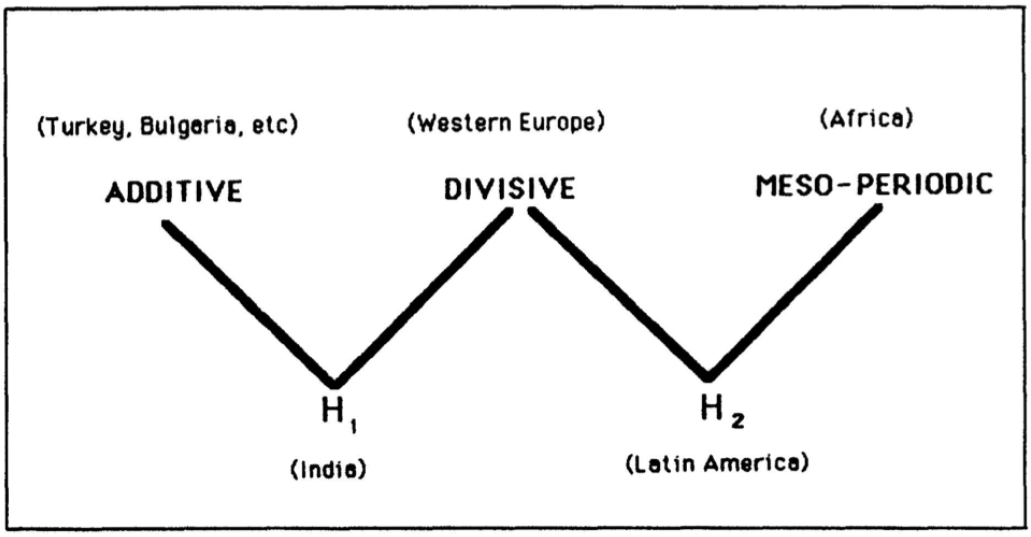


Figure 3.13: The rhythmic categories in the global theory of rhythm (Schloss, 1985, p. 38).

temporal level that is based on what I call "meso-periodicity," typically a 1-4 second long pattern. The pattern is repeated thousands of times, with very small variations in two modes: 1. rational deviations from the pattern (embellishment), and, 2. minute timing deviations from canonical pattern ("floating").

These variations can be introduced by a single player, or more typically, by several players who deviate in very small amounts from their given patterns, resulting in the bimodal deviations described above. This creates a succession (in varying temporal scope) of tension and release, which is what allows an endlessly repeated pattern to remain interesting. The repetition of this single period is fundamentally different from the other two forms, in which there is a metrical structure supporting the other aspects of the music. In African music, the meso-period is the focal point of the music.

This category is possibly the most subtle rhythmically. It is played in reference to movement, and not in reference to a pulse; that is, there is not a hierarchy of subdivided beats, but rather a parallel stream of voices (drum parts) that are "woven" together, in interlocking polyrhythm" (Schloss, 1985, pp. 40-41).

The notion that meso-periodic rhythms are streams of voices, rather than pulse-driven or bound to hierarchy of subdivided beats, suggests that the beat-bin model is the most appropriate listening model for this type of rhythm. Focus on a single "reference" stream allows listeners to identify a beat. Additionally, the concept of rhythmic tolerance in this model takes into account the inherent bimodal pattern variations, which Schloss describes as being inherent to this type of rhythm. The upper and lower temporal levels of the "middle-ground" make such variations more important in smaller duration patterns, where repeatability is higher, compared to larger duration patterns, where repeatability is lower. Therefore, such variation, combined with high repeatability of different temporal pattern lengths, will affect perceptual organisation and, subsequently, the perception of meter.

Temporal groups and patterns are the direct result of the organisation of sounds by the listener into smaller units and can be distinguished from meter, where regular patterns of different strength beats infer a reference point to which a listener contextualises the musical sound (Lerdahl and Jackendoff, 1983). As London (2012) points out, the mental organisation of sound events is based more on temporal grouping or "subjective metricization" (London, 2012, p. 15), as the listener requires a strategy to make sense of forthcoming temporal groups and patterns. Although rhythm can exist without meter, as in the case of Gregorian chant (Cooper and Meyer, 1963, p. 6), the rhythmic contextualisation of temporal groups within meter are dependant on the level of the metrical hierarchy and accented beats as focal points of the rhythm (Cooper and Meyer, 1963, p. 8). This is particularly relevant in the case of meso-periodic rhythms where Schloss argues, accents (bimodal variations) are the focal point rather than other aspects of the music that require metrical and structural cues. Essentially, meso-periodic rhythms do not use phenomenalogical accents as perceptual cues to metrical accents (Lerdahl and Jackendoff, 1983), they use them as variations for perpetuating the repeated rhythmic cycles. This demonstrates the difference in listening approach toward meso-periodic rhythms, compared to additive and divisive rhythms whose phenomenal accents are indicative of meter. From a compositional standpoint, particularly electronic composition, this raises some interesting questions.

Since the use of digital technology in creating rhythms, there have been two approaches to the manifestation of variations inherent in human musical performance in rhythm and groove. On the one hand, digital technology enables millisecond modification of sonic events to create rhythmic feel by way of adding these human performance variations. Conversely, digital technology also allows for the creation of rhythmic feel and groove that is devoid of any human variations, as in the case of electronic dance music (Danielsen, 2010). This dichotomy suggests that human variation is not necessarily required in order for a listener to perceive rhythm or groove. From an electronic music perspective, meso-periodic rhythms are the polar opposite of electronic dance music, with regards to a genre typically devoid of human variation. With human variation so ingrained into the fabric of the meso-periodic rhythms, its removal would significantly affect the perception of rhythm. Additionally, if one considers meso-periodic rhythm to be different streams of instruments, an electronic representation of human variation will require a number of timbral representations of each stream, in order to maintain rhythmic authenticity. To date, there is no electronic compositional software that represents meso-periodic rhythms from this perspective.

Schloss describes how the root of the meso-periodic pattern typically consists of a bell pattern, with two patterns being most prominent throughout much of West Africa (identified by Jones (1959), as cited in Schloss (1985, p. 44)) derived from "special subsampling of a 12-pulse meso-period" (Schloss, 1985, p. 44). Of the two common patterns identified by Jones, Schloss identifies the following pattern:

♩ ♩ ♪ ♩ ♩ ♩ ♪

as being of particular interest, as a representation of this in relative durations presents the following pattern:

2 2 1 2 2 2 1

this is equivalent to the whole (W) and half steps (H) in the diatonic scale:

W W H W W W H

Schloss also notes that if the pattern (period) is started out of phase, the resultant period represents different Greek modes. With this in mind, multiple out of phase sequences/streams produce tension in much the same way that two scales can melodically convey a sense of tension. To demonstrate this rhythmic/melodic isomorphism, Schloss presents a graphical representation of the meso-period in relation to the chromatic scale. Based on TUBS notation (Koetting, 1970), this illustration comprises of two bell patterns, including the pattern shown above (pattern 1), representing the pentatonic and diatonic scales. This is shown in Figure 3.14.

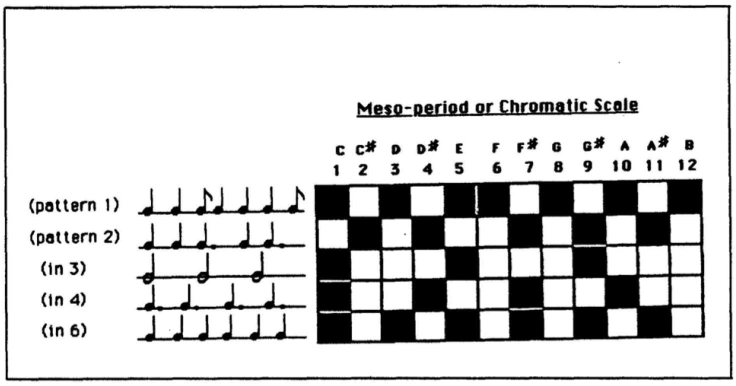


Figure 3.14: A graphical representation of the meso-period in relation to the chromatic scale (Schloss, 1985, p. 46).

Such grid representations are the basis of step sequencers, with each box representing a beat (in this case referred to as the "density referent" or the "fastest regularly occurring pulse" (Schloss, 1985, p. 46)). However, as most electronic music (and Western music) is based upon divisive rhythms, there are usually sixteen steps, rather than the twelve steps shown in this example. For this reason, step sequencers tend not to lend themselves to the composition of meso-periodic rhythms. In addition to the number of steps, the levels of potential human variation across multiple periods in traditional sequencers are not sufficient. Decomposing the structure of the meso-period in this way presents interesting isomorphic elements between the meso-period and the melodic framework of Western music. Reconstructing the meso-period electronically, and incorporating the human elements so integral to the foundation of the rhythm, is a significant challenge, hence no such system exists. The compositional software tool in this thesis fills this gap.

One of the most significant differences in aesthetic perspective between Western and African music is embedded in the social functions of the music. Arom (2004) describes this in relation to central Africa:

> "As a means of communication and an indispensable intermediary between men and the supernatural forces surrounding them, music serves to make contact the shades of the ancestors, the spirits and the djinns [demons]. That is why the Central Africans do not consider music to be an aesthetic phenomenon, even though they are quite capable of expressing their tastes

> and making very precise value judgments about both the music itself and the quality of the performance. But aesthetics remains a secondary question, and is not an end in itself. The European notion of "Art for Art's sake" has no meaning in the African tradition. Indeed, music only exists here in order to serve something other than itself, and for clearly defined purposes. That is why it is invariably a part of a more inclusive activity, a whole of which is merely a part, be it the celebration of a cult, a collective work session, a dance for pleasure on the night of a full moon or, simpler still, a mother singing to soothe her child" (Arom, 2004, pp. 7-8).

The social function of music also implies several unique aesthetics in the construction of African music. The first of these is the relationship with dance. Offering a different example of the construction of African music, Agawu (2003) describes "Rhythmic Topoi" or "Time lines" as repetitive rhythmic cycles linked with specific sections of choreographed dances in specific communities. Consequently, Agawu describes how one must understand the dance to understand the topos. In addition, the dance and music are intertwined at the same conceptual level, owing to the metrical structure of their relationship. It is this lack of understanding of dance, Agawu argues, that presents difficulties in understanding the metrical structure of African music. The music should be heard in relation to the sound of the feet during dancing. The dancing informs both the pulse selection and the pulse phasing at higher metrical levels. This is illustrated in Figure 3.15, which shows the relationship between the dance beat and the relative drumming pattern for one particular topos.

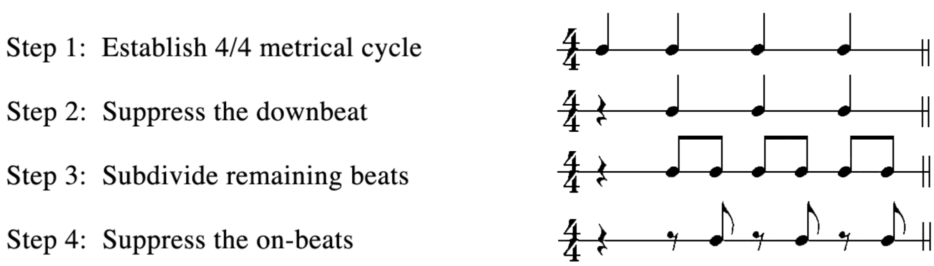


Figure 3.15: An illustration of the drum pattern construction on the topos "Highlife" (Agawu, 2003, p. 78).

Arom's account of the social function of African music, and the supporting example by Agawu, indicates that there is a contradictory aesthetic view in composition and music making between the Western and African musician.

African rhythm is distinctly different to Western music in both function and construct. In fact, difficulties faced by Western musicians and scholars in understanding African music has been ascribed to several factors including Western musical predispositions, and the lack of understanding of the complexities of the music embedded in its oral traditions (Agawu, 2003; Jones, 1959). The latter is compounded given the range of micro-cultural variation in both tradition and music, and the difficulty in assigning "stylistic traits" to localised groups (Kubik, 2010, p. 10).

Differences in African music stemming from micro-cultural variation are not easy to define, although a good explanation for these differences centres on ethnic clusters, sub-cultures, and the link to language via the oral traditional of passing on music (Arom, 2004; Kubik, 2010). With this in mind, it is difficult to compose meso-periodic African rhythms that transcend these micro-cultural differences. Therefore, it is more compositionally useful to examine the key commonalities across meso-periodic African rhythms, with a view to using these commonalities to form the key aspects of the conceptual compositional framework. It is also acknowledged that not all commonalities can be discussed in this investigation owing to the complexity of the African musical culture.

Discourse on African polyrhythm extends beyond simple aesthetic understanding and has been the source of much discussion on African music. It is not easy to notate African music, despite several attempts (Agawu, 2003; Arom, 2004; Ekwueme, 1974; Jones, 1959), consequently leading to different views of the relationship between musical devices, for example polyrhythm and polymeter (Arom, 2004).

In determining a formal structure for African musical form, Ekwueme (1974) describes four characteristics: constant (elements that occur without much change); variable (elements that appear in changed at different parts of the piece); essential (elements critical to the piece); and non-essential (ornaments). In what he calls a "hierarchy of usefulness", Ekwueme (1974, p. 47) identified four distinct groups (essential constant, essential variable, non-essential constant, and non-essential variable) and discussed their usefulness in determining form:

> "An analysis of FORM will take into account such delimiters as musical phrases and accompaniment patterns. Repetitions of melodic or rhythmic fragment should be taken into consideration in determining the structure and, when possible, the reduction should be made graphically or represented in mathematical symbols" (Ekwueme, 1974, p. 47).

With this in mind, a graphical representation is shown in Figure 3.16.

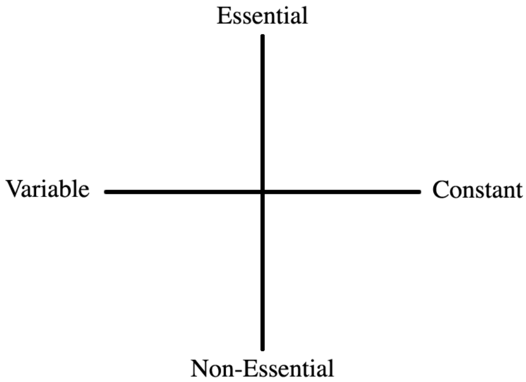


Figure 3.16: An illustration showing the hierarchy of usefulness. Adapted from Ekwueme (1974).

Interestingly, Ekwueme omits one critical point in relation to the hierarchy of usefulness. That is, the change in hierarchy of an element over time. For example, the bell sequence being critical to African music (in Afro-Cuban music, the functional equivalent is the "Clave" pattern (Washburne, 1998)), could transform over time to occupy different hierarchies of usefulness. Therefore, this graphical representation of the hierarchy of usefulness allows musical elements to be transformed over time. A representation of this is shown in Figure 3.17.

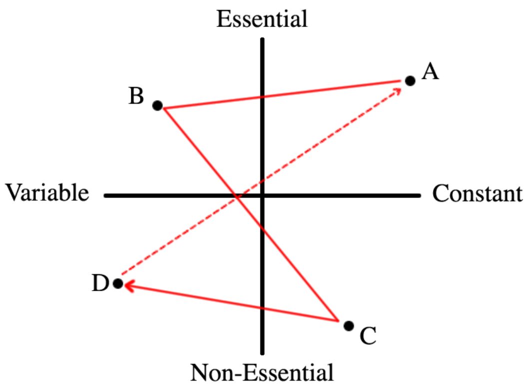


Figure 3.17: The transformational stages of a musical element's usefulness over time. Adapted from Ekwueme (1974).

To describe the hypothesis in Figure 3.17, one example could be the usefulness of the bell transformed over time. The bell pattern begins at point A by providing the temporal reference point to which all other instruments rhythmically lock. Then, to continue this hypothetical example, as the other instruments become more rhythmically and temporally independent from the bell pattern, a reduction in the bell pattern towards irregular but key components of the pattern would see the usefulness move to point B. The bell pattern could then be reduced further, to the playing of one or two regularly recurring strikes with little rhythmic significance, which would place the bell pattern at point C. Finally, the regularity of the strikes at point C is reduced to irregular strikes, with irregular accents, which has little or no musical relevance to the overall structure of the music (point D). The bell pattern could then repeat the starting pattern and return to its original usefulness at point A (represented by the dotted line). This representation is compositionally useful in order to define the structure of African music, away from the Western lens. However, it is important to define the variables that are essential and non-essential, and also to quantify the variability.

Given the social functions of African music, it is unsurprising that a performance of African music involves many performers. Herein lies some tension between the compositional framework of the meso-period and the performance model. The performance model in this investigation simulates a single human percussionist playing nine individual percussion instruments simultaneously. In African music, however, one person plays each instrument. Furthermore, as the constraints of the performance model include the physical limitations of the performer: there is a maximum of four instruments that can be played simultaneously.

As a method for extending the performance model to meso-periodic rhythms, the repeat use of the same instruments in a single representation of the performance model, does little to compositionally explore the efficacy of the performance model in different musical paradigms, particularly meso-periodic rhythms. Given that there are nine instruments in the performance model, the addition of two further representations of the performance model is sufficient to allow for each of the nine instruments to be used at any given time. Therefore, a minimum of three simultaneous representations of the model will be implemented, representing a maximum of 12 instruments. This will require the user of the model, and the listener, to consider each instrument as being independent rather than as a component of a set of instruments to be played by one person (as in the jazz drum set). One advantage to this change in instrumental perspective is that each instrument can be thought of as a stream, in which sequences of successive notes containing similar timbral properties are perceived as stemming from a single source (Parncutt, 1989).

The decoupling of instrumentation from the performance embedded in the model has deeper underlying effects on the construction of meso-period rhythms. The independence of each performer in meso-periodic music creates musical tension (Schloss, 1985, p. 51). Schloss also describes independency, where each player represents a feedback loop that creates complex independent variations in the complexity of the timing and tone, or "flux", where the independant variations cumulatively affect the resultant meso-periodic pattern (Schloss, 1985, p. 52). In the context of the performance model, the timing and timbral variations are independent for each instrument. In the case of three representations of the model, there are twelve independent algorithms for determining temporal and timbral variations, which will cumulatively affect the resultant meso-periodic patterns. Furthermore, the exploitation of independent streams allows greater flexibility regarding the structure or form of a given piece.

The notion of independency also supports the hierarchy of usefulness, with the analysis of hierarchy being relevant to different streams of instruments, as well as individual or sequences of elements. Such an approach can also be considered a methodology toward an extended technique of the jazz drum set. As each player creates independent variations in timing or timbre from a predefined pattern, changes in rhythmical tension are created in higher-order rhythmical sequences, which can be used as a structural device in a given composition.

The previous section described how the performance model is able to simulate independency by virtue of the timing fingerprints and timbral variations. However, in the case of timbre, the amount of timbral flux generated by the performance model is context-dependent. This presents an interesting compositional challenge when considering the specification of three representations of the performance model, where multiple instruments are selected across the representations. One limitation of the performance model in the independency paradigm, relates specifically to the independent levels of flux for instruments in a representation, which are intrinsically linked to the selection of another instrument in the sequence. An example of this is where, in any given representation, the local and global parameters produce lower levels of flux. Therefore, the composer must be mindful of the resultant flux when selecting independent streams of instruments across multiple representations of the performance model.

In African music, an assortment of different percussion instruments are typically used. In the performance model, nine percussion instruments are used, ranging from membranes with low frequencies, and membranes that convey a greater sense of pitch, to idiophones that display non-linear characteristics containing multiple frequencies with chaotic behaviour. The term broad bandwidth is used to describe the range of frequencies in a collection of percussive instruments and, more specifically, "the physical correlates of the source" (Schloss, 1985, p. 55). Schloss suggests two reasons why broad bandwidth is important in meso-periodic rhythms, particularly those rhythms that give rise to trance-like states in listeners:

> "The broad bandwidth is important for two reasons: physiologically, in that it may result in a wider breadth of neural excitation patterns, and cognitively, in that it obscures pitch, which results in the abstraction of melody and the resultant strengthening of the pure rhythmic impact" (Schloss, 1985, p. 55).

Based on this description, compositional uses of the concept of broad bandwidth can be defined. That is, the greater the broad bandwidth of the instruments at any given time, the greater the rhythmic impact, and the reduction in the perception of melody. The latter is particularly relevant in the case of sequences of instruments conveying different senses of pitch, which may result in an implied melodic phrase. This interrelationship is illustrated in Figures 3.18 and 3.19.

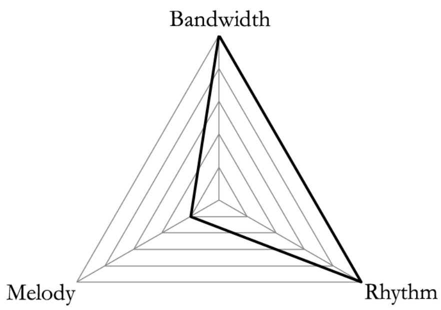


Figure 3.18: The interrelationship between broad bandwidth, rhythm and melody. Adapted from Schloss (1985, p. 55). In this example, the bandwidth and perception of rhythm is *high*, and the perception of melody is *low*.

It is worth reiterating that this interrelationship is an hypothesis presented by Schloss. Consequently, there is no strict linear relationship between the three variables and

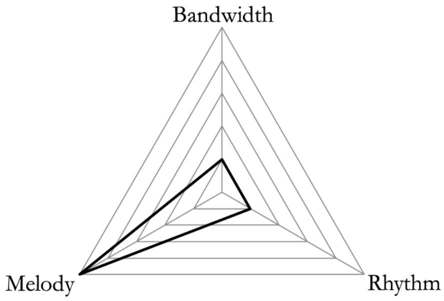


Figure 3.19: The interrelationship between broad bandwidth, rhythm and melody. Adapted from Schloss (1985, p. 55). In this example, the bandwidth and perception of rhythm is *low* and consequently, the perception of melody is *high*.

no guarantee that this model is robust enough to fit all potential percussive genres. However, from a compositional perspective this hypothesis has the potential to form the basis for some interesting structural implementations by re-conceptualising some of the important elements of the meso-periodic rhythms of Africa. This relationship will be explored in the portfolio of compositions.

3.3.4 Composing using Feature-Based Parameters

Loudness is a common measurement of dynamics in Western music. Loudness is also a useful parameter for the creation of both structural devices and local phrasing. The loudness of an individual strike within human performance aids in accentuating and conveying expression. One of the most common effects caused by loudness is that of masking. Masking is defined as the "complete or partial "drowning out" of one sound by another" (Parncutt, 1989, p. 174). This is particularly relevant in the context of independency, where part of an instrument's stream is not audible owing to another instrument. This may have some perceptual implications for the listener, where a stream at the top of the hierarchy of usefulness is masked by another less useful element in the hierarchy.

Spectral flatness is indicative of how closely a sound resembles white noise. Since white noise has a perfectly broad bandwidth, there is a positive correlation between the spectral flatness of a sample, and its broad bandwidth. Using this parameter as a basis for composition in musical form or structure will, to a certain degree, affect the perception of melody and rhythmic impact. However, a single instrument can have extreme spectral flatness, irrespective of the broad bandwidth of all of the instruments.

In such cases, the total bandwidth can be either wide or narrow, resulting in a localised effect similar to that of broad bandwidth, which will produce interesting aesthetic results. A compositional exploration of spectral flatness from both a local and global perspective will be explored in the composition portfolio.

The spectral centroid of an audio signal is the centre of gravity or weighted mean of the frequency distribution, and is often cited as an indicator of the audio signal's brightness (Eigenfeldt, 2010) where a higher centroid correlates to higher spectral frequency distribution. There are many instances where spectral centroid does not correlate to physical performance. For example, a strong strike of a large membranophone (e.g. kettle drum) will produce a lower centroid, whereas a very weak strike will produce a higher centroid, as there is less energy to excite the lower modes. In the case of idiophones the reverse is true. As the strike strength increases, more frequencies are excited, leading to an increase in spectral centroid and subsequently brightness.

Since each instrument has inherently different levels of brightness, and since the increase in brightness between each instrument is not correlated with each other, this will increase the perceived independence of the instruments. As a compositional parameter, spectral centroid offers the potential to create interesting aesthetic results, where the effect of such parametric organisation creates instability in the broad bandwidth, and subsequently instability in the perception of rhythmic impact and melody. Spectral centroid as a compositional parameter will be explored in the composition portfolio.

3.3.5 Electroacoustic Application

Before computers were readily available and applied to composition, many of the European avant-garde composers were composing electroacoustic tape music (music concrete, pioneered by Pierre Schaeffer). Notable examples of this are Stockhausen's Konkrete *Etüde* (1952) and Schaeffer's *Orfée* 53 (1953). At around that time, tape was being spliced into small pieces and concatenated to create "micromontage" works such as John Cage's *Williams Mix* (1952), Xenakis' pieces *Concret PH* (1958), and *Analogique B* (1985-1959) (Xenakis, 1992, p. 54; Roads, 2004, pp. 64-66; Sturm, 2006).

Once the micromontage technique was extended into the digital domain, the computer began to have distinct advantages over tape splicing, particularly when manipulating sound on micro levels. Two of the main proponents of microsound and, more recently, "granulation" (Roads, 1978) and "mosaicking" (Zils and Pachet, 2001) are Horacio Vaggione and Curtis Roads (2004, 2005). These new methods of computer composition led to an exploration of sound palettes, texture, and different levels of time, which utilise different computer-based effects to manipulate the textures, such as waveshaping, time stretching, and phase vocoding (Roads, 2005; Sedes, 2005). A notable piece using these techniques based on a sound palette of percussion is Vaggione's *Nodal* (1997). However in light of the transformation of many of the source sounds via digital techniques, many are difficult to identify, and the piece is very abstracted from the source palette. Other notable techniques similar to mosaicing include the CATERPILLAR system for data-driven concatenative sound synthesis developed by Schwarz (2004).

The range of complex timbres associated with a nine-piece jazz drum set provides a useful palette of sounds to digitally manipulate in order to create new textures for use in electroacoustic compositions. The application of the model for electroacoustic purposes extends the compositional relevance of the tool, and explores the timbral limits of the software and the implementation.

3.4 Pulse Code Modulation (PCM) Sampling

Sampling refers to the digital copying of sounds from one source to another (McKenna, 2000). Sampling creates a digital representation of a signal for storage and playback, and differs from wavetable synthesis by storing a large wavetable containing "thousands of cycles" (Roads, 1996, p. 117) rather than a single cycle (Kahrs and Brandenburg, 2002, p. 318). This is done by taking "snapshots" of a signal at a prescribed sample rate (Huber and Runstein, 2005), and quantizing the resulting numerical time domain representation, whose numerical length corresponds to the bit depth.[36] A high resolution sampling process allows high quality approximations of the original signal (Goldberg and Riek, 2000, p. 1; Klingbeil, 2009; Zagaykevych and Zavada, 2007) and, combined with its convenient storage, manipulation, and playback capabilities (Bongers, 2000, p. 42), has led to popular use within the music industry in both hardware and software forms.

The first commercial sampler was the Fairlight in 1979 (Roads, 1996, p. 120). By the 1990s, sampling was the basis of music genres such as hip-hop, particularly drum

[36] For further information on sampling, the reader is invited to read Puckette (2007); Roads (1996); Rocchesso (2004); Russ (2004); Smith (2004); and Zölzer (2008).

machines (Greenwald, 2002, p. 265), the first of which to use sampled drums being the LM-1 by Roger Linn (Souvignier, 2003, p. 30). Improvements in modern technology have allowed a variety of software-based sampling applications, including "FXpansion BFD, Toontrack EZdrummer, DigiDesign Strike, Reason Drum Kits, Native Instruments Battery" (Tidemann and Demiris, 2008, p. 145) amongst others. These sampling programmes often include large sample libraries (Tidemann and Demiris, 2008), emulating different performance dynamics, implemented using "complex mappings, cross faded overlays and keyswitch programs" (Klingbeil, 2009, pp. 5-6).

Another area of research that has grown with the development of sampling is the management, archiving, and exploitation of large databases of sound files, which although relating primarily to the field of Music Information Retrieval (MIR) has much wider practical applications, for example, sample banks and sampler operating systems (Herrera et al., 2002, p. 69). Feiten and Günzel (1994) describe how automatic sample selection based upon specific auditory attributes could be problematic with large databases of sound files. More specifically, problems arise where there are dissimilarities between the file descriptors of auditory attributes and the actual auditory attributes that extend to the physical sound representation, and include multiple attributes of the sound space. Feiten and Günzel pre-processed the sounds in order to reduce the descriptors of the sample feature set and applied a Kohonen Feature Map (KFM), a computational technique for dynamically arranging two-dimensional data. The pre-processing stage had the effect of improving memory space and computational overheads, although the authors noted that the computational cost was still very high. This was apparent in increased execution times during basic database functions requiring infological level reorganisation (Sockut and Goldberg, 1979, p. 375).

Automatic classification and feature selection was extended to percussion sounds by Herrera et al. (2002) who, in contrast to Feiten and Günzel, began with fifty descriptors and gradually refined the feature set to twenty. In addition, the authors devised a three-level taxonomy of the percussion instruments, including "super-category level" (classification based on physical properties), "basic level" (based on the instrument), and "sub-category level" (based upon more specific features of the instruments, e.g. high and low tom) (Herrera et al., 2002, p. 71). Both the descriptors and the taxonomy were applied to five classification techniques, with no appreciable performance differences between them under the current taxonomy. This research was then extended by Herrera et al. (2003) to include un-pitched acoustic and synthetic percussion instruments, using both temporal and spectral descriptors. However, the authors note a high error rate of between 10-30%.

Van Steelant et al. (2004) used a support vector machine (SVM) (an algorithm for analysing data, recognising patterns within the data, and subsequently classifying the data) for the classification of percussive sounds. However, this approach differs from Herrera et al. in that the sounds were taken from existing recordings. This led to different classification structures, based on the contextual occurrence of the sound rather then the physical properties of the sound, and included "Isolated drum sounds", "Overlapping drum sounds", and "Overlapping drum sounds layered with other instruments" (Van Steelant et al., 2004, p. 4). The authors compared two types of SVM implementation by using a linear SVM and a Gaussian SVM. The authors found that the linear SVM was more computationally efficient by a significant amount, compared to the Gaussian SVM.

One example of the exploitation of a large database of samples for synthesis can be seen in Schwarz (2004). Schwarz developed a system called CATERPILLAR, which unlike the automatic classification techniques described above, synthesised musical sounds from the database of samples. The concatenation of the audio samples required analysis and segmentation of the audio samples into smaller units and, using content classification data from the database of samples to drive the synthesis parameters, the author was able to preserve the spectral integrity of the synthesised tone as a single sound. Dynamic time warping and hidden Markov models (HMM) were used to maintain smooth transitions and the accuracy of the synthesis technique. The advent of sampling technology also had an impact on other synthesis methods, particularly with regards to improving the instrumental articulation (Horner, 2003). An example of this is hybrid sampling-wavetable synthesis (Yuen and Horner, 1997) where samples were used for the attack portions of a sound, and wavetables (including wavetable interpolation) used for the sustain and decay portions, with crossfading between the two applied to harmonic sounds. Other hybrid techniques using sampling include granular synthesis where the samples provide the source material for the granulation (Truax, 1987), with extensions and applications to other signal processing architectures using different digital signal processing techniques such as time-stretching (Lippe, 1994).

Sampling without modification and application with other synthesis techniques has come under significant criticism due to the inflexibility of the medium. Vaggione (1994) differentiates sampling from synthesis, describing it "difficult to consider" them as the same thing (Vaggione, 1994, p. 74) or, more specifically, sampling a type of synthesis. This perspective is taken in the context of micro-timbral structural composition, where unmodified samples inherently compositionally operate within more macro time levels. In fact, Vaggione draws upon an argument by Smith (1991) regarding the implementation of spectral modelling approaches (STFT) to understand how sound operates and transforms over time in order to allow greater micro-structural composition of samples.

Comparing sampling and (additive) synthesis, Jaffe (1995) describes sampling as having "weaker" parameters, whereby the manipulation of both musical and physical parameters of the synthesis technique are limited (Cook, 1997, p. 38). Jaffe also describes sampling as an "identity synthesis technique" (Jaffe, 1995, p. 78), where the samples themselves are parameters of the synthesis. This is supported by Kahrs and Brandenburg (2002) who described an interesting expressivity vs. accuracy dichotomy between sampling needing greater expressivity and synthesis needing greater accuracy. In addition, Kahrs and Brandenburg propose a third axis: implementation cost. Describing the third dimension, cost, Kahrs and Brandenburg states that:

> "Sampling on the other hand is extremely efficient for sound designers to produce instrument data sets for. While still requiring a fair amount of technical skill, it is very straight forward to produce a data set for a sampler once a representative set of the desired instrument sounds is recorded" (Kahrs and Brandenburg, 2002, p. 317).

This dichotomy is presented in Figure 3.20, the evolution of an ideal between the two approaches. Undoubtedly, the lack of expressivity of sampling is a significant limitation to the technique, despite improvements through the use of filters and other methods (Cook, 1997). In addition, the limitations imposed through accurate instrumental modelling using sampling synthesis are summarised by Cook:

> "infinite memory would be required to store all possible samples, the equivalent of a truly exhaustive physical model would have to be computed in order to determine which subset of the samples and parameters would be required to generate the correct output sound, and the requirement that the correct samples be loaded and available to the synthesis engine would tax any foreseeable real-time sound hardware/software" (Cook, 2002b, p. 2).

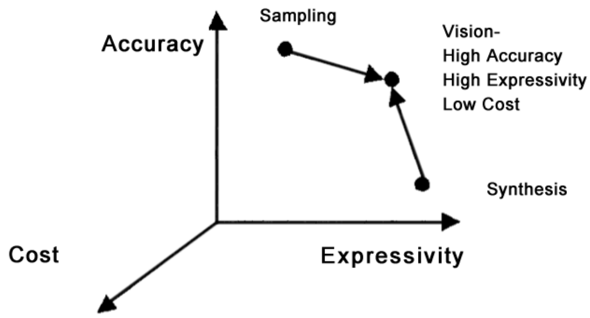


Figure 3.20: Accuracy, expressivity, and cost for synthesis and sampling approaches (Kahrs and Brandenburg, 2002).

As has been previously highlighted, the dis-uniform tuning of a membrane fundamentally changes the timbral characteristics of a drum sound, depending on the excitation location. Such a variety of timbral differences would require significantly more samples in order to represent the complete spectrum of variations, and therefore more computer memory, which supports Cook's more theoretical approach to instrumental modelling. However, the use of all possible samples for instrumental modelling may not be required, due to finite differences in timbral variation between strike locations that would mitigate the perceptual differences between the two.

Limiting the samples used for instrumental modelling could reduce the computational overhead. Such an approach was taken by Loureiro et al. (2004) who analysed the samples and, through a series of mathematical procedures, classified and clustered them using analysis and data reduction techniques, determined a more general representation of the instrument that remained robust under auditory scrutiny. The result was a data reduction rate of 64:1 (de Paula et al., 2004). Combined with increases in computational overhead since 2002, a re-evaluation of Cook's position is needed, although the challenge remains in reducing the data for the computational overhead, while retaining convincing instrumental modelling.

3.5 Summary Evaluation

The first part of this chapter initially discussed the differences in the vibrational characteristics of membranophones and idiophones closely associated with a nine-piece jazz drum set. Drawing upon previous empirical research, the discussion has shown that there are also differences in the vibrational characteristics and behaviour within each instrument classification. The intra-class differences in vibrational characteristics that lead to subsequent timbral variation are the result of many different factors, most notably the construction of the instrument, either through differences in shell size or plate, membrane configuration (single or double head), or through supporting mechanisms. Despite the timbral variation afforded by the differences in vibrational characteristics of these factors, the vibrational behaviour of equivalent drums with similar or slightly deviating properties is largely the same. This is of particular relevance in the case of a performer's ability to predict and exploit the vibrational behaviour of a drum on an unfamiliar drum set. From a theoretical standpoint, these similarities facilitate the synthesis of these instruments by implementing the similarities as generalised parameters in the synthesis process. This is particularly relevant to physical modelling synthesis that aims to simulate the physical behaviour of an instrument to produce a sound.

Conversely, and more problematically, is the generalised parameterisation of context in synthesis, where there are some physical characteristics unique to different drum sets. These are usually due to the configuration of the drum set and drums. One example relates to the choice and location of each drum's supporting mechanisms (e.g. a bass drum mounted tom-tom, or an open/closed hi-hat), however this example could also be related to preferences in individual configuration. By far the most significant factor in the production of timbral variation lies with level of uniformity of the tuning of a membrane. A dis-uniformly tuned drum can alter the vibrational behaviour of a drum set and, at a fundamental level affect the timbre production of the drum (particularly the micro-timbre) by exciting different modes than observed in theoretical ideals. Ultimately, this leads to the potential for an infinite number of micro-timbral variations across the surface of a membrane for a single instrument, and has significant implications on the repeatability of, and timbral consistency of, performance, as well as the computational methods associated with synthesis paradigms, such as some physical modelling techniques.

From a synthesis perspective this presents significant challenges. The chosen synthesis technique must not only be able to produce a sonic representation of the constituent drum types (membranophones and idiophones), but also different vibrational characteristics and timbre within the classification types, and potential micro-timbral variations across the instruments' performance space. While some synthesis techniques are more suitable to modelling membranophones, others are more suitable to idiophones. Some techniques produce more accurate representations of the sounds, while others have better control parameters for implementations. As one of the key motivations of this investigation is to develop a performance modelling compositional tool that simulates human performance variation, the development of a new technique for synthesizing percussion is outside of the scope of this thesis.

One critical factor in the ability of this model to simulate performance variation lies in the representation of micro-timbral and temporal variation. This chapter has described several determinants of timbral variation in membranophones and idiophones. The diversity of these underlying factors, combined with the differences in acoustical variation, highlights the difficulty in modelling the potential timbral variations of all of the drums simultaneously, and in an efficient manner that does not impede the real-time implementation of the model. An intermediate position to this dilemma is finding a solution that creates a successful balance between the quality of the sound reproduction and the algorithms that control the sound generation. For this reason, and despite the drawbacks to this technique, PCM sampling synthesis will be the technique adopted for this investigation.

There are several reasons for this. For the most part, the quality of the reproduction of PCM sampling in reproducing micro-timbral deviations in both membranophones and idiophones, is the most accurate synthesis technique. This is important because it has implications on the perception of micro-timbral variations. A performance model that does not accurately represent micro-timbral variations in either of these instruments can significantly affect the efficacy of the performance model. With advances in modern computing, the effect of multisampling on computer memory and hard drive space is becoming less problematic. It is intended that the performance model in this investigation will also include timing deviations, which makes real-time implementation of this model time-critical. An important requirement, therefore, is to ensure that the synthesis technique chosen has the least amount of latency possible. PCM sampling achieves this goal.

For the purposes of this investigation, there are two exclusive benefits that can be drawn from this approach. Firstly, the creation of a sample bank allows for the analysis and extraction of musical information from these samples for compositional purposes. Such feature extraction would allow for the re-parameterisation of drum timbres. Additionally, such analysis allows for the creation of a database of percussive hits for classification, and could be adapted for application to other musical synthesis systems like CATERPILLAR (Schwarz, 2004). Secondly, because there is little empirical data on the micro-timbral variability of percussion, this sample bank will provide pertinent material for researchers and scholars for analysis, perhaps with a view to further work in the areas of analysis-resynthesis of percussive timbre.

One drawback of the PCM sampling technique is that it is the least expressive of all methods discussed in this chapter. As Jaffe (1995) points out, each sample can be considered a "parameter". Therefore, an underlying computational model must be a more abstract representation, which can be guided by classification and features of the samples themselves. Another drawback of PCM sampling, particularly multi-sampling, is described by Smith (2004) as the difficulty in capturing a complete range of playing conditions of the instrument. This is particularly relevant here as, unlike Smith's description of a piano, there are multiple excitation locations across a drum. Conceptually, multisampling can be thought of as a representation of an instrument, with a larger number of samples providing a higher resolution of the representation. A balance must be found in the collection of the sample database between practicality of sample control and computational overheads, and the resolution of the sonic representation of the instrument. This is shown in Figure 3.21.

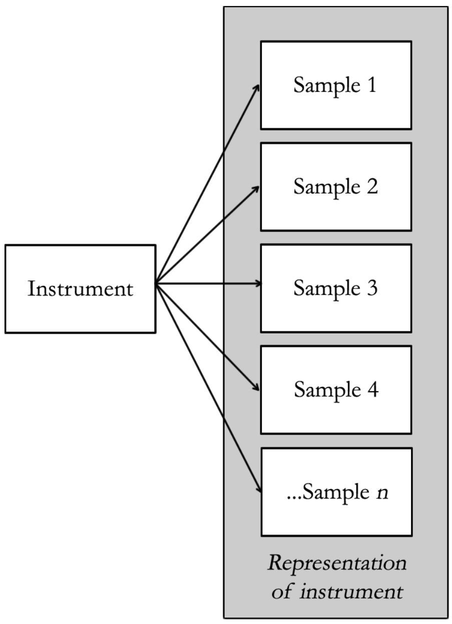


Figure 3.21: A conceptual representation of an instrument using multi-sampling.

On their own the samples do not resemble the full range of the instrument. Instead, context must be created, and the relationship between each of the samples, as parameters, must be defined in order for the representation to be accurate. Without this, the samples are an unorganised database of sounds. This has two implications. Firstly, how are differences between the sounds defined, and how are they organised? Earlier in this section it was mentioned that feature extraction is useful for re-parameterising drum timbres for composition and database classification. Extracting various features from the database, therefore, is an effective way of generating parameters and classifying the samples in the database. The problem here is defining a relationship between the feature to be extracted and the physical cause of the difference between the minimum and the maximum in the database.

The second implication lies in defining the relationship between *how* the samples operate together in the real world. That is to say, the relationship must be defined by identifying the causal connections between strikes producing sounds correlating to sample 1, compared to sample 4. Based on the discussion in the first part of this chapter, there are many variables that can affect the relationship between the samples. Arguably, excitation is the most fundamental causal factor but, in order to understand and subsequently define this computationally as rules, the underlying reasons that cause variations in excitation must be understood.

The temporal relationship between samples is also an area for investigation. This is because the states of dynamic systems often change over time, sometimes displaying vast differences between states. Music is no exception. Systems employing *nth*-order Markov chains use previous states to inform the current state, with higher numbers of orders having greater state predictability. Such systems are usually aimed towards machine learning. Conceptually, the model presented here should not be capable of learning, as ultimately such a system may theoretically learn to not produce the expressive variation this model seeks to create. This suggests a lower dimensionality of control constrained by the choice in synthesis technique and by the objective of the model.

The second part of this chapter introduced three broadly defined, interdependent development goals of drummers: physical control, coordination, and endurance. In order to narrow the analytical focus of these development goals, a tri-level analysis framework, primarily used to analyse computer simulations of human behaviour, was applied to these development goals, resulting in a focus on the computational and intrinsic analysis of physical control and coordination in human movement, specifically percussive performance.

The order of human analysis followed a bottom-up approach (local to global variables), focussing on stick management, rebound control and strike control, extending to the coordination of posture and stability control of the player, and trajectory and movement control in relation to the player's body. Stick management was found to play an important part in the interaction between the stick and the drum, by changing stick contact times that alter the vibration of the drum (and the subsequent timbre). Similarly, stick grip influences the rebound of the stick from the drum, which has two effects on drumming: force contact dampening of a drum after a strike; and positive and negative rebound use for the subsequent strike, particularly in sequences of drums operating at different angles and locations relative to the torso. Although, in most cases stick control was evident during the strike, much of the rebound and strike control was done during preparatory movements.

During the downward motion of a strike, a curvilinear trajectory was observed in Dahl (1997a). This can be accounted for by the phasing of muscle activity in the homologous muscle groups of the arm (Kelso et al., 1991). In-phase muscle activity produced greater arm stability and economy of movement, which is a contributory factor towards stick control. At the top part of the strike, a fishtail motion was described (Dahl, 1997a), which further exploits the existing synergy between muscle activities by taking advantage of the upstroke to minimise additional muscle activity in the upper arm.

In bimanual and unilateral arm movement, which are common occurrences during drumming, anticipatory postural muscle adjustments (APAs) were observed in order to maintain postural stability. These involved small muscle movements that compensate for changes in force (e.g. changes in the centre of gravity) resulting from arm extension. The effect of this, in a seated position, is that the trunk is responsible for postural stability in the upper body. With more complex arm movements in drumming sequences, compared to the simple arm movements in previous studies, the potential need for constant postural anticipation and control was highlighted, particularly in arrhythmic unilateral strikes at non-opposing angles and at different distances from the torso.

The problem identified in this section has presented is not only akin to, but includes the DOF problem. Fundamentally, the main problem in modelling a jazz drum set is that arm movements (e.g. reach distance, height, and angle) are often unequal, and the rhythmic striking of these can be irregular. The inequality of arm location, and irregularity in drumming, constantly changes the joint torques and the force interactions that affect trajectory control, movement stability, and posture and stability, which subsequently affect strike control, strike accuracy, rebound control, and stick management, ultimately causing variations in timbre and timing.

This problem is compounded by an almost infinite number of combinations of movements between cartesian strike coordinates during drumming and, if one takes into account the DOF problem, there is an extreme abundance of potential system representations. Such an abundance of potential representations would not only be hard to computationally implement, but a selection of a smaller number of representatives is difficult to theoretically justify. As Abend et al. (1982) note, there would need to be an inverse kinematic transformation of the Cartesian to joint coordinates and then, using inverse dynamics, the joint torques would need to be calculated. Therefore, modelling performance variation in drumming must take a pragmatic view. In this thesis, the modelling will be guided by the discussion in this chapter. One way this will be done is by taking a broader and more basic view of the key findings in this discussion, and creating a representative abstraction. The broader view to be adopted in this thesis, for the purposes of modelling percussive performance, is that large multijoint movements operating in multiple planes of motion are more likely to generate performance variations.

3.5.1 Compositional Summary Evaluation

The final part of this chapter described the compositional application of the software. The initial part of the discussion defined the compositional approach of this thesis by broadly summarising ways of evaluating compositional systems based upon their relevance to a given application. Then, a compositional perspective was presented in order to justify the exclusion of certain compositional approaches. The conclusion was that, because the compositional software can be applied to a variety of different genres, it should be evaluated as a tool with diverse application.

With that in mind, a variety of different genres were described in greater detail, beginning with a broad discussion on computer-assisted composition. A variety of compositional software systems, that compose music with percussion, were described. Notably, these systems differ from the models described in Chapter Two because they were not intended to humanize percussion. Their main relevance here is two-fold: firstly, in the approach to composition; and secondly, in the relationship between the composer and the computer in the generation of musical material. One of the main conclusions drawn from the discussion in this section is that, for the purposes of creating a software tool with a diversity of application, complete compositional control must not be yielded to the computer. Such a system should *support* the compositional process. Following on from this, the discussion then focussed on the application of systems in live improvisation, thus continuing the theme of compositional control. Specific examples of live improvisational systems were described, including the conceptual reasons for the human-computer relationship regarding compositional control. The discussion in this section reinforced the notion that the computer should control certain functions and, in extending this to live improvisation, provision should be made to enable the composer to bypass certain functions and retain complete compositional control of the software.

From a musical perspective, this part of the chapter also described the musical context in which the software will be applied. Different types of rhythm were described, with a particular emphasis on the meso-periodic rhythms of Africa. Parallels were drawn between the construction of rhythmic sequences and the chromatic scale, which also presented a method for constructing meso-periodic rhythms. The difficulties in defining an aesthetic perspective of African music were also described owing to the subcultural variations in Africa, resulting in a description of analytical methods common to most West African music. These included the hierarchy of usefulness, independency, broad bandwidth, and flux. Each of these analytical methods is useful in not only composing meso-periodic rhythms, but in qualitatively evaluating the resulting compositions in that genre using the software tool.

The final part of the discussion focussed on using the spectral analysis of samples in the database as compositional parameters. It was suggested that re-ordering the samples by spectral content will produce interesting compositional results. The selected spectral features were loudness, spectral centroid, and spectral flatness. The final compositional application was described in relation to electroacoustic music. Notably. the creation of electroacoustic music with this model tests the limits of application, and situates the compositional tool presented in this thesis as one component within a larger compositional process.

The final part of the discussion also focussed on describing PCM sampling, with emphasis on the considerations of using this synthesis technique in the context of capturing a timbral representation of a jazz drum set, and in creating the humanizing algorithm. Particular attention was paid to the expression vs. accuracy dichotomy of sampling. Notably, this section presented overall considerations to the methodology presented in the next chapter regarding the chosen synthesis approach.

3.5.2 Implications of the Methodology

Adopting a physical-based approach presents us with two levels of conceptual representation of the system. The first relates to David Marr's representational level, whereby the relationship between the samples in a database conceptually represent an instrument. The second level further abstracts performance, and presents a more contextual understanding of the variables that can affect relationship between the samples by inferring a relationship between the instrumental representations themselves. This is described in Figure 3.22, a simplified diagram showing the context of two instrumental representations. This illustration extends Figure 3.21.

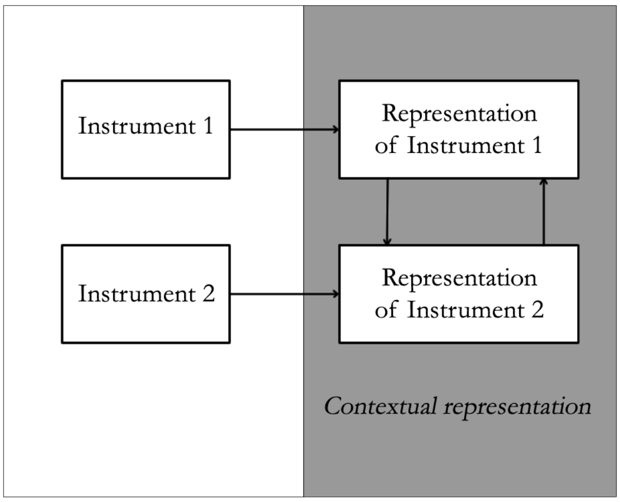


Figure 3.22: The interaction of instrumental representations in context.

Holistically, the instrumental representations are part of a larger conceptual construct related to performance context. With this in mind, a theoretical model is presented in Figure 3.23, which shows a compositional framework derived from information in the sample data, augmented by the performance model. This figure also shows the conceptual flow and linkage between components in the theoretical framework.

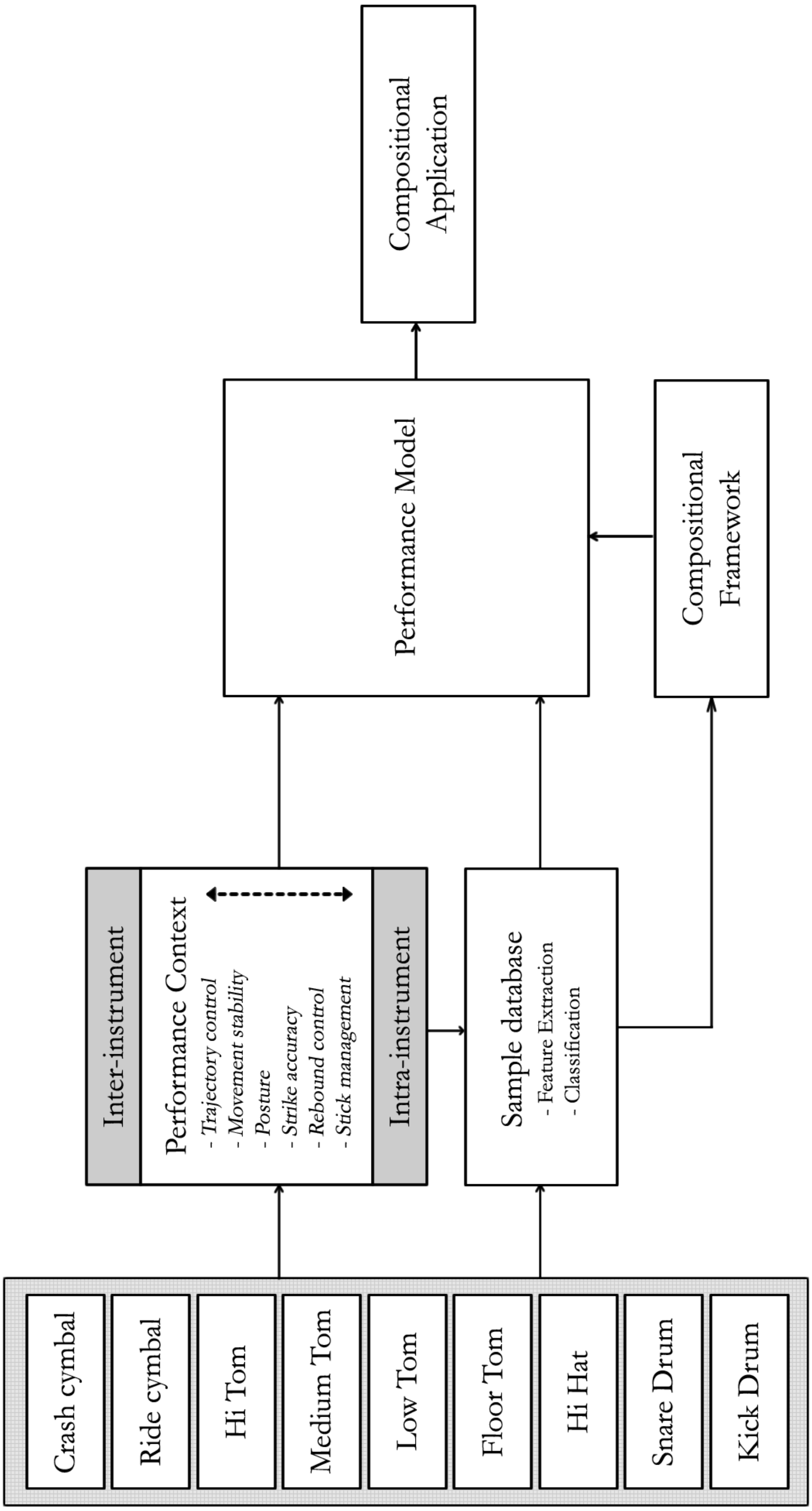


Figure 3.23: A theoretical model showing the relationship between the implementation of the performance model and the compositional application.

The link between the instruments and the performance context is characterised by the interaction of the performer and the instrument, and comprises both intra- and inter-instrumental interaction. This draws upon the discussions in parts one and two of this chapter.

The link between instruments and the sample database is an inherent feature of the chosen PCM sampling technique, and was driven by the diversity of acoustical behaviours of the instruments as discussed in the first section of this chapter. However, it is important that the feature extraction and classification methodologies in the sample database are informed by the performance context, to ensure that the parameterisation and control of the sample database is consistent with human performance, particularly as the sample database is critical in ensuring an accurate representation of instrumental performance in the model. In addition, performance context is not implied by the presence of a sample database alone. Therefore, the control of the sample database in the performance model must be informed by the performance context, at both intra- and inter- instrument levels, in order to imply a human performance on the control paradigm.[37] The methodology for the implementation of the performance model will be dealt with in the following chapter. The compositional framework has two links: the sample database as being the source of a compositional framework; and the influence of the performance model on the compositional framework. The feature extraction of the sample database for use in the performance model can be exploited to create new compositional parameters that extend the performance model into new creative realms. This will be discussed in the following chapter.

It was noted earlier that performance characteristics change across different genres of music. The implications of the performance model, as a representation of a performance in a particular genre, on composition is significant. The re-contextualisation of instruments and the representation of any musical structure implied in the model, coupled with the production of expressive aspects of performance in the model (as implied by the timbral and timing variations of these instruments) present enormous opportunities for developing new aesthetical paradigms of rhythm. The discussions in this chapter have identified the fundamental aspects of performance context relating to the production of micro-timbre in an instrument, and the potential causes of micro-temporal strike variations. In order to represent these, the performance model must make use of relevant approaches. These will be discussed in the following chapter.

[37] Similar to note, intra-note and note-to-note transitions (Maestre et al., 2009).

Chapter 4

Research Design and Methodological Approach

"Personally, for my conceptions, I need an entirely new medium of expression: a sound-producing machine"
– Edgard Varèse, 1966

This research took a two-part experimental approach. The first part of the methodological approach, which is discussed in the first half of this chapter, concerned the collection of timbral data from the instruments of a nine-piece jazz drum set. The main aim was to assess the effects of strike strength and location on drum timbre in order to develop a methodology with which to create an effective micro-timbral model. Existing research designs were evaluated in order to formulate a research design method for the collection and extraction of percussive timbral information. This comprised a brief overview of existing approaches, with specific emphasis on limitations of feature extraction methods with un-pitched percussion and subsequent difficulties in applying this for compositional purposes. The second part of the methodological approach, which is also discussed in the first half of this chapter, concerned the collection of human percussive performance information captured from three differently skilled drummers. The main aim was to assess the relationship between performer skill level, movement, performance variation, and timing variation during a performance, in order to develop a methodology with which to create an effective micro-timing model. Timing fingerprints are presented, and the methods for creating these timing fingerprints using analytical software are evaluated, in this section.

The latter half of this chapter begins with a discussion of the preliminary results of the data, including instrumental micro-timbre, data reduction and classification by strike location, timing fingerprints, and participant video and accelerometer data. This is followed by a discussion of the creation of the performance model, PD-103, with particular emphasis on computational design. The chapter concludes with a discussion on the compositional design aspects for the PD-103.

4.1 The Performance Model Defined

This section will detail the methodological approach of the performance model. It will also further define this methodological approach in relation to the theoretical model presented in the previous chapter (see Figure 3.23). Having presented a conceptual overview of the model and the role of the instruments, performance context, sample database, and compositional framework, attention will now turn to defining the flow of control data in the performance model, and describing the various operations and functions. The model presented here is iterative, and dynamically changes the timbral output depending on the user input. This methodological approach is shown in Figure 4.1.

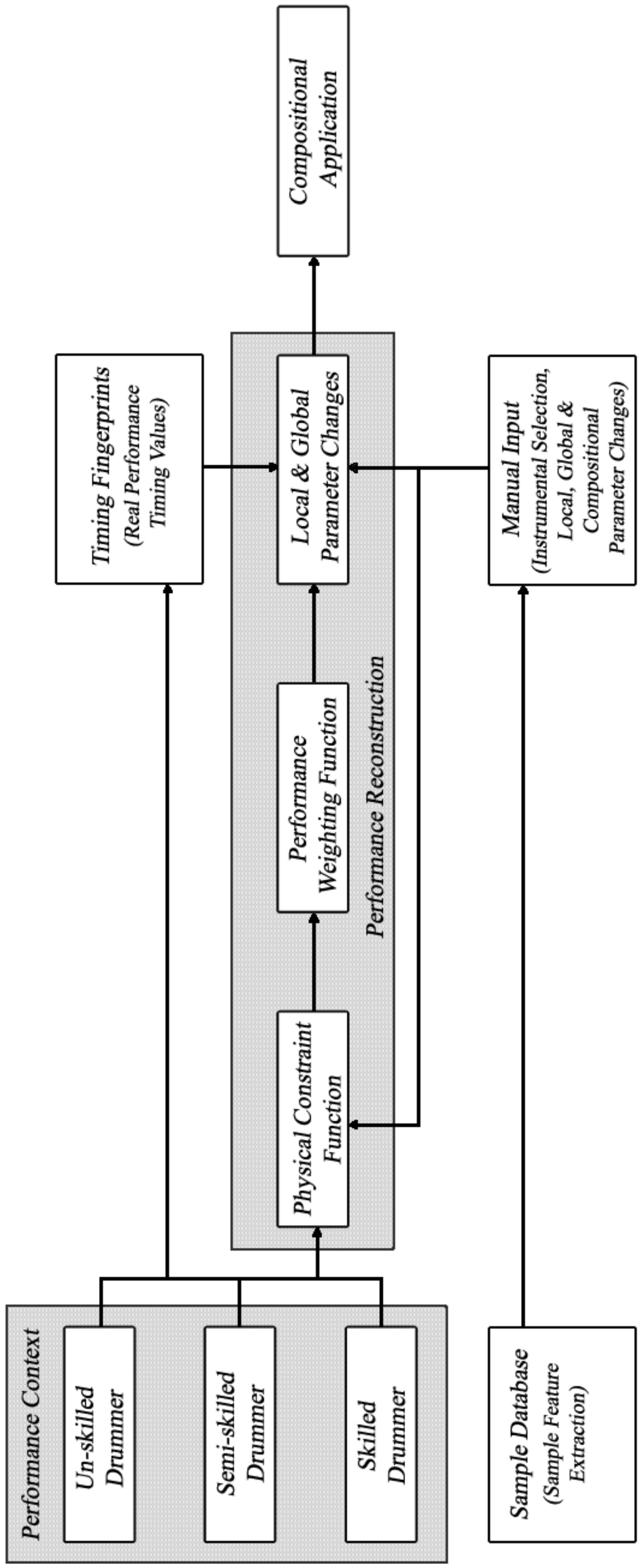


Figure 4.1: An illustration showing the methodological approach to the performance model.

This methodology contains four significant elements that distinguish it from other methodologies. The first signifant characteristic of this method relates to the performance context. Represented here are three different skill levels of drummer, the function of which was to provide three performance sub-environments. The manual selection of three drummers of different skill levels comprised the historiographical context of the drummers, with differing proficiency towards development goals. More specifically, the effect of rhythmic complexity and compound movement on performance variation. The creation of three sub-environments from real human performance had four key advantages. The methodology:

1. allowed for the capturing of various percussive performance parameters;
2. facilitated the comparative analysis of performance across skill levels;
3. enabled the definition of rules from the comparative analysis of several different parameters; and
4. yielded real performance data for use in the implementation of the performance model.

It was important, therefore, to ensure that capturing the percussive performances was approached experimentally, using appropriate methods. These methods included video, audio, and accelerometer data. Current research shows that there are no methodological approaches to simulating percussive performance that use this combined methodological approach, particularly in the use of extracted data from real human percussive performances.

The second characteristic of this model relates to the performance reconstruction. This was a three-stage process, consisting of a physical constraint function (PC), a performance weighting function (PW), and local and global parameter change functions (LGP). Once the three performance contexts had been defined, and appropriate rules had been identified from the analysis of the experimental performances, high level physical constraints were applied to the percussive performance model based upon relevant rules in the each of the three performance contexts. The PC function was the first step towards realising the physical context of human performance. Since this model is concerned with creating compositions and live performances, the PC function was designed to accept instrumental selection via manual input, and either accept or reject the input according to the predefined rules in the function. The PC function was also designed to apply these rules discretely to the manual input.

Notably, the PW function is a major component in defining the performance reconstruction. Firstly, the PW function was designed to accept the constrained manual input from the PC function. Then, it was designed to analyse the current and previous output of the PC function, and hierarchically smooth between the discrete points (first-order smoothing). Where a previous point does not contain a physical output, a zero-order smoothing method was applied to the current output. This hierarchical smoothing then applied weighting to the sample selection criteria. Relevant samples were then selected according to the weights which are governed by rules generated from the performance analysis.

The third significant element of this model manifests itself in the form of timing variations using real performances of differently skilled drummers. The LGP function was designed to select a temporal deviation from the timing fingerprint of the currently selected performance context, and apply that deviation to the weighted output of the PW function for the current operation. In addition, a LGP function then adjusted the duration of the temporal variations in the fingerprint according to the manually selected beat level. This method of creating temporal variations in percussive performance modelling differs from existing systems, where the timbral and timing variations here are dynamically selected at runtime.

The resultant performance reconstruction, from the collection of functions and parameter changes, then produced the output to be applied compositionally. Consequently, the manual input and the resulting performance reconstruction directs the compositional application, which is the fourth element that distinguishes this model from other systems. Compositional parameter changes, occurring from manual input, consist of two types: the re-classification of samples based upon spectral content; and the structural organisation of the content in performance reconstruction, including the repetition of percussive sequences.

4.2 A Method for Representing Micro-Timbre

This section will describe the required properties of the sample database, including the database size and feature extractions that will serve as compositional elements. It is not intended to be an exhaustive discussion on analytical MIR methods, as this is not the focus of this investigation. Upon conclusion of this section, the key properties of the sample database will be defined, together with a description of the features that will be extracted from the database, as well as the computational methods using MIRToolbox (Lartillot and Toiviainen, 2007) that will be used to extract these features. A brief description of the compositional application of these features in relation to the instruments of the jazz drum set is also described.

4.2.1 The Sample Database

In order to fully represent the number of possible micro-variations of a drum, an infinite number of samples must be taken. However, this is neither practical owing to computational constraints, nor are such levels of micro-timbral detail necessary from a perceptual perspective. However, it is important to address the limitations of existing sample databases when considering the sonic representation of an instrument. Such limitations include the number of samples taken per instrument. Hellmer (2006) used a database totalling 98 samples across nine instruments, of which a maximum of 28 samples represented the hi-hat, a minimum of 4 samples represented the bass drum, and 4 samples represented a crash cymbal. This falls short of the multisampling approach of many commercial software drum programs, whose databases use on average 127 drum sounds per instrument - the maximum number of samples assignable to a MIDI note. However, the samples in these databases are intended to produce smooth timbral changes across the MIDI note velocity, with minimal micro-timbral variation.

In capturing percussive gestures, Tindale et al. (2004) recorded 1,260 samples of a snare drum using a brush and a stick at different locations on the membrane, with a view to analysing the spectral features of the samples for timbral classification. Tindale acknowledges the timbral variation caused by excitation location, by specifying five different strike locations, each location struck 20 times by three different subjects. These strike positions are shown in Figure 4.2. Although the specification of these strike locations provides a more detailed timbral representation of the snare drum, the locations are only representative of part of the membrane. As noted in chapter two, dis-uniform tuning can create areas of higher tension in different parts of the membrane, and cause changes to the speed of the vibrations and boundary reflection times. It is important, therefore, to take samples from a range of locations across the drumhead, and not just specific locations across one line of a single polar axis.

Tindale's results demonstrate the successful automatic recognition of samples with

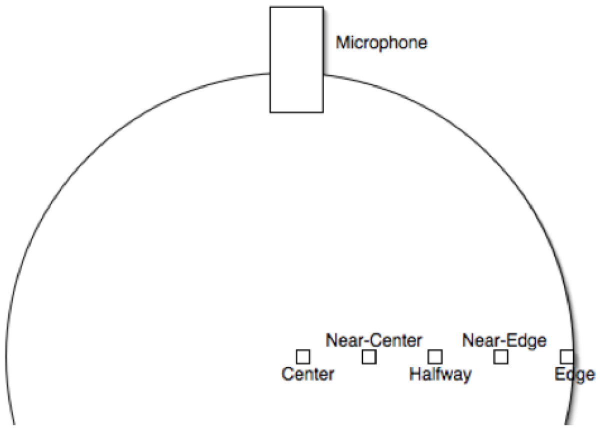


Figure 4.2: The strike locations specified by Tindale during data collection of snare drum timbres (Tindale et al., 2004, p. 543; Tindale, 2004, p. 23; Tindale et al., 2005a, p. 5).

differences in timbre resulting from different strike locations (Tindale, 2004). These results suggest that recording of samples across the entire surface of the drum, together with an increased sample size of at least 1,000 samples per instrument, is the next logical step in extending Tindales methodology. With a total of nine instruments, the sample database in this model contained a minimum of 9,000 samples. In order to ensure that micro-timbral variations are present in the membranophones, the drums were not tuned uniformly.

4.2.2 Timbral Feature Extraction

Feature extraction is an area of inquiry within MIR that has seen a number of systems and tools developed for extracting information from sound, for a variety of audio and music classification purposes. As a result, some systems are more general with broader high-level features, while others are designed with low-level features for more specific use. Since this thesis is concerned primarily with the performance of un-pitched percussion (including those instruments that evoke a greater sense of pitch, such as the tom-tom), there are limitations in the types of features that can be extracted, and the suitability of these tools for extracting relevant features. Finding a set of features that are robust across all of these instruments is a challenge, particularly with features intended for compositional application. This section will define the features that will be extracted from the audio, and describe the tools to be used to undertake the feature extraction.

Extracting pitch values from un-pitched samples is extremely difficult owing to the autocorrelation algorithms that attempt to find stable pitches where none exist. This results in artefacts and misleading pitch values for these sample types. Feature extraction programs such as *Praat* (Boersma and Weenink, 2013), a specialised speech analysis and synthesis tool,[38] and MIRToolbox (Lartillot and Toiviainen, 2007), a suite of modular music analysis functions available for use with the commercial software Matlab,[39] both use autocorrelation in their pitch algorithms. Furthermore, pitch analysis in Praat is difficult with sounds containing higher noise floors (Boersma and Weenink, 2013), of which cymbals can be considered "noisy" due to their nonlinear and chaotic behaviour.

In MIRToolbox, however, there are a number of other algorithms for extracting other features that do not use autocorrelation methods,[40] and are therefore suitable for use on both pitched and un-pitched percussion. The extractable musical features in the original implementation of MIRToolbox (2007) are shown in Figure 4.3, with a complete list in the latest version shown in Figure 4.4.

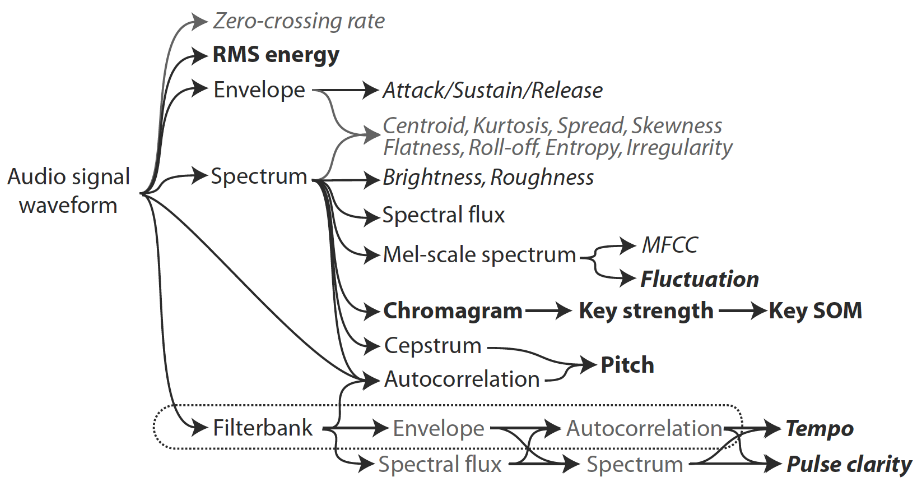


Figure 4.3: Extractable musical features in MIRToolbox (Lartillot and Toiviainen, 2007).

[38] Although Praat is mainly concerned with speech analysis, it has been successfully applied to a flute for pitch estimation in Hindustani classical music (Ramesh and Sahasrabuddhe, 2008).

[39] For further reading concerning Matlab, the reader is referred to Hunt et al. (2001); Kiusalaas (2005); Mirza (2010); Quarteroni et al. (2010); Register (2007) and Trefethen (2000).

[40] Later versions of MIRToolbox contain additional features, such as inharmonicity, that use autocorrelation methods (Lartillot, 2010).

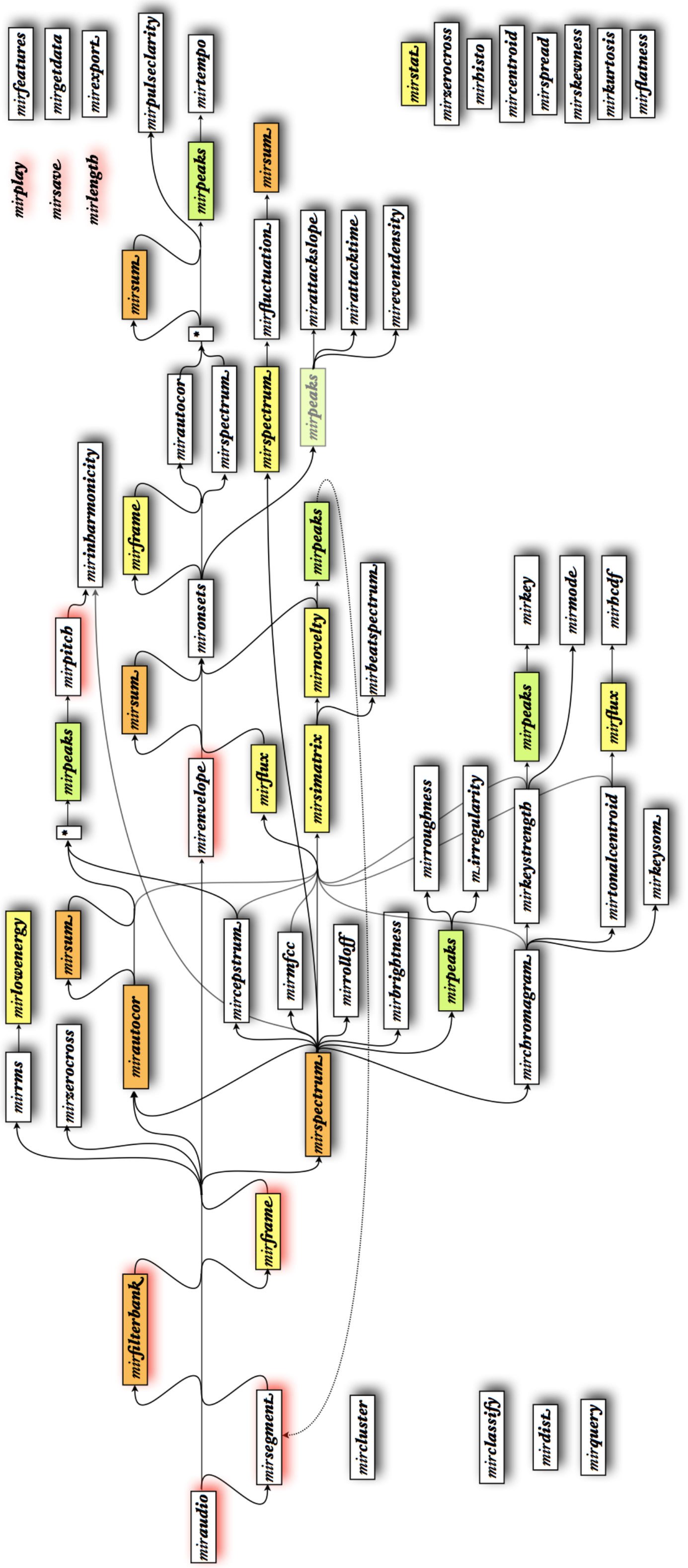


Figure 4.4: An overview of the MIRToolbox functions (Lartillot, 2010, p. 9).

Strike strength is a common parameter used to map samples to MIDI keyzones in commercial software, particularly for simulating dynamics where spectral changes occur (Dahl, 1997b; fxpansion, 2005, p. 22). One function that computes the global (RMS) energy of the sample in MIRToolbox is *mirrms*, which uses the following calculation:

$$x_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} = \sqrt{\frac{x_1^2+x_2^2+...+x_n^2}{n}}$$

Although using loudness as a parameter for composition is not new, it was useful to include this parameter for two reasons. Firstly, this parameter allowed for the recreation of performance for evaluation of the model. Secondly, and more importantly, from a compositional perspective this parameter is important for critical compositional devices, for example, crescendos, decrescendos, dynamics, sound duration, and timbre. Since average values per sample are required, segmentation and framing were not necessary.

4.2.3 Compositional Feature Extraction

When looking at timbral parameters for musical form, it is useful to consider the most perceptually important timbral dimensions. One such dimension is brightness (Aramaki et al., 2006; Barthet et al., 2008; Darke, 2005; Donnadieu, 2010; Giordano and McAdams, 2006; Marozeau and de Cheveigne, 2007; Pressnitzer and McAdams, 2000; Schubert and Wolfe, 2006; Turcan and Wasson, 2003 and Risset and Wessel, 1999, pp. 147-148). The *mirbrightness* function in the MIRToolbox measures the proportion of spectral energy over a designated cut-off frequency (Juslin, 2000, as cited in Lartillot, 2010).[41] The nine instruments in the jazz drum kit provide a rich palette of sound for compositional exploration, ranging from the low frequencies of the bass drum, to the higher frequencies of the cymbals.

Another useful timbral parameter is spectral flatness. In MIRToolbox, *mirflatness* determines the flatness of the frequency distribution from a ratio between the geometric and arithmetic mean (Lartillot, 2010, p. 148) using the following equation:

$$\frac{\sqrt[N]{\Pi_{n=0}^{N-1} x(n)}}{\left(\frac{\Sigma_{n=0}^{N-1} x(n)}{N}\right)}$$

[41] The default cut-off frequency in *mirbrightness* is 1500Hz.

Spectral flatness is indicative of how closely the sound resembles white noise (Dubnov, 2006). This particular parameter produced interesting compositional results given the propensity for cymbals to have very noisy spectra, for example, a high flatness value (Brent, 2009).

4.3 Data Collection: Micro-Timbre

The overall specification of the sample database was described in section 4.2.1, and was derived from a review of limitations of previous methodologies in percussive research. A sample database consisting of at least 1,000 strikes per drum was suggested, leading to a sample database of 9,000 samples. The number of strikes per instrument was selected as an optimum amount to balance constraints in computational overhead, with the need for higher resolution sonic representation of the instrument. It was also noted that the samples must represent different strike strengths and surface locations and, in order to produce micro-timbral variety, the drum should not be uniformly tuned.

4.3.1 Protocol and Procedure

Recordings were taken of the isolated drums inside a studio reinforced with sound absorbing curtains in order to mitigate early reflections and outside noise.[42] The drums were arbitrarily tuned by the author with the intention of ensuring that the membrane tuning was not uniform, so as to generate as much micro-timbral variation as possible. A set of "capture" rules were devised to ensure maximum timbral variation, and to ensure a random sample of strikes from the widest possible surface area of the drum. Firstly, each drum should be struck at least 1,000 times (1,000 +5%). Secondly, each strike must be: a) executed with a stick; b) devoid of expression; c) a single stroke; d) in a different location to the previous stroke; e) at a different strike strength to the previous strike; and f) commence only when the sound from the previous strike is perceived to have ended. Any mis-hits (for example, strikes that include contact with the rim) will be included in the database. Owing to the excitation method of the bass drum using a bass drum pedal and beater, rules (a) and (d) do not apply to this drum.

4.3.2 Equipment

A full list of drum set instruments can be seen in Table 4.1. Each drum sound

[42] The sound absorbing curtains and tracks were made by JANDS (www.jands.com.au).

EQUIPMENT LIST			
Instrument	**Diameter (in)**	**Depth (in)**	**Height (in)**
Bass drum	22	16	23
Snare drum	14	7	27
Hi-hat	14	3	32
Floor tom	16	25	23
Low tom	12	11	36
Medium tom	13	8	36
High tom	10	10	38
Ride cymbal	20	3	48
Crash cymbal	16	3	49

Table 4.1: A list of drums used, their size, depth, and height (top to the floor).

was recorded using two Neumann KM140s, due to their cardioid characteristics, non-coloured reproduction, and suitability for close-miking with percussion. In addition, these were also chosen due to the quality of the reproduction against other microphones.[43] The microphones were positioned in an X-Y configuration and were positioned exactly 11" (approx 30cm) above the centre of each drum. Other equipment used during the recording process included:

- 2 x Canon HG21 HDD video Cameras. One camera was situated directly above the drum to capture the strikes from above, and the other camera was placed to the side to capture an alternative angle;
- 1 x 24 Apple iMac; and
- 1 x Millennia pre-amp.

All strikes were recorded continuously (in one take) into Adobe Audition as 44.1kHz, 16-bit, .wav files. Some photographs showing the experimental setup are shown in Figures 4.5 - 4.11. The low, medium and high toms were kept on their respective stands during the capture of the strikes, to ensure that accurate representations of the instrumental decay through the stands were maintained. The low and medium toms were mounted on the bass drum, and the hi-tom was mounted on the crash cymbal.

[43] A pair of B&K pressure sensitive microphones and a HATS device was also used to record the individual strikes, although the audio was discarded in favour of the audio from the KM140s.

Figure 4.5: Microphone placement above the ride cymbal.

Figure 4.6: Microphone placement for the bass drum.

Figure 4.7: Microphone placement above the low tom.

Figure 4.8: Microphone placement above the medium tom.

Figure 4.9: Microphone placement for the high tom.

Figure 4.10: Microphone placement above the crash cymbal.

Figure 4.11: Microphone placement for the hi-hat. Note that the placement is slightly off-centre to avoid contact with the centre-rod.

4.3.3 Data Preparation

Once 1,000 drum sounds were recorded, the resulting audio files were imported into Pro Tools 9, where the attack points were identified using Pro Tools “Beat Detective”. Each attack point (the first zero crossing before the attack) was then visually accounted for, and exported as a region into a separate audio file in a wav format. Each of the individual audio files (samples) was manually accounted for and truncated in Wavelab 5, in order to compensate for perceived differences in the sample end point, between the strike during recording (see procedural rule (f) above) and the actual recording and representation of the recording in a digital system. There were no changes to the audio file format, sample-rate, or bit depth during the use of Pro Tools or Wavelab.

4.3.4 Feature Extraction

The feature extraction method is outlined in section 4.2.2. MIRToolbox was chosen to analyse the audio samples, owing to the large number of musical feature extraction functions, particularly those related to timbre (*mirrms*, *mirbrightness*, *mirroughness*, *mircentroid*). MIRToolbox was also chosen because of its capability to conduct batch operations, as well as its capacity for a range of analysis output formats.

4.3.5 Analysis

Although it was noted in section 3.3.4 that loudness, spectral centroid, and spectral flatness were the key features to be used in this model, additional features were extracted. The reason for this is two-fold: firstly, at the time of analysis the efficacy of the approach and analysis of the instruments were untested and, in the event that analysis of these features failed, time was saved in the re-analysis; and secondly, having completed the analysis, data exists for further work, where different features can be applied to the model. The features extracted from the audio files can be seen in Table 4.2.[44]

Feature	**Measurement/Output**
Loudness (RMS)	DBFS
Bark Envelope	Float value (0.0=min, 1.0=max)
Brightness	% of spectral energy >1500Hz (0.0=min, 1.0=max)
Centroid	Frequency (Hz)
Cepstrum	Quefrency (Hz)
Flatness	Float value (0.0=min, 1.0=max)
Roughness	Float value (0.0=min, 1.0=max)

Table 4.2: A list of features extracted from the audio files using the MIRToolbox.

The resultant values of the measurements for each instrumental sample were aggregated into a text file representing each instrument, and are shown in Appendix B.1. These text files were then used to evaluate the variation in features among the captured strikes.

4.3.6 Data Preparation

Once the audio files were analysed using MIRToolbox, the data showed that a number of individual samples had failed to produce a valid measurement. However, these processing failures amounted to only 0.04% of the total samples processed, well within the +5% tolerance that was factored into the number of strikes taken. Samples that produced failed measurements were removed from the database.

[44] All values are a mean average of the total audio file.

4.4 A Method for Representing Micro-Timing

This section describes the timing paradigm for the performance model. This will include describing the methods and tools used in the extraction of empirical performance data.

4.4.1 Timing "Fingerprints"

Each person has their own kinematic "fingerprint", as a result of their unique anatomical proportions, dimensions, and physical constraints (Runeson and Frykholm, 1983, p. 592). Musically (and by no means strictly limited to physiological or kinematic conditions), each drummer has their own unique way of playing the drums. This is manifested as "the unique time-feel of each person that is like a fingerprint or signature" (Keil, 1995, p. 11). This model used extracted timing fingerprints of three different skill levels of drummer. Capturing and analysing percussive performance from drummers of different skill levels allows for a comparative analysis of the performances, in order to establish general causes and effects between skill level, performance, and timing. In order to achieve this, the selected skill levels under investigation are: unskilled, semi-skilled, and skilled.

These timing fingerprints were extracted from recordings of the performances of three jazz drummers, playing to reference material in a controlled environment. The onset times of every strike in the performances were compared to the reference material, in order to determine the exact temporal deviations of each participant. The timing fingerprints represent these unique differences in temporal performance. Analysis of the participants' recordings were done using Sonic Visualiser, and are described in greater detail in the next section. These fingerprints were then analysed, and the temporal variation of the model was derived from these timing fingerprints. These timing fingerprints contain both positive and negative asynchronies, and also include both systematic and expressive variations. As a result, the model is able to produce both early and late temporal values, and the representative method for comparing the onsets to the reference material was capable of incorporating both types of variations.

However, for a more accurate assessment of the performances, it was necessary to capture additional performance data, such as video information and sensor data, to assess performance movement and the underlying causes of performance variation. Beyond the creation of timing fingerprints via recorded audio the observation of performance from video and the analysis of movement from sensor data, assisted in the development of performance rules, which more accurately represent real human percussive performance. The data extraction methods for the audio, video, and accelerometer data are described later in this chapter.

4.4.2 Temporal Data Extraction

Another system developed for analysing the content of audio is Sonic Visualiser (Cannam et al., 2006). Sonic Visualiser allows for visual multidimensional analysis of several simultaneous features using layers, and the subsequent extraction of information from each layer. Although spectral analysis can be undertaken in Sonic Visualiser, one of the strengths of this system is the beat detection function, which enables comparative beat detection of different layers and the export of each beat as a time value. In addition, the markers that represent the beats are manually adjustable, allowing for visual inspection should the beat tracking algorithm fail to accurately identify a beat (as in the case of Naveda et al., 2009). These advantages have led to Sonic Visualiser being used in the analysis of micro-timing and expressive asynchronies in different musical performance (Dodson, 2011; Naveda et al., 2009, 2011). In the case of extracting timing information from real performance, Sonic Visualiser was used to extract the onset times from the performances, relative to the source material.

4.5 Data Collection: Performance and Micro-Timing

The first phase of the experiment was the collection of the 9,000 individual strikes of instruments of a nine-piece jazz drum set, and the analysis of the sample data. The second phase of the construction of the performance model involved the collection of data from real human performance, to construct timing fingerprints, and to generate rules for the implementation of the model. This was done using captured audio, video, and accelerometer data of three drummers of differing skill levels. The following sections will describe the participants, methods, protocols, and procedures used in capturing the performances, and present an analysis of the data in order to generate the rules for the model.

4.5.1 Participant Overview

In order to collect the timing fingerprints for the model, and to collect performance data for rule generation, it was necessary to recruit three participants with different levels of drumming skill. Potential participants were asked to complete a questionnaire, with the selected participants chosen based upon their drumming experience and preferred musical style. The three selected participant responses to the questionnaire are shown in Appendix B.2.[45] The questionnaires revealed different levels of playing experience in different styles, and with differing levels of musical education.

Although each of the participants had experience in composing and playing music in different genres, both acoustically and electronically, their experience in playing the jazz drum set differed significantly, from no experience, some experience, to professional experience. The skilled drummer was the only participant that could read percussive notation. Their musical experience indicated that each participant had at least a rudimentary understanding of rhythm and timing, and knowledge of the role of, and constitution of, the drums despite different experiences in their physical performance. As a result, these participants made suitable references for a comparative analysis for rule generation.

4.5.2 Audio Data Collection

In order to create a timing fingerprint, the timing deviation values of each strike, relative to a specified beat, were obtained. These values were then used to represent the timing fingerprint, as each performer had different values. This was done using audio recordings of the performances. Previous studies that have examined human musical performance focussed on IOIs and note durations of recorded audio to determine deviations in timing and synchronisation of musical performance (Butterfield, 2010; Friberg and Sundstrom, 2002; Kilchenmann and Senn, 2011; Petrini et al., 2009; Prögler, 1995). A similar approach was taken here, as this was a previously tested empirical method.

Other methods, such as capturing MIDI performance data, present problems with high latencies. A latency of 5ms (Repp, 1994) can significantly affect the timing fingerprints, which can increase during the capture of fast performances owing to the serial nature and baud rate of the MIDI protocol. However, disadvantages to using recorded audio are the separation of strikes from waveform data, which can be labour intensive, and the difficulty in extracting performance dynamics data (Goebl et al., 2005).

Since not all participants had sight-reading ability, it was concluded that each participant would play along with an audio recording of a jazz piece for reference, in order

[45] Recruitment and participation of the three drummers was subject to the University of Sydney Human Ethics Committee, protocol number: 12943 (2012/1714).

to maintain similar experimental conditions. The onsets of each strike were analysed in relation to a defined "beat" in the reference material in order generate the timing fingerprints.

4.5.3 Movement Data Collection

In order to infer performance context, it is important to understand how each performer moved at various points during the performance. To perform such an analysis, video data of each performance was captured. In order to understand the movement context, the angles of trajectory and strike velocity, and data from hand-mounted dual-axis accelerometers was also captured. The use of accelerometers as a direct measurement technique of human movement, in gesture tracking of percussive performances, and in musical interface design, is well documented and, in some cases, these are augmented with video recordings (Dolhansky et al., 2011; Friberg and Sundberg, 1999; Hajian et al., 1997; Tindale et al., 2005b; Wagner, 2006; Winter, 2009; Young and Fujinaga, 2004).

4.5.3.1 Audio Reference Material

The audio reference material chosen was *Bird's Lament*, by Moondog (Hardin, 1969) and is shown in Appendix B.3. The piece was chosen for this investigation because it is well known, highly expressive, with an isochronous pulse originating from the basso ostinato (see also Figure 4.12), and has different metrical levels that allows performers of all skill levels to play along. In addition, the short duration of the track was of practical benefit for data management, given the potential to extract large quantities of performance information.[46]

4.5.3.2 Protocol and Procedure

All of the performances were recorded in the recording environment described in section 4.3.2., using the same equipment. The microphones were placed above the drum set to capture the whole performance. Participants were given a ten-second preview of the reference material via headphones. Following the preview, the reference material was then played in full through the headphones (no monitoring of the performance was played through the headphones), and the participants were asked to play along as closely

[46] Philosophically, the profound influence of African music on jazz allows the model to indirectly embody this influence through the choice of reference material. This influence will metaphysically augment the meso-periodic compositional approach to rhythm.

as possible. Participants played wearing gloves that had been modified to include small pockets sewn into the back of the hand to hold the dual axis-accelerometers. The wires from the accelerometers were then taped to the participants' arms, so as to minimise interference during performance.

4.5.3.3 Equipment

The same drum set used in the sample collection methods described in section 4.3.2 was used in the performances. The audio of each performance was recorded into Pro Tools on a 24" Apple iMac via a Millennia pre-amp as 44.1kHz, 16-bit, .wav files, using a pair of Neumann KM140s positioned over the drum set. The microphones and their positioning were favoured over the individual recording of instruments with separate microphones for two reasons. Firstly, it reduced the number of microphones required to collect the audio data. Secondly the audio captured during a pilot was sufficient for analysis purposes, as it distinctively captured each strike. In addition, high quality audio data was not required, because only the extracted timing deviation values are to be used in the model.

The video was captured using two Canon HG21 HDD video cameras. One was placed directly above each participant, and the other to the side of each participant. This approach, previously employed in section 4.3.2, captured the performance movement in the different planes of rotation.

The hands were measured using two T302D/E dual-axis accelerometers,[47] which measured the roll (X axis) and the tilt (Y axis) of each of the participant's hands during the performance. The physical range of the accelerometers was $+/-5$g on both axes, with a response time of 20ms. The accelerometer data were read by an ElectoTap Teabox[48] sensor interface at 4000kHz, which transmitted the sensor data optically to a MacBook Pro, which was then recorded into Max/MSP using the Tap Tools suite of objects. The sensor data were then saved in text files for analysis purposes. The analysis algorithm is described in more detail in section 4.5.4.2.

4.5.4 Data Extraction Methods

Once audio recordings of the participants' performances were captured, they were imported into Sonic Visualiser together with the reference material. The bar and beat

[47] Manufactured by Analog Devices (www.analog.com/en/index.html).
[48] Full technical details of the ElectroTap Teabox can be found at (www.electrotap.com/teabox/).

tracker analysis plugin was used to plot time instants of beats per bar. This is shown in Figure 4.12. The exact timings for each instance of the beat were then exported as a text file. Each of the performance audio files was imported to Sonic Visualiser and, using an onset detection algorithm, each strike was initially highlighted and then personally verified. A low onset detection threshold was initially used, and any unidentified onsets were manually added. This is shown in Figure 4.13.

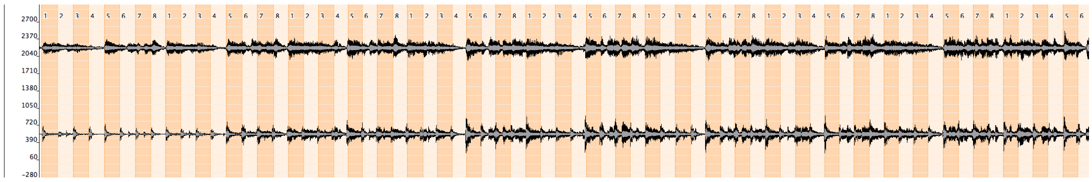


Figure 4.12: The beat tracked audio reference material in Sonic Visualiser.

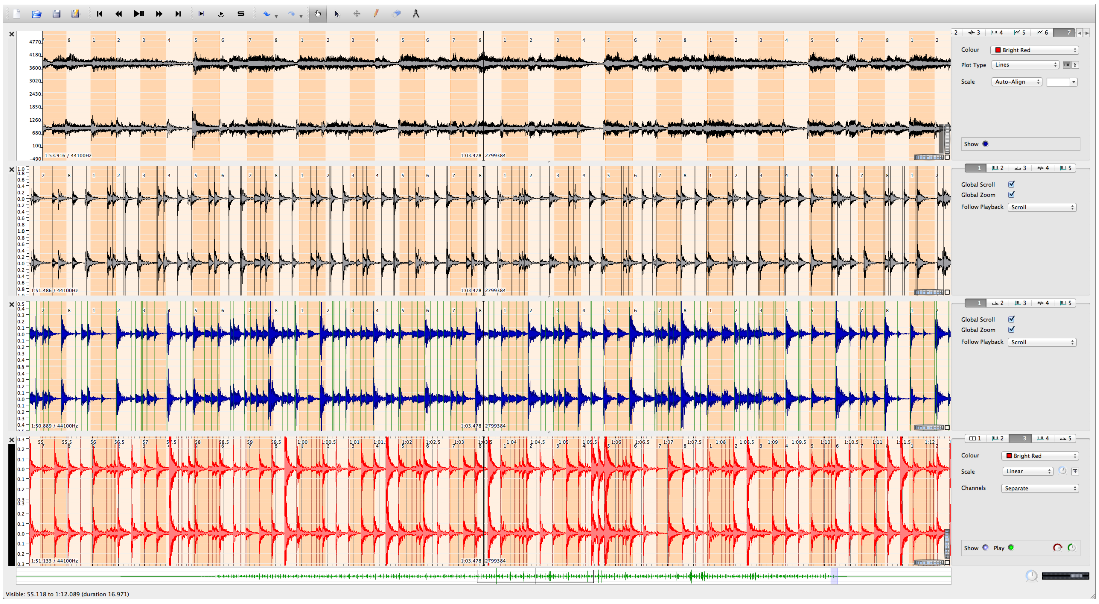


Figure 4.13: A screenshot showing the beat-tracked participant performances in Sonic Visualiser. The top pane shows the reference material from which beats are indicated in alternating background colours. The second (grey) pane shows the unskilled participant, the third (blue) pane shows the performance audio from the semi-skilled participant, and the bottom (red) pane shows the performance audio from the skilled participant. Each vertical line in the participant panes represents the onset of a drum strike.

The exact temporal locations of the start of every strike were exported into text files, together with the tempo at each strike point relative to the last strike. The relative strike-to-strike tempo was extracted in Sonic Visualiser by obtaining the temporal location of each strike, and calculating the tempo relative to the reference tempo, in order to determine the stability of the strikes.

4.5.4.1 Video Recordings

The video recordings were copied from the camera and visually inspected. In addition, the video data was analysed through the Jitter component of Max/MSP. Movement levels were calculated by comparing the inter-frame movement in the video. The algorithm reads the video file, converts the video into greyscale, and then compares the current frame of the video with the previous frame using the Tap Tools object *tap.jit.motion* to output the global measure of per-pixel frame-to-frame difference. The output is then saved as a text file. This empirical extraction of movement allows for a comparative analysis between the three drummers. A screenshot of the Max/MSP video analysis algorithm is shown in Figure 4.14, and the patch is available in Appendix B.4.

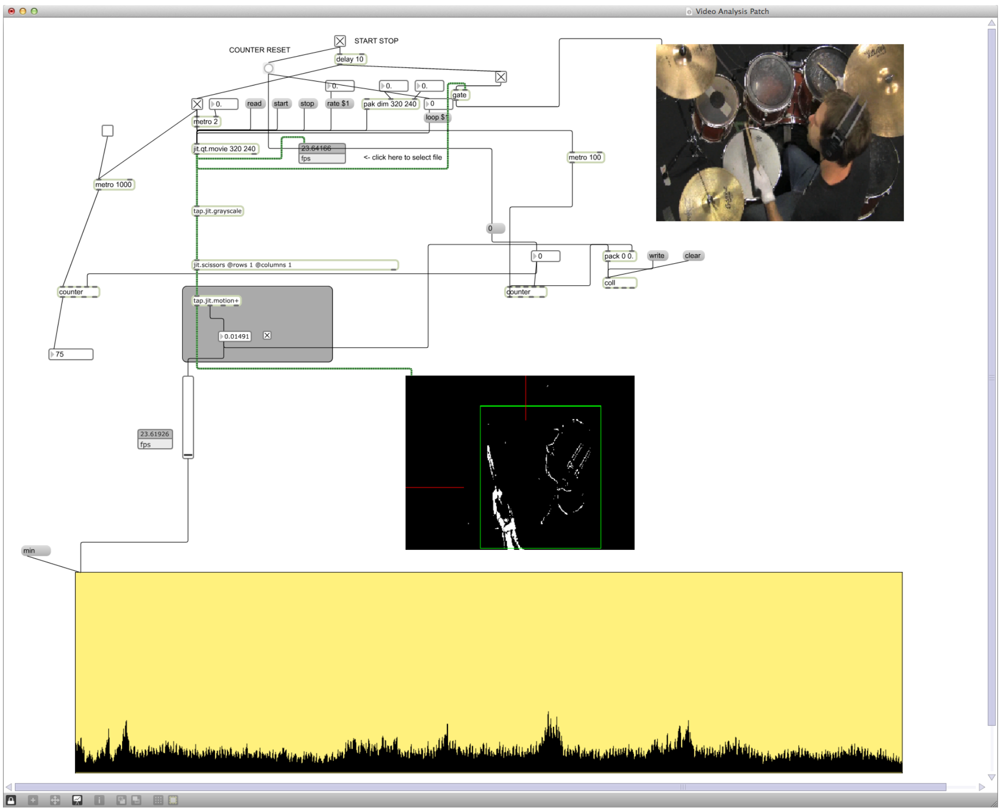


Figure 4.14: The Max/MSP video analysis algorithm.

4.5.4.2 Accelerometer Data

The accelerometer data was captured in Max/MSP at 20ms intervals and saved in time-stamped text files with the corresponding tilt and roll values. Both of the participants' hands were recorded simultaneously using objects from the Tap Tools suite of objects. A screenshot of the Max/MSP patch, which analysed the accelerometer data, is shown in Figure 4.15. The Max/MSP patch is available in Appendix B.5.

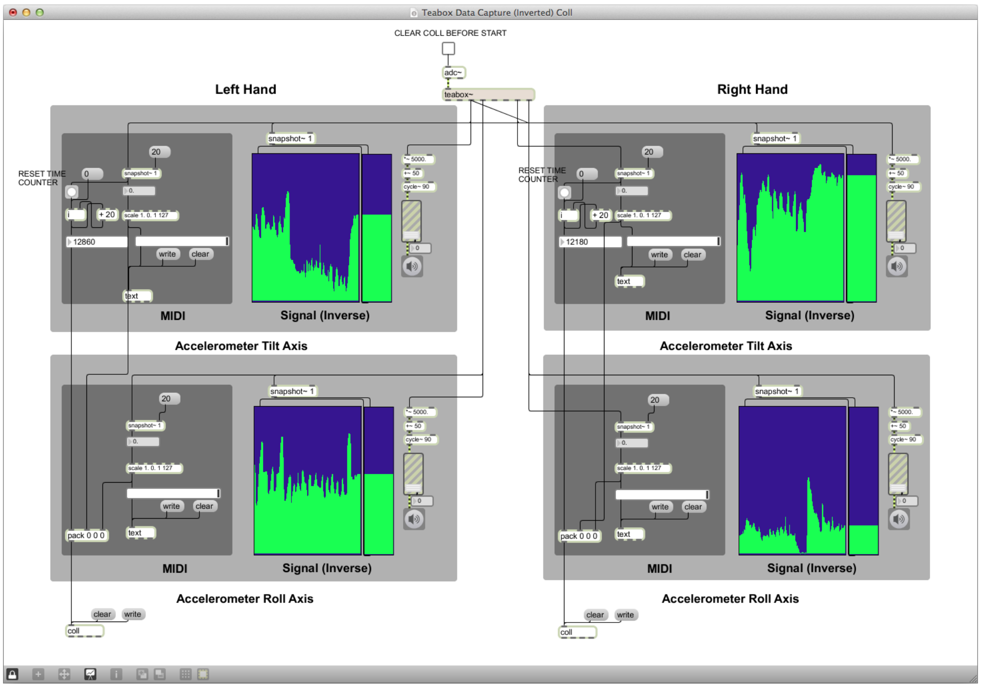


Figure 4.15: The Max/MSP patch used to capture the accelerometer data.

4.6 Preliminary Results

This section will provide a summary interpretation of the data analysis and, the relevance of these findings to the performance model. It will begin by analysing the timbral feature extraction of the sample database. The results of this discussion will inform the approach to using the sample data for both the performance model and the compositional approach.

The second part of this section will continue analysing the timing fingerprints of the participants. It will describe the differences between the percussive performances from a temporal perspective, and will define the implications of these differences in relation to the model. The timing fingerprint data is presented using graphs showing the temporal deviations from the reference beat and their respective inter-onset-interval times, coupled with time-series graphs showing the IOTs for each of the participants. These are shown in Figures 4.36 and Figures 4.37 to 4.39. Following on from the analysis of the timing fingerprints, attention will turn to an analysis of the movement levels between the participants, shown in Figure 4.40. These time-series representations of movement will add context to the timing fingerprints by way of identifying the global movement patterns of the participants relative to the timing fingerprint.

The final part of the analysis will focus on hand movement, and will involve the analysis of the accelerometer data for the participants' left and right hands. This will help to identify the relative hand/arm movements between the drummers, particularly in relation to the three dimensional performance space in which the performances were undertaken: a three-dimensional performance space determined by the instrumental configuration. This analysis will allow for localised three-dimensional movement variations to be considered in the model's rule generation process. The analysis undertaken here is not intended to be exhaustive. Instead, it seeks to provide a greater understanding of percussive performance across skill levels for generating computational rules for the performance model.

4.6.1 Initial Results of Instrumental Micro-Timbre

Results of the MIRToolbox analysis showed variations in loudness, flatness, and spectral centroid for each of the instruments' captured samples. These are shown in Figures 4.16 to 4.24, and are grouped by membranophones (Figures 4.16 to 4.21) and idiophones (Figures 4.22 to 4.24). The results demonstrate that the proposed methodology for capturing variations in percussive strikes produced a sample database that adequately represented the micro-timbral variations in each instrument.

Although it is not the intention of this section to explicitly describe these graphs, as this is beyond the scope of this investigation, the results show adequate variation in the ratings of samples, between the loudness, flatness, and spectral centroid features. A comprehensive range of values in each parameter has been captured, allowing for interesting compositional devices, such as crescendo, where the database is classified based upon loudness ratings.

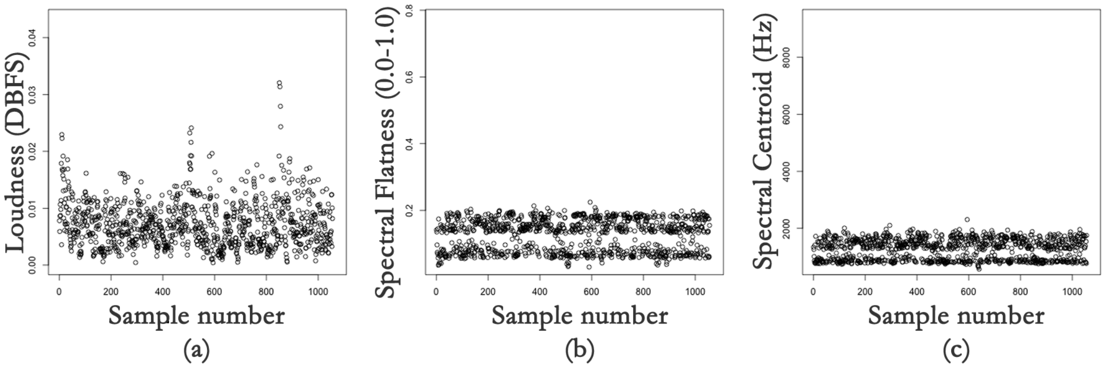


Figure 4.16: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the bass drum samples captured.

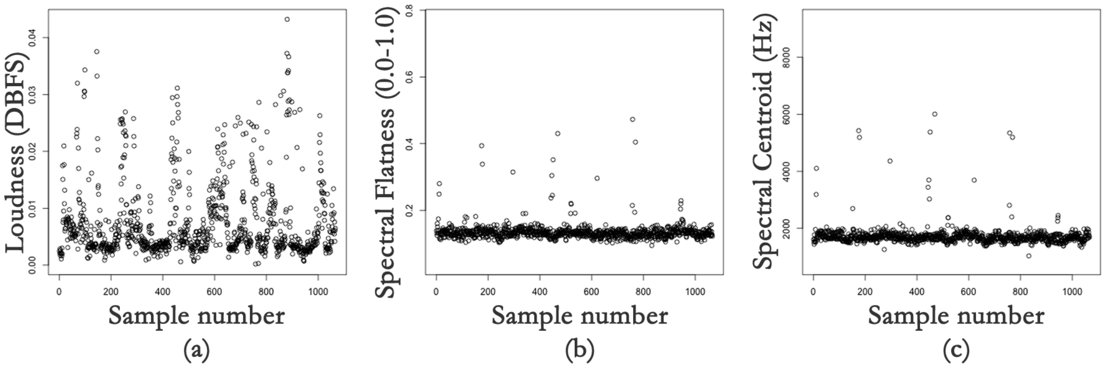


Figure 4.17: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the snare drum samples captured.

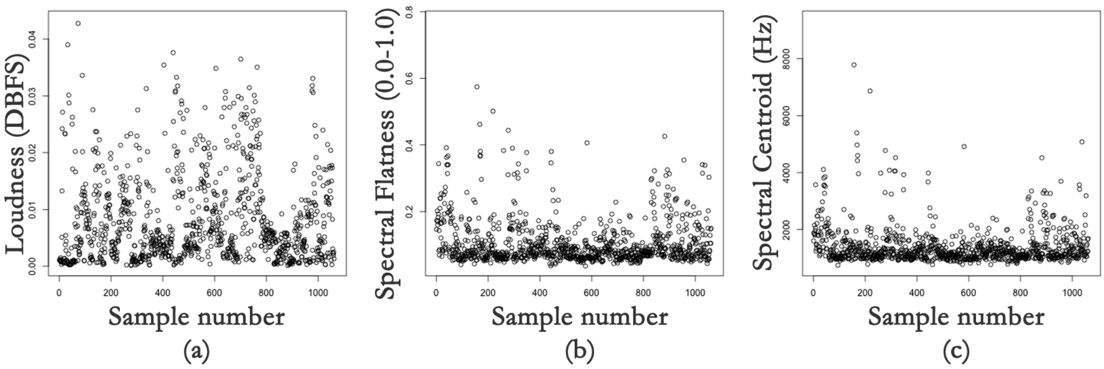


Figure 4.18: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the floor tom samples captured.

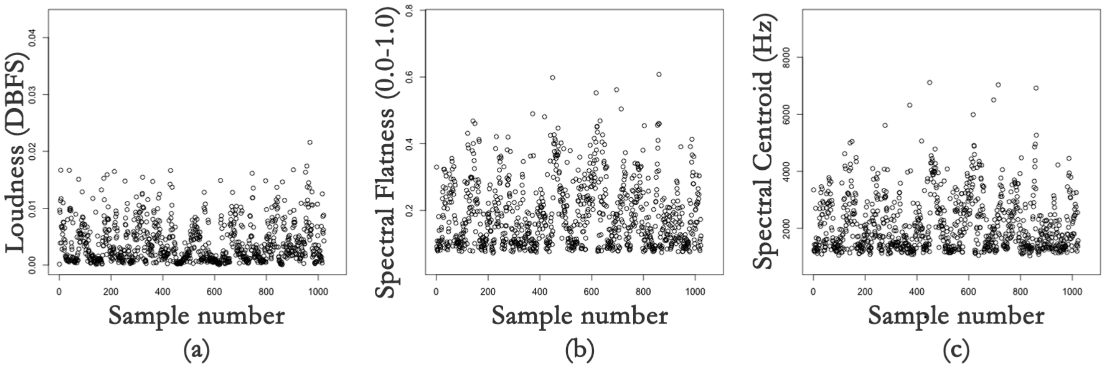


Figure 4.19: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the low tom samples captured.

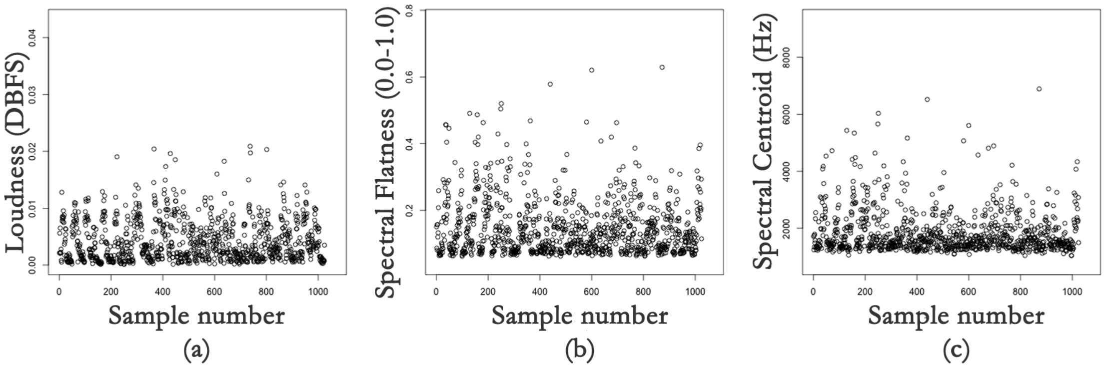


Figure 4.20: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the medium tom samples captured.

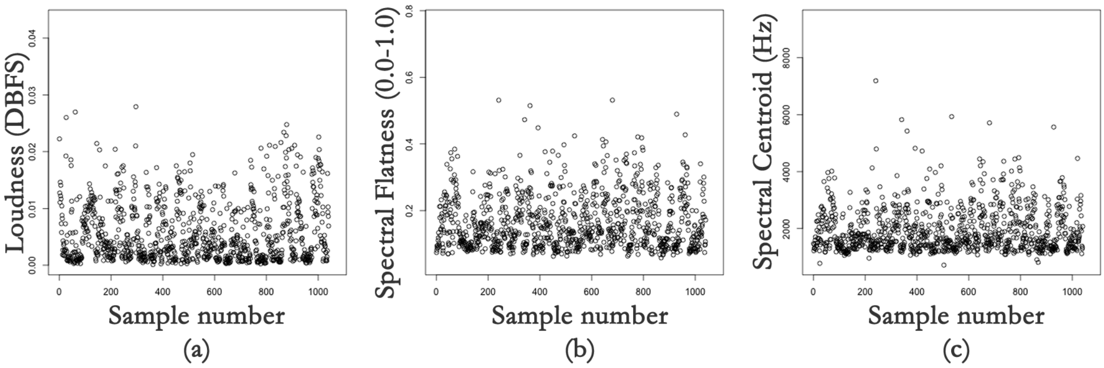


Figure 4.21: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the high tom samples captured.

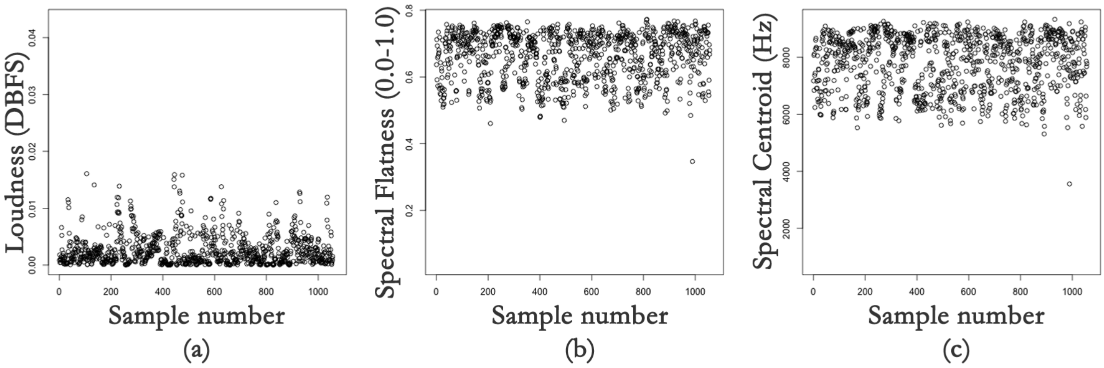


Figure 4.22: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the hi-hat samples captured.

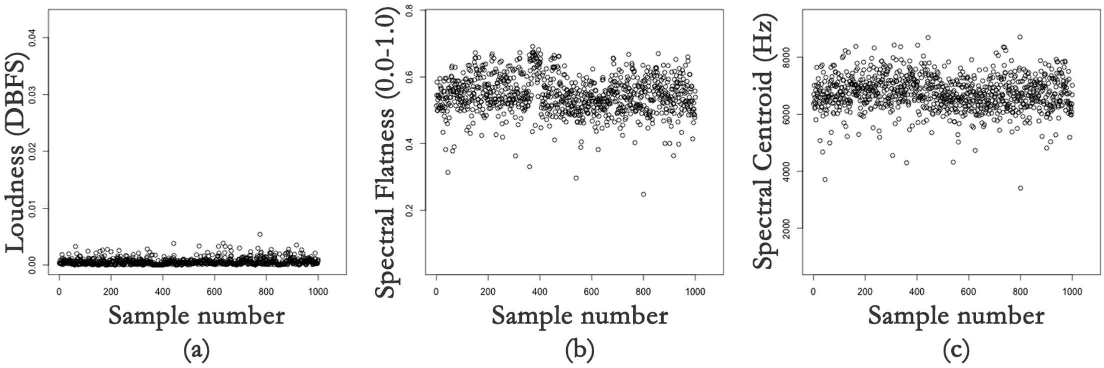


Figure 4.23: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the ride cymbal samples captured.

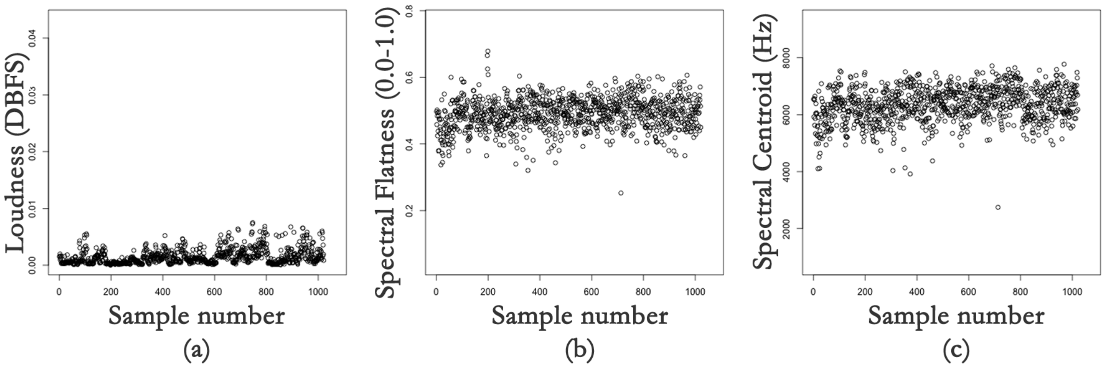


Figure 4.24: Variations in (a) loudness; (b) spectral flatness; and (c) spectral centroid in the crash cymbal samples captured.

The presence of parametric variation in the sample database, within and between features, indicates that any classification of parameters will feature different schematisations, resulting in differing sample selections. Furthermore, these findings also highlight the timbral variation inherent within each instruments' dataset. The findings also suggest that there are sufficient differences between each of the parameters to produce different database structures, particularly where the database is ordered by minimum/maximum feature values (as seen in most MIDI volume key assignments). This will result in different samples being played depending on the parameter, irrespective of whether the parameters are ordered in the same way. However, it is acknowledged that this is not a conclusive timbral representation of the instruments due to the limited feature extraction and the multi-dimensional nature of timbre. This limited representation of the instruments presents problems with the database in its current form.

These figures do not represent relationships between other parameters excluded from the analysis and, therefore, do not provide a comprehensively detailed timbral description of the strikes. As a result, it is possible that two samples with similar values in a database may exhibit differences in other timbral features. Conversely, it is also possible that two samples with different database values have similar secondary timbral features. Since one of the main objectives of the model is to simulate human percussive variation, any timbral variation between two strikes of similar value in these parameters will add to the perceived variability of human performance. However, such a large dataset has the potential to create too much variation, with three contextual effects: instrumental; performance; and compositional.

Instrumentally, each database is a representation of the sounds produced when striking across the entire surface of the instrument. In the case of membranophones, excitation location is a significant cause of the vibrational characteristics of the drum and, although locally these variations may be consistent, across the surface of a drum the vibrational differences may produce large differences in other timbral features. This is relevant when considering large differences between minimum and maximum values in a database. One example of this relates to the vibrational differences between strikes at the centre of the membrane and close to the rim, where membrane stiffness may be higher, increasing the pitch with the reflection of transversal waves affecting the decay time.

From a performance perspective there are two implications. Firstly, the dataset consists of samples at different strengths, causing changes in the vibrational behaviour and subsequently changes in secondary timbral features. One example of this relates to changes in frequency observed in tom-toms with greater strike strength, in which pitch was difficult to measure given the autocorrelation function of the analysis method. The selection of a sample with a higher variation in salient timbral feature may result in the model playing a strike that is out of context (e.g. the performance of an accent in a structurally atypical place). This is tightly linked to the second implication, which relates to problems in individual sample selection, where the control algorithm, as an abstraction of the performance context, must repeatedly select a specific value to play a sample, representative of the current performance context.

A 3D correlation of the spectral features, as shown in Figure 4.25, provides further insight into the timbre model, particularly in relation to instrumental behaviour and compositional implication. Concerning instrumental behaviour, there are clear correlations with centroid and flatness tending to increase with loudness, particularly among the idiophones. The tom-toms also display a characteristic “boomerang” or “hockey stick” curve, in which there is a tendency for higher levels of flatness with lower loudness, and a higher centroid with increased loudness. Similarly, the bass drum also displays higher flatness with lower loudness, although the correlation is less distinct compares with the tom-toms, possibly owing to the bass drum pedal mechanism. The snare drum also exhibits higher flatness with lower loudness, and the centroid is fairly consistent across the loudness and flatness dimensions. With instrumental behaviour governing the distribution of points in each of the instrument’s 3D loudness/timbre space, the sample selection paradigm allows for the arbitrary selection of any point within the timbre space, thus intersecting instrumental behaviour with composition. Understanding the correlation and between the spectral features of an instrument’s behaviour, also allows for greater compositional freedom, as correlated timbre space (e.g. higher loudness and high spectral centroid) can be used in conjunction with orthogonal point selection in more uncorrelated timbre space (e.g. lower loudness and high spectral flatness).

The potential timbral variations arising as a consequence of the sample collection methodology, supports the compositional validity of employing database classifications based upon the timbral features of loudness, flatness, and spectral centroid for the purposes of exploring the micro-timbre of the percussion instruments. Since the compositional perspective consists of an exploration of percussive micro-timbre using the

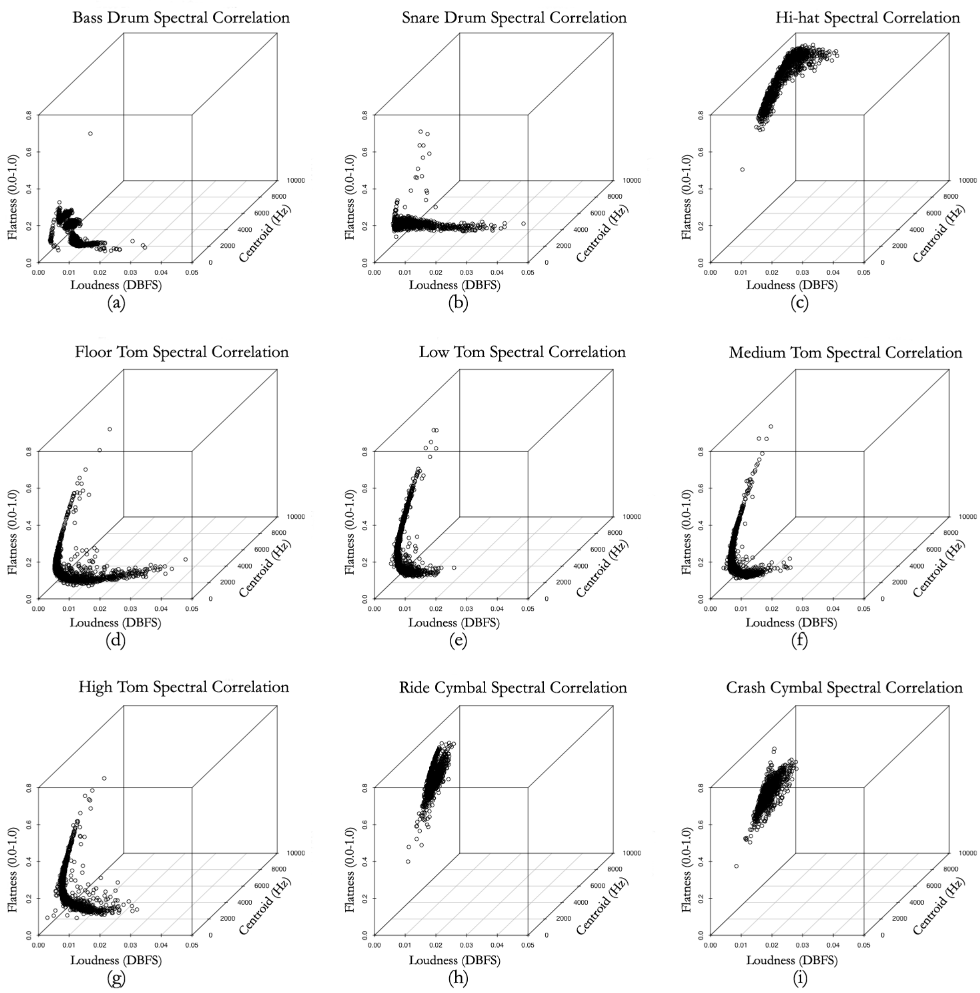


Figure 4.25: 3D correlations of the spectral features for (a) bass drum; (b) snare drum; and (c) hi-hat; (d) floor tom; (e) low tom; (f) medium tom; (g) high tom; (h) ride cymbal; (i) crash cymbal.

database schemas as compositional parameters, there is a necessity to mitigate the effect of uncorrelated salient timbral features, which are analogous to the implications in instrumental context. In summary, these findings support the use of these parameters for the modelling performance variation, and for compositional use, although the database must be timbrally constrained in a way that is useful in both applications. It is therefore necessary to perform data reduction on the dataset of each instrument in order to confine the timbral variation.

4.6.2 Data Reduction and Classification by Strike Location

The dataset of each instrument lacks both instrumental and performance context, which must be addressed by way of data reduction. One way the dataset can be reduced is to classify the samples according to strike location, where excitation location can potentially excite different modes of vibration, thus causing changes in timbre. This is particularly the case for membranophones, where the impact of dis-uniform tension is likely to produce localised timbral similarity. Furthermore, as vibrational modes operate both concentrically and diametrically, it is useful to consider such an approach to classifying the samples, where more complex movement affects strike accuracy on different radial planes on the skin, that traverse both diametric and concentric modes. Furthermore, a change in strike location may be deliberate in trajectory planning.

However, classifying samples based on existing vibrational modes are not conducive to inferring a performance context owing to extremes in precision that the modes would infer on performance, from a precise (6,3) mode (see Figure 3.4), to the less precise (1,1) mode (see Figure 3.3). Consideration must be given to the reclassified sample sizes. Therefore, a performance-based demarcation, adapted from Fletcher and Rossing's (2,2) mode, has been chosen to re-classify each instruments' samples. This is shown Figure 4.26, with: (a) the original (2,2) mode; (b) the adapted membranophones demarcation; and (c) the adapted idiophone demarcation. The centre area in (b) relates to the tendency for a drummer to strike in the centre and the option for centre spots in the tom-tom. This area corresponds to 1/3 of the diameter. In (c) the centre area represents the bell in the centre of the cymbal.

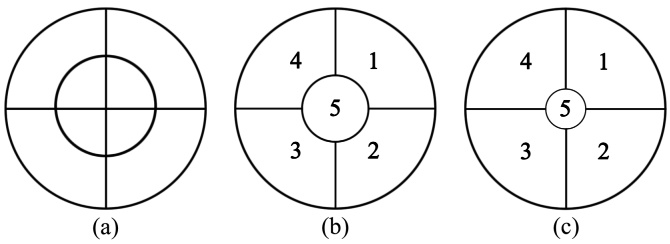


Figure 4.26: The (2,2) mode from Fletcher and Rossing (1998) (a); and the adapted demarcations for (b) membranophones and (c) idiophones.

To re-classify the sample database, a trace of each drum from the overhead video footage was put onto acetate, with the demarcation points calculated from scaled measurements of the video at full screen. Each strike was visually accounted for, and manually assigned a number based upon the location of the strike, relative to the demarcation zones for the respective instruments in (b) and (c) above. Since the strike location of the bass drum pedal is constrained to one location by the mechanism, such demarcation does not apply. The bass drum is the only instrument with such a large database. The resultant sample re-classification increased the total number of potential zones in which a performer can strike, and reduced the sample size in each zone, while maintaining adequate timbral variations in each zone. Graphical representations of each parameter and zone are shown in Figures 4.27 to 4.31 (membranophones) and Figures 4.32 to 4.34 (idiophones). This approach also facilitates the use of stochastic methods to infer performance inaccuracies, contextualising strike locations relative to other drums, with zonal striking weighted according to rules generated from the performance analysis.

It is not the intention of this investigation to describe the differences or trends between each of the instruments. However, there are a few observations that should be noted from the graphs, which are useful for this investigation. The findings suggest that in general there are consistencies between the parameters in each zone in each instrument. For example, the flatness and spectral centroid of the snare drum are generally the same across the five zones, with similar ranges between minimum and maximum values, even though each zone consists of different strikes. The data reduction method has retained a large diversity of values in some zones, most notably in the tom-toms and idiophones. Additionally, there are large differences between idiophones with the hi-hat having higher average flatness and spectral centroid values compared to the ride and crash cymbals, which display similar characteristics. This can be attributed to the interaction between the top and bottom cymbals. However, on listening to the membranophone samples there are clear differences between the five zones in other timbral attributes, such as pitch. These variations are related to the dis-uniform tuning of the drum and suggest that this approach to zonal demarcation is relative to timbral variation across the entire instrument.

A significant implication of this data reduction method relates to the allocation of samples to zones. The findings indicate that there are different numbers of samples in each demarcated zone. In some instances there are very few samples compared to other zones in the same instrument.[49] An example of this is shown in Figure 4.33 (d), ride cymbal mode 4. In the context of samples from this mode being played in sequence with samples from the other modes, timbral variations may be produced as values are

[49] This is a result of the random nature of the sample capture, in which data reduction via the demarcation points were neither required, nor conceived.

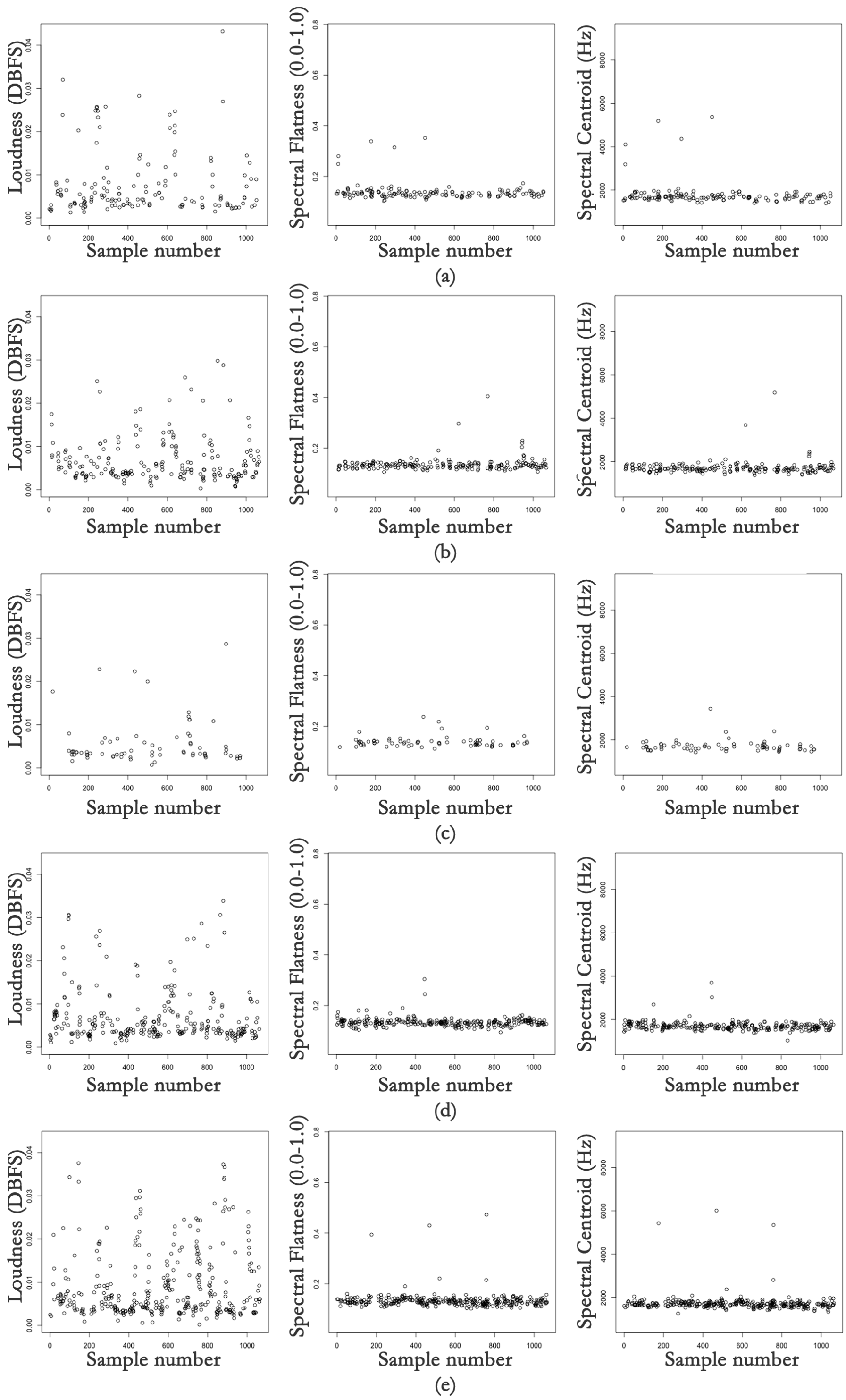


Figure 4.27: Snare drum sample features after modal demarcation (loudness, flatness, and centroid, left to right). Modes 1-5 are (a) to (e) respectively.

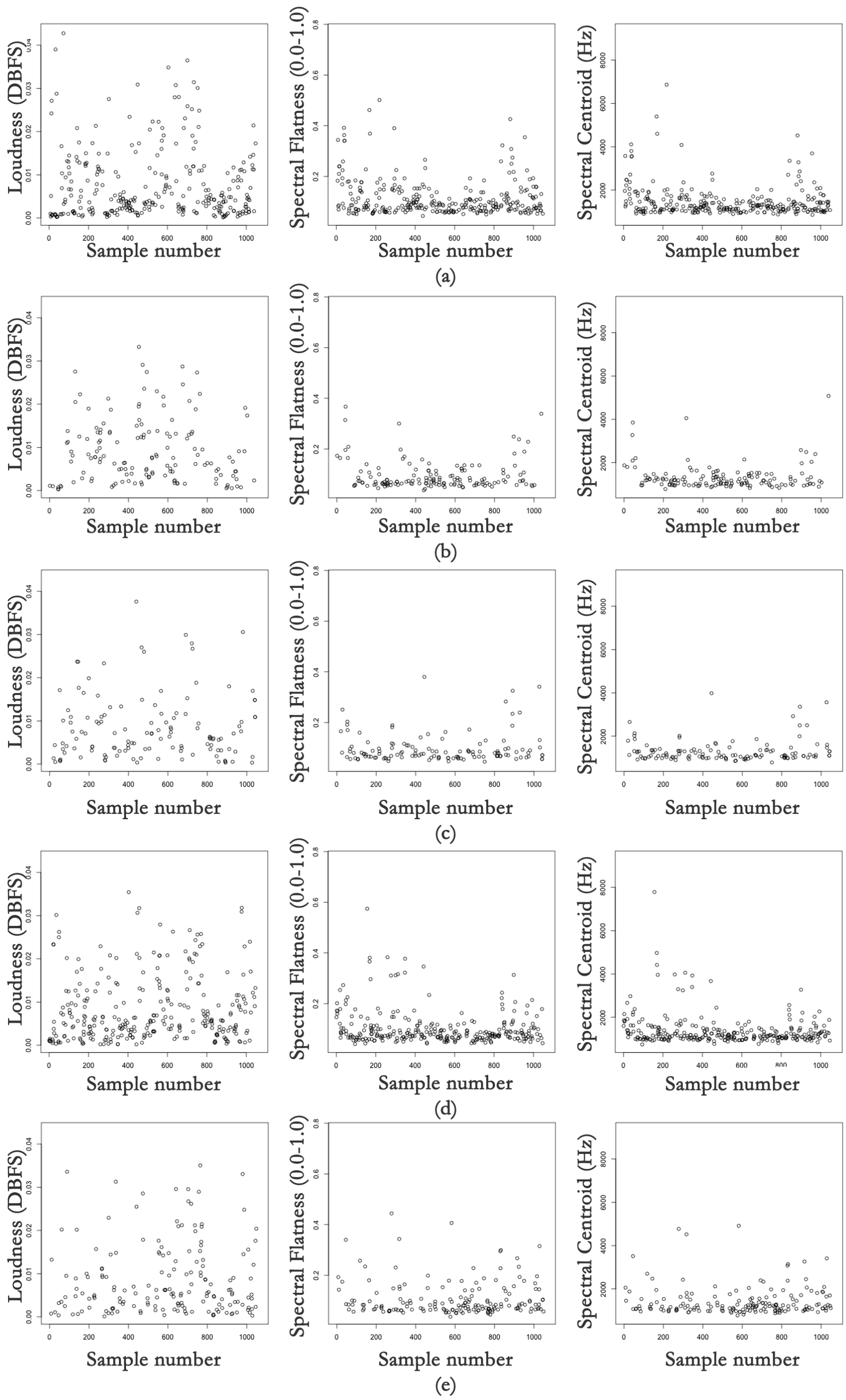


Figure 4.28: Floor tom sample features after modal demarcation (loudness, flatness, and centroid, left to right). Modes 1-5 are (a) to (e) respectively.

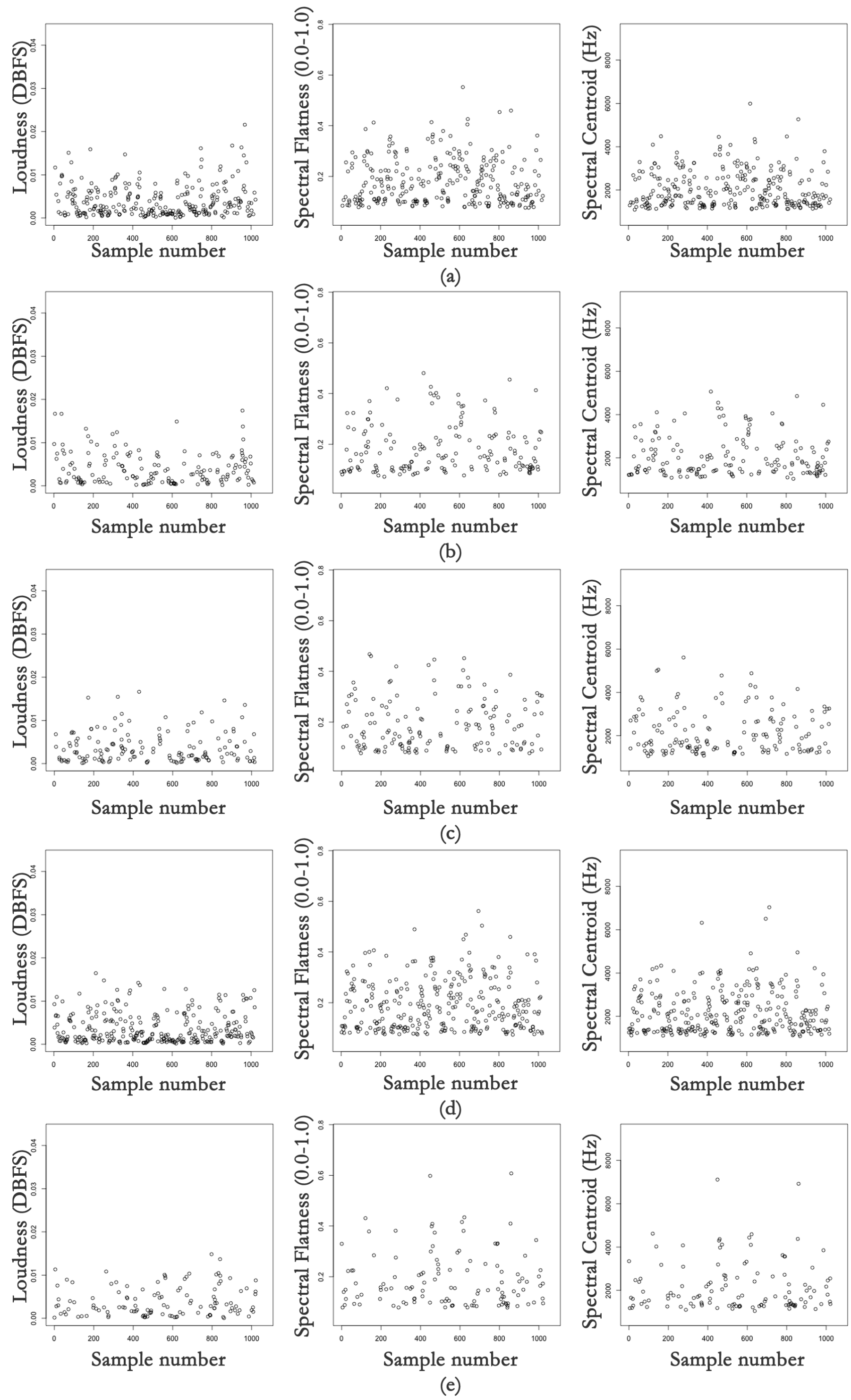


Figure 4.29: Low tom sample features after modal demarcation (loudness, flatness, and centroid, left to right). Modes 1-5 are (a) to (e) respectively.

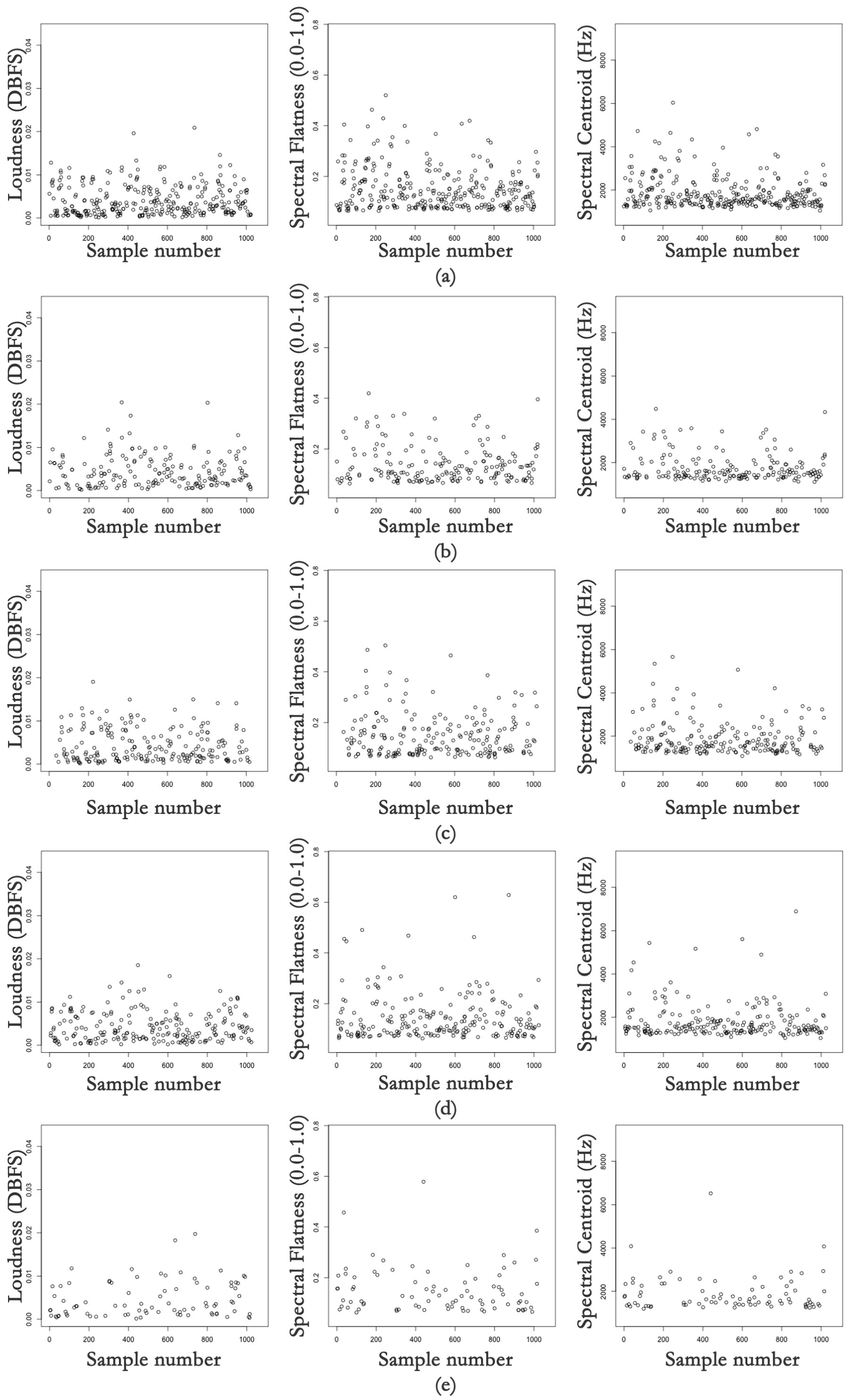


Figure 4.30: Medium tom sample features after modal demarcation (loudness, flatness, and centroid, left to right). Modes 1-5 are (a) to (e) respectively.

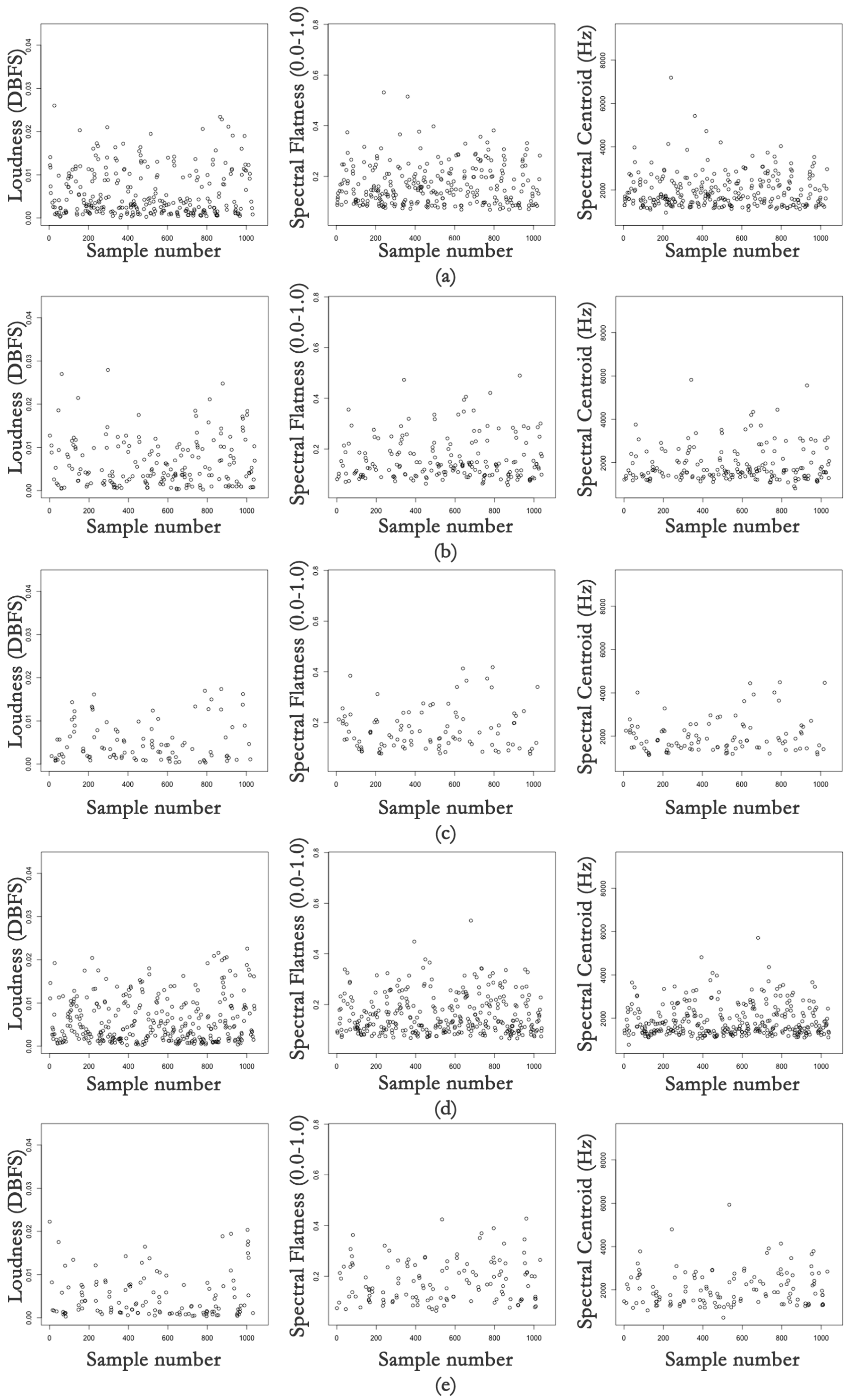


Figure 4.31: High tom sample features after modal demarcation (loudness, flatness, and centroid, left to right). Modes 1-5 are (a) to (e) respectively.

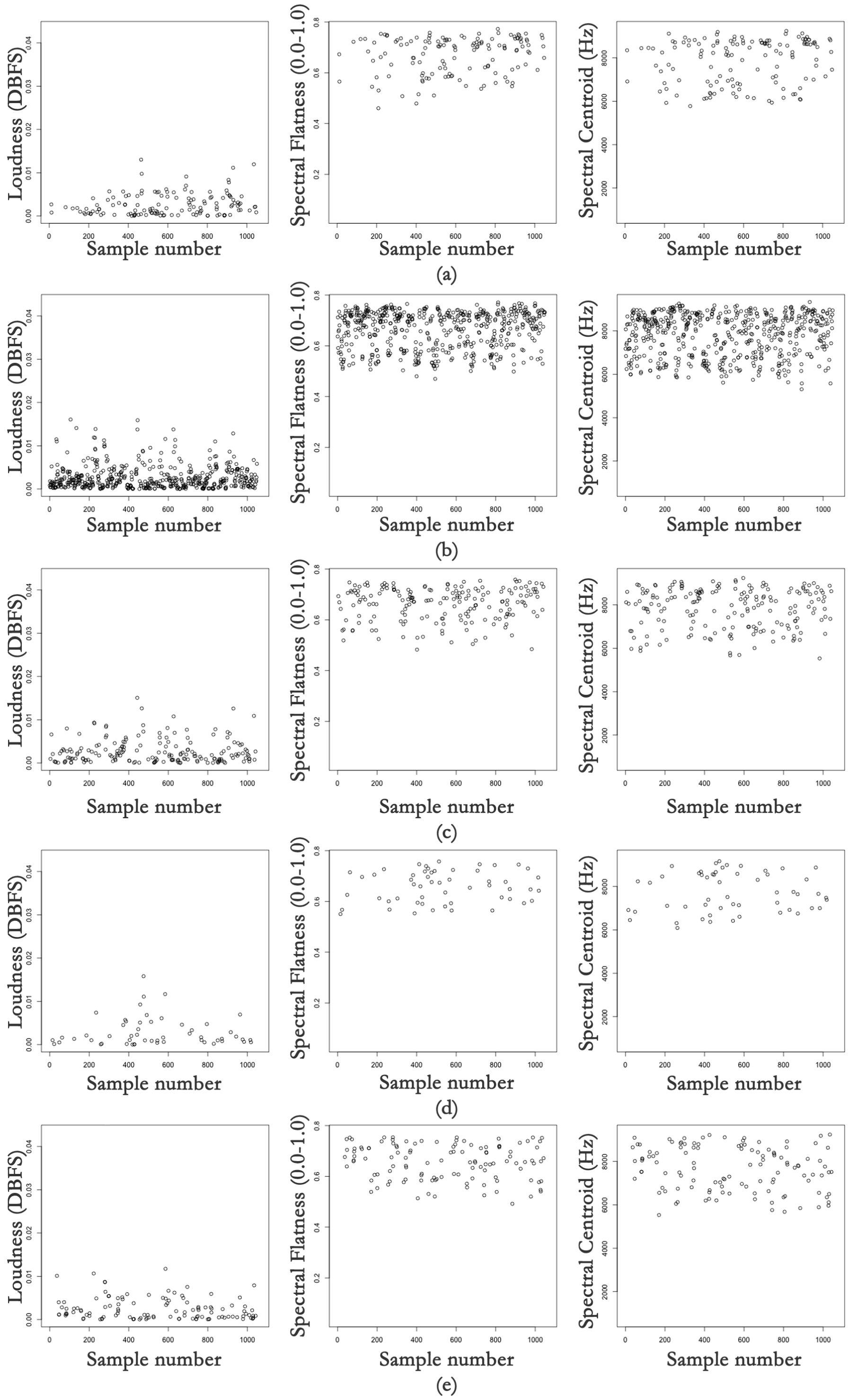


Figure 4.32: Hi-hat sample features after modal demarcation (loudness, flatness, and centroid, left to right). Modes 1-5 are (a) to (e) respectively.

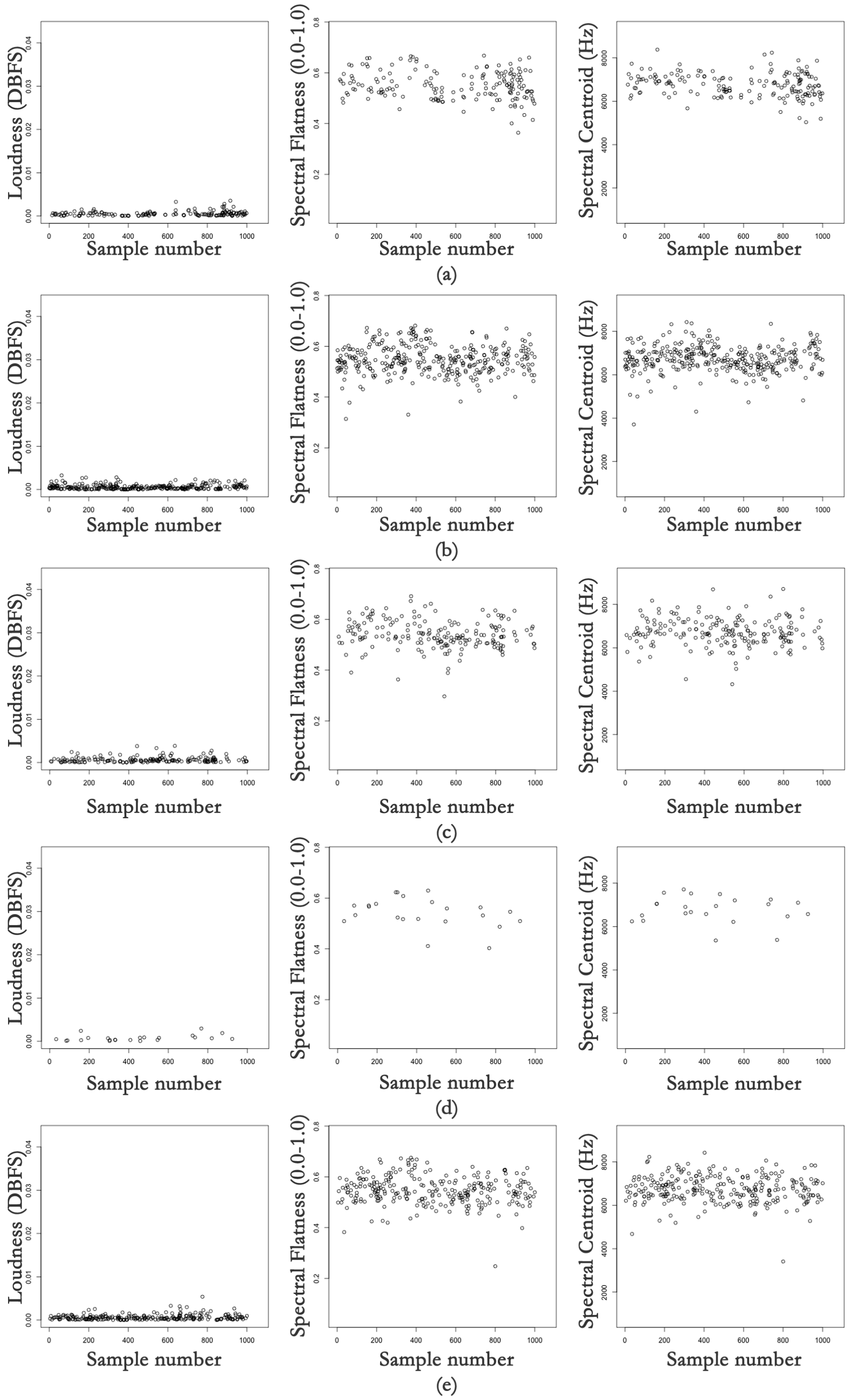


Figure 4.33: Ride cymbal sample features after modal demarcation (loudness, flatness, and centroid, left to right). Modes 1-5 are (a) to (e) respectively.

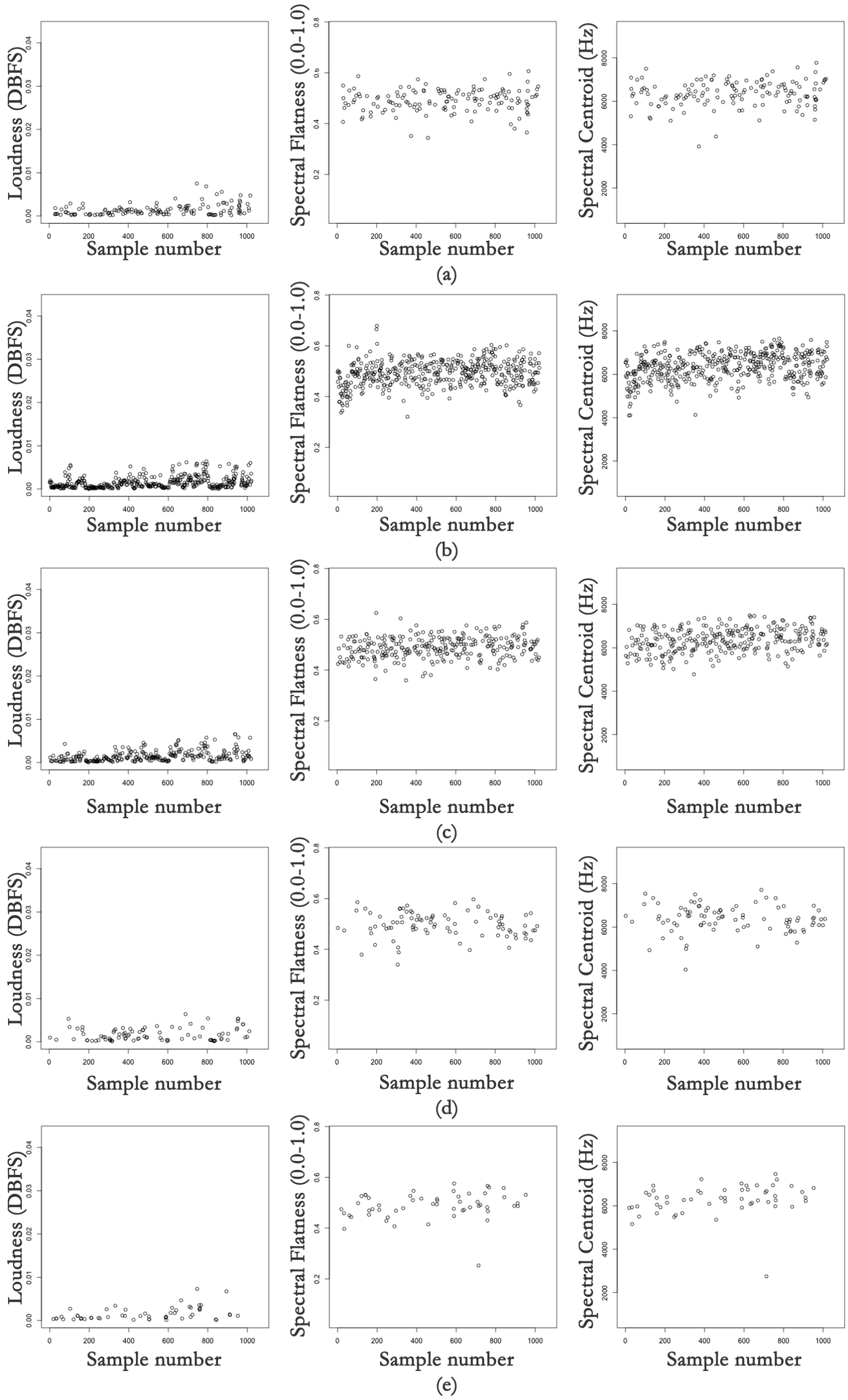


Figure 4.34: Crash cymbal sample features after modal demarcation (loudness, flatness, and centroid, left to right). Modes 1-5 are (a) to (e) respectively.

skipped in the parametric mapping, resulting in the selection of relatively higher (or lower) parametric values for similar mapping values. This will result in an ordered database of a zone producing variation between values in the same database, and between other similar values in other zones of the same instrument. One hypothesis is that the variations produced by the different sample numbers in each mode will produce variations significant enough to be considered "accented strikes", subsequently assisting in the model conveying a greater sense of human performance. This is in contrast to the complete dataset producing "irrational" strikes which, are strikes so timbrally and dynamically irregular that they convey a greater sense of artificial performance. It is expected that the data reduction method will localise secondary timbral features, thereby making the timbre of the samples in each zone more consistent, particularly for membranophones. The demarcation of the drums based upon a combined approach to instrumental mechanics and performance lends itself to the structural implementation of the audio in the performance model.

4.6.3 Timing Fingerprints

The onset deviation values from the reference beat (quarter notes) were obtained by calculating the mid-point of each reference beat IOI. Quarter notes were used as the reference because this was the regularly marked recurring beat (pulse) contained within the larger IOI durations for whole and half notes. Consequently, any data analysed at these higher metrical levels will include deviations from the more prominent quarter note beats. Each value within the temporal space between each IOI mid-point (the magnetic area) is then assigned to the reference beat within that area, and the relative asynchrony calculated as the difference in milliseconds from the reference beat position. This is illustrated in Figure 4.35.

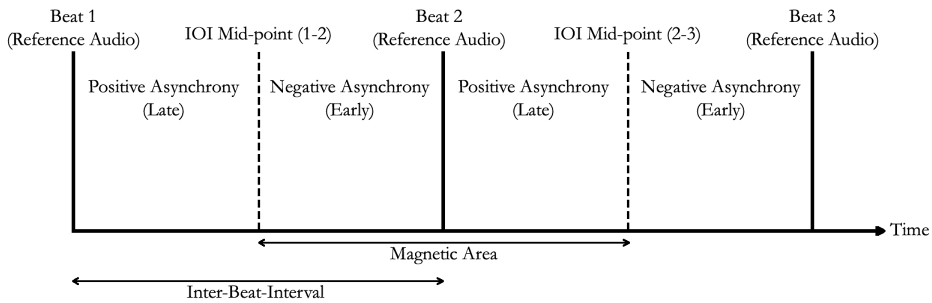


Figure 4.35: An illustration of the method for calculating strike onset asynchrony.

Since the values for the inter-onset-tempo (IOT) were previously calculated in Sonic Visualiser, it was possible to plot the asynchrony and IOT per strike, or timing fingerprint, for each performance. Figure 4.36 presents the timing "fingerprints" of the three participants. Timing "fingerprints" can be defined as a performer's unique temporal performance characteristics, and are shown as graphic representations in relation to tempo (BPM) and deviation value (ms). The findings suggest that in general, as the skill level of the drummer increased, the variations in tempo and timing reduced. This is apparent in the deviations from the reference beat, and in the increased stability and regularity of the IOT series across the participants, where the unskilled drummer displays the least discernable temporal strike patterns and IOT groupings.

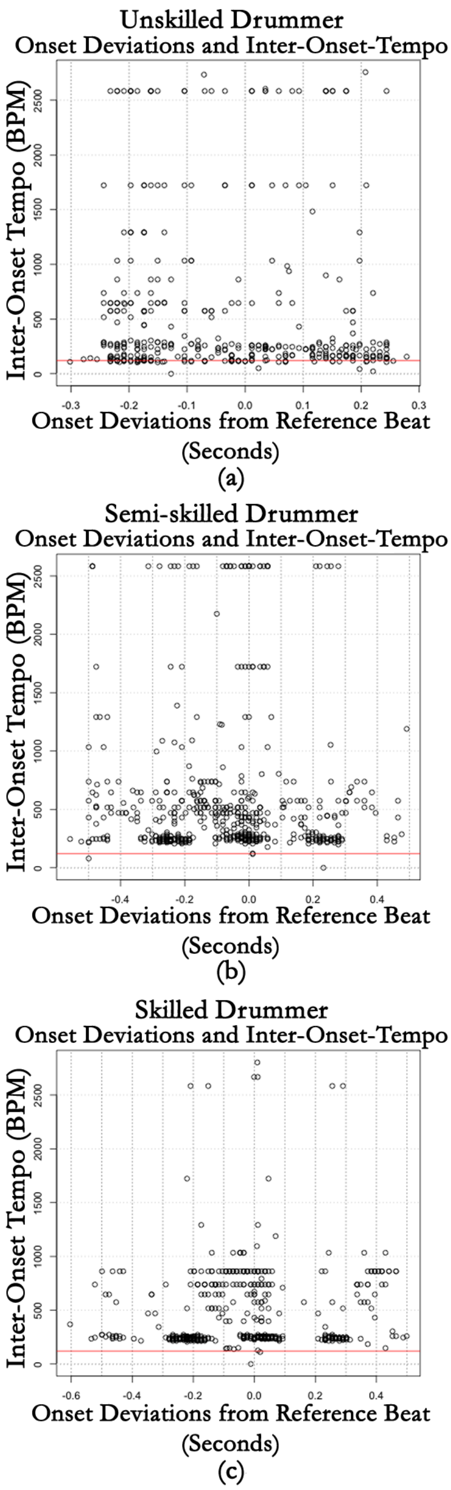


Figure 4.36: Timing fingerprints for the (a) unskilled; (b) semi-skilled; and (c) skilled drummers. The red line indicates the tempo of the reference material at 119BPM.

A pattern in onset deviations becomes apparent in Figure 4.36 (b), where the semi-skilled drummer shows three significant temporal strike locations: −0.25s, 0s, and +0.25s. From an IOT perspective, these temporal locations are also visible at high tempos. This finding indicates that the semi-skilled drummer is more temporally consistent in sequences of greater complexity. However, there is greater temporal instability at mid-range IOTs, suggesting that the participant may have had difficulty with certain aspects of the percussive sequence. These findings are also supported in the IOT time series graphs for the unskilled and semi-skilled drummers (Figures 4.37 and 4.38), where the semi-skilled drummer shows regions of IOT flatness at approximately 30-45 seconds, with some spikes in IOT indicative of periodic changes in strike density. Interestingly, the unskilled drummer is more consistently closer to the reference tempo, although this is largely due to double-time components in the reference audio in which the unskilled drummer was unable to execute. Relative to the reference beat, the unskilled drummer is inconsistent, and displays erratic IOTs between the reference and double-time tempos.

The onset deviations are more targeted for the skilled drummer in Figure 4.36 (c), indicating greater consistency across the strikes. Furthermore, there is significantly less IOT variability than both the unskilled and semi-skilled drummers. This is particularly visible at higher IOTs (>2,500) that are consistent with the onset deviations at lower IOTs (−0.2s, 0s, +0.2s). The IOT stability is supported by the greater consistency in the time-series graph (Figure 4.39), where strikes are temporally more regular with the regular spikes in tempo being indicative of recurring patterns that have consistent IOT times. In Figure 4.39 the skilled drummer has two prominent tempo spikes between approximately 75 and 85 seconds. Critically, the subsequent IOTs very quickly become stable, a feature that is not as prominent in either the semi-skilled or skilled drummers, whose fast IOT times are often succeeded by IOT instability with prominent increases and decreases in IOT. For the unskilled drummer this is apparent at approximately 26 and 46 seconds, and for the semi-skilled drummer this is apparent at approximately 93 seconds.

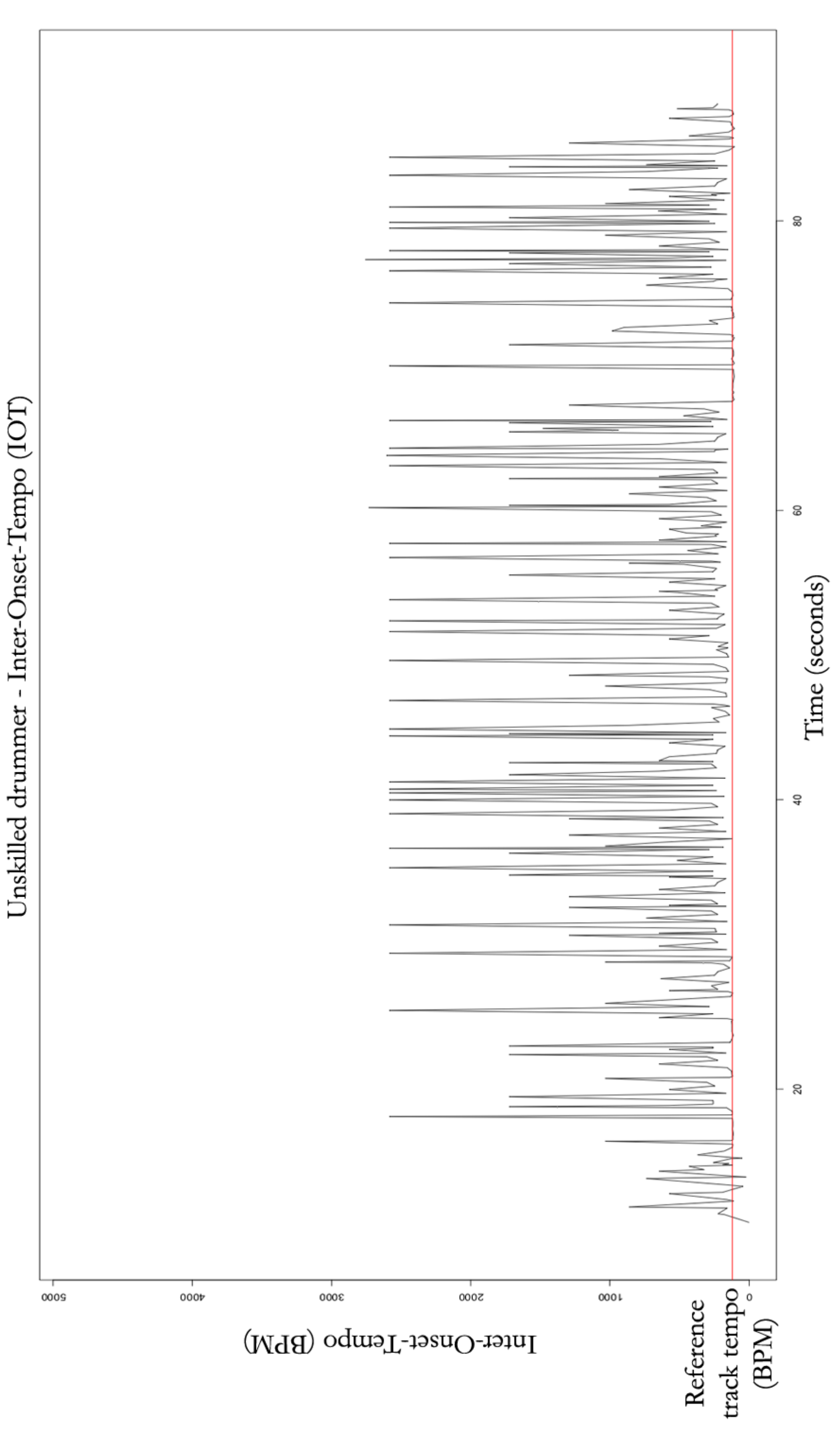


Figure 4.37: Inter-Onset-Tempo of the performance by the unskilled drummer. The red line indicates the tempo of the reference material (119BPM).

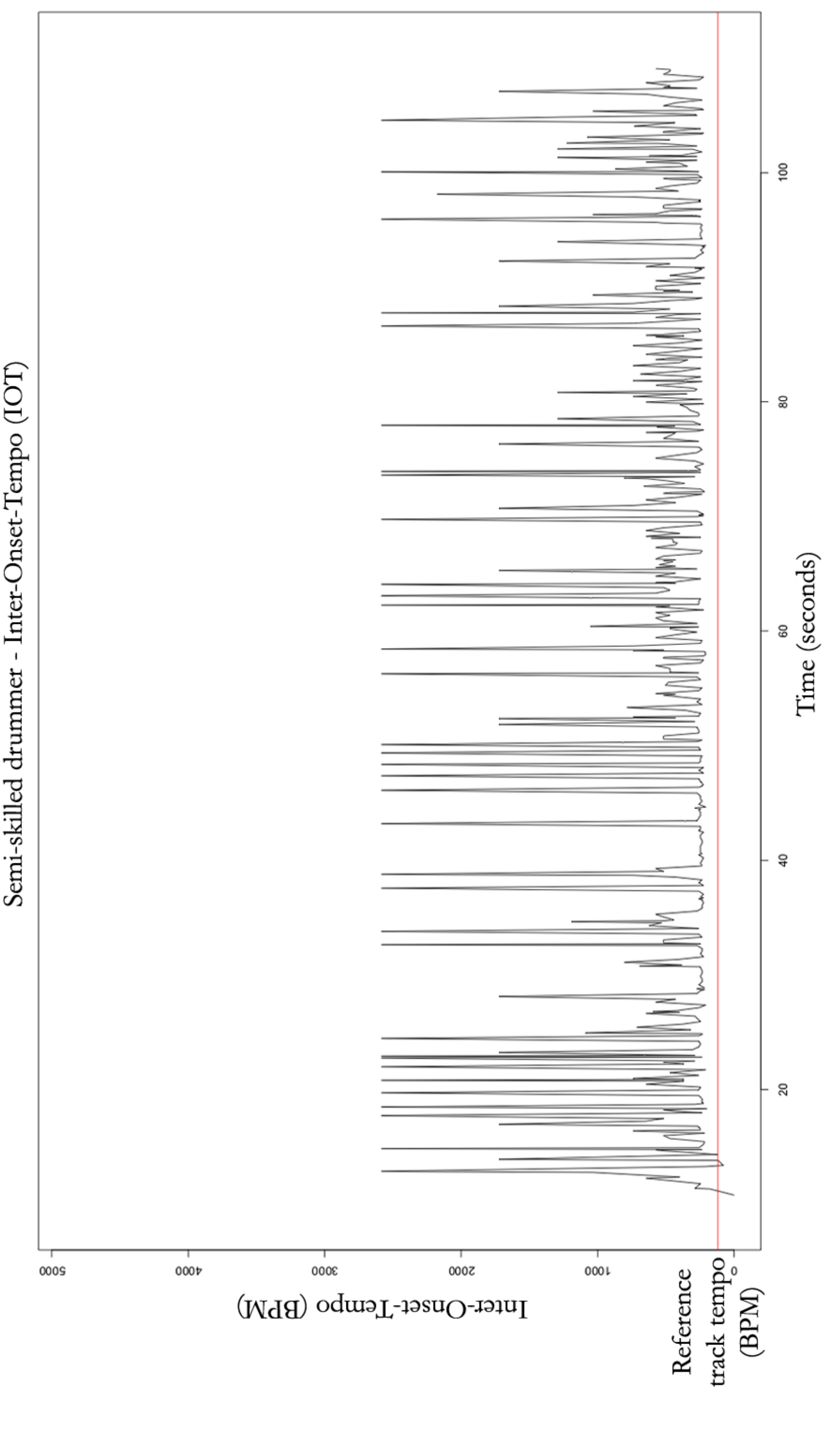


Figure 4.38: Inter-Onset-Tempo of the performance by the semi-skilled drummer. The red line indicates the tempo of the reference material (119BPM).

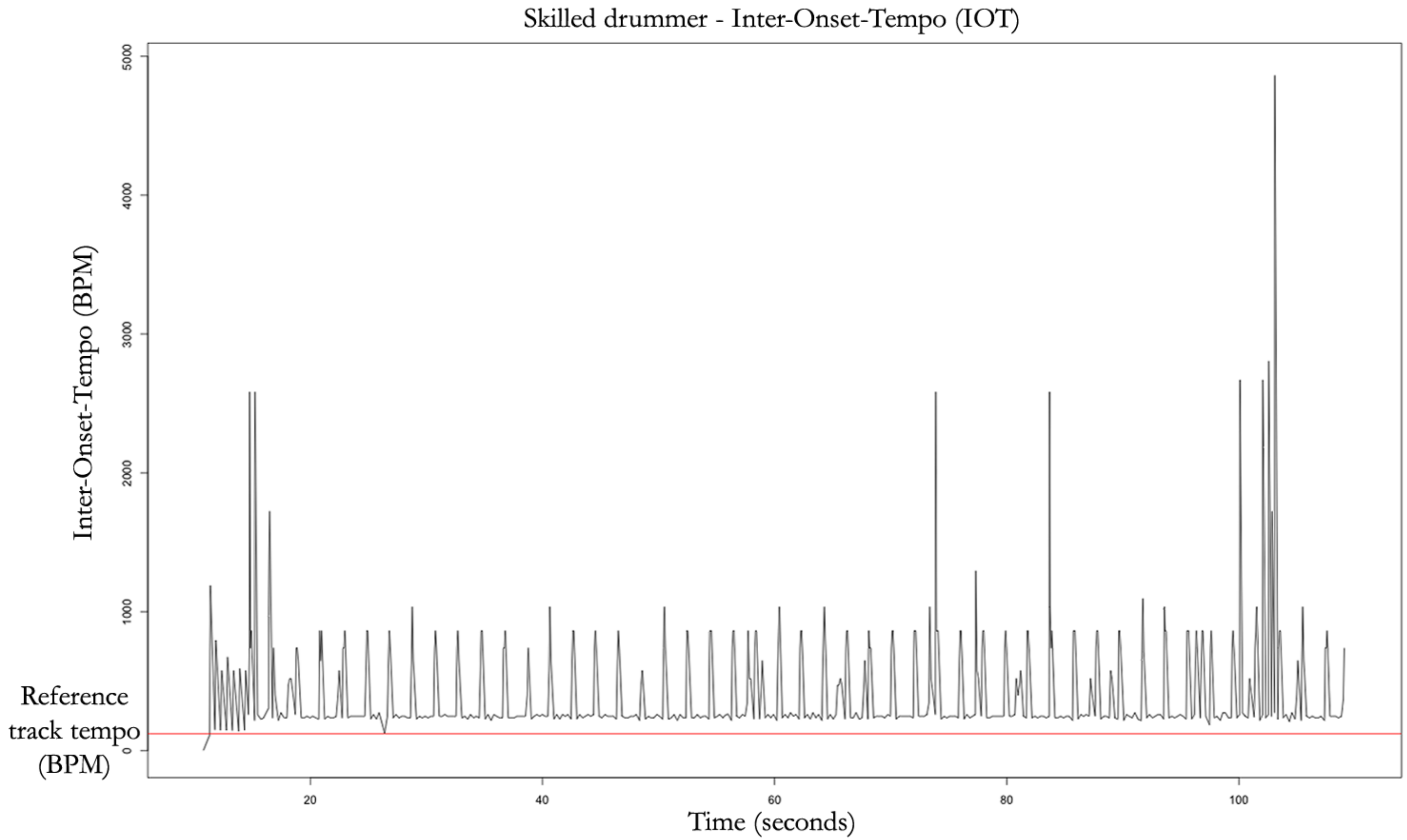


Figure 4.39: Inter-Onset-Tempo of the performance by the skilled drummer. The red line indicates the tempo of the reference material (119BPM).

4.6.4 Participant Video Data

Having identified the differences between the participants' timing fingerprints, it is important to understand the effect of bodily movement on temporal deviations and IOTs. Figure 4.40 shows the relative movement amounts between the three participants. Unsurprisingly, the amount of movement correlates to the IOT amounts, where increased movement corresponds to increased occurrences of higher IOT events. An example of this can be seen Figure 4.40 (a), in the spikes in movement and the increased activity in IOT at 40 seconds and 80 seconds for the unskilled drummer. These can be attributed to the participant moving to strike a different drum in order to create a fill, and subsequently attempting another fill, but being indecisive about the planning of the strike, and consequently waving his sticks in the air. These can be seen in the overhead video footage in Appendix B.6.

Interestingly, this correlation is less apparent for the semi-skilled drummer and, although there are points of high movement corresponding with increased higher IOT activity (at approximately 65-70 seconds, demonstrated by a change in strike location and pattern in the video in Appendix B.7), and relatively lower movement and lower IOT activity (approximately 40 seconds, Figure 4.40 (b) and another change in strike pattern in the video), there are several points with an inverse correlation. Two examples of this are at approximately 30 and 50 seconds, both of which correspond to the participants' changes in drum pattern. In the first instance, the old and new patterns are temporarily interlocked in order to preserve timing, although the merging of these two patterns causes extra movement. In the second instance, the pattern changes to a variation on the "Boogaloo" groove, which produces high IOTs due to the additional 16th-beat snare strikes.

With the most overall amount of movement, the skilled drummer's performance has two IOT spikes at approximately 75 and 85 seconds, corresponding to high levels of movement (Figure 4.40 (c) and Appendix B.8). This is particularly interesting, considering that there were noticeably less temporal variations in the timing fingerprint of the skilled drummer. These findings indicate that strike sequence density corresponds to movement, but not vice-versa. This is due to the combination of instruments in the sequence, and their location. Some percussive techniques can generate dense sequences of strikes with high IOT amounts that require minimal movement (e.g. a drum roll), while other less dense sequences require greater amounts of movement (e.g. a slow, short drum fill, using a variety of instruments).

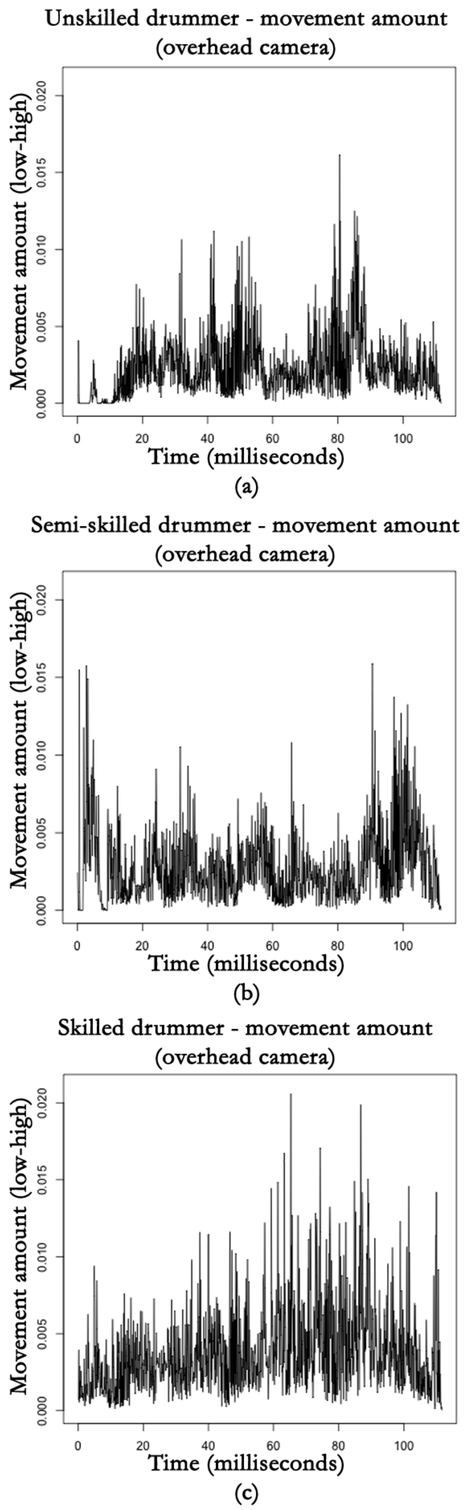


Figure 4.40: A visualisation of the performance movement for the (a) unskilled; (b) semi-skilled; and (c) skilled drummers.

Differences in movement amount can also be attributed to the skill level of the performer, and their technique. This is particularly relevant where the unskilled drummer may lack confidence in their playing and, owing to a lack of experience in performing actions associated with playing the drums, their body becomes rigid as a result of not being relaxed. In addition, an unskilled drummer may find decision making difficult because their drumming technique has not been developed enough to make quick improvisational decisions, or to make drumming more instinctive. Of course, this is fairly speculative, but this relates significantly to the development goals and drum rudiments that were discussed in section 3.2.1.

A further variable linked to the skill level of the drummer is the difficulty of the audio reference material. The reference material may have impacted the performance across the participants where it was too complex a piece for the unskilled and semi-skilled drummers, but still very easy for the skilled drummer. This particular variable was difficult to control, as assessing the skill of a performer relative to another is not easily quantifiable. Finally, it is worth noting that the video also included movement of the drums. Although this is the same for all participants (with the camera being in the same location for each participant), there may be differences given the choice of the instruments, where cymbals may move for longer post-strike periods, and given the average strength of the strikes across the participants, where the skilled drummer may cause greater residual movement in the struck drums.

Due to these reasons, the findings of the movement analysis are largely inconclusive. However, this analysis does extend the knowledge from the timing fingerprints for the model. Since sequences employing the use of different drums give rise to greater movement, which was found from higher IOT values being closely associated with minute fluctuations in the surrounding IOTs, any rules generated for inferring the effect of movement of percussive variation must take two things into consideration: multi- and cross-instrument sequence density. The next section will discuss the accelerometer data to further define the effect of instrumental location on the spatial coordination of bimanual movement for increased rule specificity.

4.6.5 Participant Accelerometer Data

Figure 4.40 shows the accelerometer data from the participants in time-series box scatterplots. One initial finding of this data lies in the similar increases in the levels of movement between the unskilled, semi-skilled, and skilled drummers. Despite this, direct comparisons between the overall spatial movement of each of the participants' hands is impossible due to the differences in strike sequence and hand/drum use. Therefore, specific differences in the spatial location between the hands of the participants must be contextually examined.

For the unskilled drummer, the left and right hands show higher average periods of elevation, characterised by a higher average tilt value. This can be seen in Appendix B.6, where the participant unconventionally uses rim shots primarily on the snare drum and medium tom. This participant's use of rim shots limits the variations in recorded spatial locations, particularly for the right hand. Another key feature of the unskilled drummer's performance lies in the stick technique and the management of the stick control, where the stick is held in the palm halfway towards the tip, horizontally and facing the floor (although elevated and pronated due to the drum position).

This can be seen by the consistently higher tilt and roll values. The data for the unskilled drummer's left hand indicates a tendency for supination, where the roll values indicate inward tilt. At approximately 53 seconds in the video performance, the unskilled drummer adjusts the position of the left hand on the hi-hat from a horizontal position to more vertically pronated position. This is reflected in the data where the variations in roll become more scattered after a period of constraint, at about halfway through the performance.

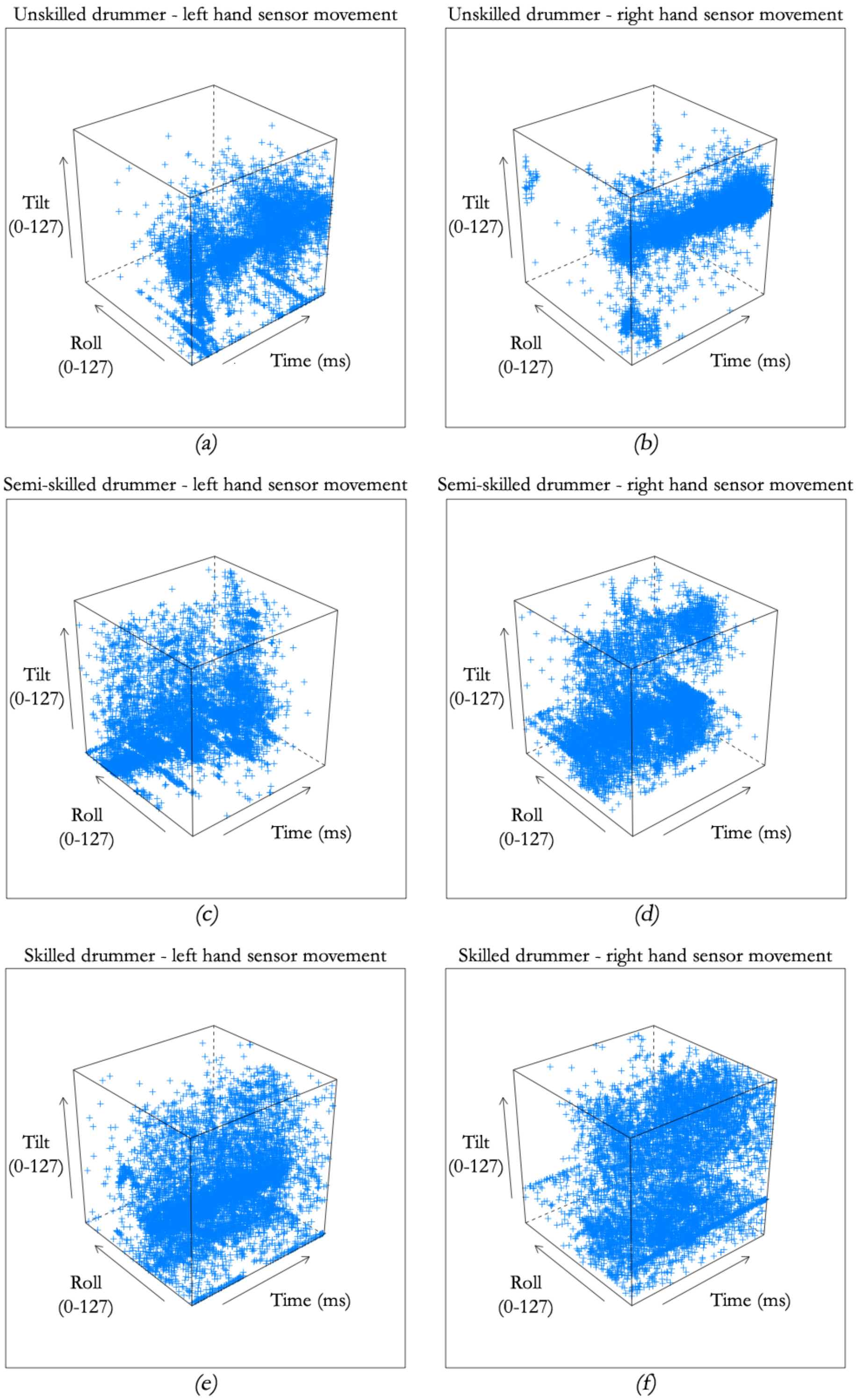


Figure 4.41: Tilt and roll accelerometer data for left hand (a, c, e) and right hand (b, d, f) of each participant's performance of the reference material.

4.7 Creation of the Performance Model

This part of the chapter is concerned with presenting an analysis of the performance data, followed by a section describing the computational implementation, drawn from the data analysis, and within the confines of the performance data extracted. The computational implementation will formalise the performance rules that have been developed directly from observation of the participants, and from analysis of the video, audio and accelerometer data. The rules here also define how the two methodological approaches, described in the previous chapter, will be amalgamated by contextualising the sequences of events, and applying different functions to the data flow. The final section in this chapter describes how some of the aspects within the compositional framework will be incorporated into the construction user interface of the compositional software tool. This section concludes with an overview of minimum system requirements for the compositional software tool.

4.7.1 Computational Design

So far this investigation has covered a range of different aspects of human percussive performance, ranging from instrumental behaviour, to variables in physical performance in Chapter Three. This discussion led to the theoretical framework in Chapter Four, which presented key areas of performance context containing three drummers of different skill level, physical constraint, and performance weighting functions that conjugate to represent the performance field. In the previous section, discussion focussed on the analysis of results collected from performance data of the three drummers, with a view to describing the performance context, and to assist in the development of computational performance rules for the performance reconstruction. This section will begin by defining performance rules and describing their computational implementation.

4.7.1.1 Performance Rules

One of the most significant findings of the performance analysis was the relation between timing variations and movement, in which a skilled drummer with more movement generated significantly less timing variations relative to drummers of different skill levels. This finding has implications on the performance model, as timing variation cannot be easily implemented based upon abstracted rules. Since the collected timing fingerprints are inherently weighted (particularly with the skilled drummer), the implementation of this in the model has two characteristics. Firstly, rules will not apply to timing variations. Instead, the rules will apply only to the generation of timbral variation. The second, and most important, characteristic of this model is a direct result of the first characteristic. That is, there is no direct relationship between the timing and timbral variations. This raises interesting philosophical discussions particularly concerning the perception of human performance. Despite this, it is expected that there will be little impact on the perception of the model, particularly as asynchronies less than 20ms are difficult to perceive (and a large proportion of the strikes for each fingerprint are within 20ms).[50]

The performance rules have been developed based upon previous discussion and focus on the physical aspects of human performance, and the physical nature of the interaction with the instrument. As a result, these rules are divided into three categories: the physical constraints of the performer; the physical constraints of the instrumental configuration; and the context of simultaneous movement, which will be described in the following three sections.

4.7.1.2 Physical Constraints of the Performer

At any given time the hands can play only two instruments with a stick. Since the hi-hat and bass drum are played using the feet, four of the nine instruments can be played simultaneously, two of which are the bass drum and hi-hat. Therefore, the first function that simulates the performer must prohibit the manual selection of simultaneous strikes to simulate this physical constraint (the physical contraint or PC function). This function is called *MultiSwitch*, the Java code for which is shown in Appendix B.9.

4.7.1.3 Constraints from Instrumental Configuration

Simultaneous bimanual cooperative and disjointed movements produce different biomechanical effects on the body. These biomechanical effects are more apparent in compound disjointed movements, in which playing the drums requires the simultaneous striking of instruments in a disjointed manner, and at different heights.[51] In order to abstract the level of biomechanical effects, the drums will be divided into three classes based upon relative height from the floor. This has three benefits. Firstly, these instruments are usually positioned in the same height, irrespective of location in the drum set

[50] In addition, the collection of the sample data was done independently of the participant performances, rendering the model an approximation of the performance.

[51] One exception to this is the bass drum, which is solely foot operated. Since there is only one excitation location on a bass drum (the beater contact area), it is difficult to computationally abstract constraints from the use of the bass drum. As a result, the bass drum will be excluded from this rule.

configuration and their relative position from the drummer. Secondly, this approach does not infer a specific drum-kit configuration, so can conceptually represent different drum set configurations. Finally, this approach allows for the assignment of values representing different combinations of simultaneous drum strikes, from cooperative to disjointed movements. Consequently, the three drum set classes are as follows:

1. High Drums
   - Crash cymbal
   - Ride cymbal
2. Medium drums
   - High tom
   - Medium tom
   - Low tom
3. Low drums
   - Floor tom
   - Snare drum

4.7.1.4 The Context of Simultaneous Movement

Manual selection of simultaneous drums can include drums from different classes, so it important to define the relationships between these classes and implement this relationship as a PC function. Since cooperative movement occurs when both hands function similarly, simultaneous strikes at similar heights will be considered to have a *low* level of movement. Where drums from opposing heights (low and high) are selected, the disjointed action will be considered to have a *high* level of movement. Since an instrument in any class can be selected first, from within the three classes, there are nine potential combinations of movement value that can be manually selected. The levels of movements and their respective combinations are shown in Table 4.3.

Where instruments from only one class is selected, the movement level will be considered to be *low*. The "rounding down" of the movement levels (e.g. *High-Medium* movement becomes *Medium* not *High*) has been done for two reasons. Firstly, there are biomechanical considerations of arm trajectory and motion planning where, as described in Chapter Three, there is a tendency to use the most economical movement

TRUTH TABLE FOR INFERRED MOVEMENT			
Drum Set	**High**	**Medium**	**Low**
High	Low	Medium	High
Medium	Medium	Low	Medium
Low	High	Medium	Low

Table 4.3: A truth table for defing simultaneous movement.

(Abend et al., 1982). Secondly, although there are combinations of instruments that produce high levels of movement under certain drum set configurations, these are relatively few, and addressing this by "rounding up" the movement level to *High* would affect the majority of combinations that would fall into the *Medium* level. The Java code implementation for the PC function, *Fuzzy*, is shown in Appendix B.10.

4.7.1.5 The Context of Sequential Movement

In Chapter Three, the sequences of strikes and the implications on the interaction of the drum were discussed in relation to trajectory formation and bodily control. It was noted that sequences of movements spanning multiple planes of motion and axes of rotation were more likely to affect performance variation more specifically, the timbre. Assigning a movement value (*Low, Medium, High*) to the current instrumental selection, as defined in the previous section, does not take into consideration sequences of movements, and the transition between the previous instrumental selection and the associated class, or the current instrumental selection and the associated class (e.g. low movement to high movement). Depending on the current temporal location in a performance, each instrumental selection will eventually become the current instrumental selection and, since the movement of the current class is only directly associated with the movement of the previous class (e.g. a strike in a sequence is only affected by the previous strike location, and not the 10th previous strike), only a *first-order* dependency is necessary.

As the previous instrumental selection and associated class has been defined according to the truth table in Table 4.3, the most appropriate method for smoothing the previous movement value with the current value is through the use of recursion; that is, to apply the same logic to the values in the current and first-order classes. This is described in Figure 4.41. In this example, the movement level of the previous selection was *Medium*,

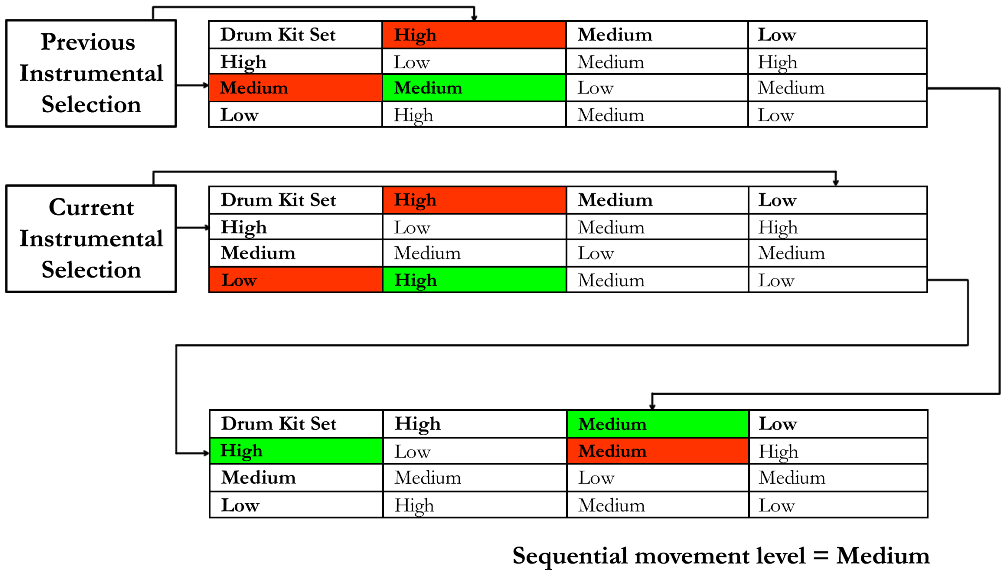


Figure 4.42: The truth tables for defining sequential movement levels.

and the movement level of the current selection was *High*. Recursion of these movement values produces a sequential movement level of *Medium*, which is transitively applied to the current selection. The rationale for this originates in the height relationship in the initial drum classes of both selections. In the previous and current selection, a drum in the *High* class, was selected. The movement across the planes of motion and the axes of rotation of these instruments over the sequence is relatively low and, based upon the discussion on trajectory formation and bimanual coordination from Chapter Three, as well as the results from the analysis of the performance data in the previous sections, it is most likely that any struck combinations of these instruments will employ the use of the same hand. In contrast, the other classes in the selection change from *Medium* to *Low* (e.g. a tom-tom to a snare drum). Although the movement level in the current class is *High*, the overall levels of movement across the sequence, when both hands are taken into consideration, is *Medium*.

Since music exists linearly over time, it is necessary to specify how the logic will transition at different temporal points in the composition. This can be done by considering that the next period in the sequence will become the current selection, and the current selection will become the previous selection, and so on. In the event that a selection is changed, the simultaneous movement value is updated and automatically propagated for recursion. This method of auto-propagating recursion differs from existing percussive control algorithms because this method allows the system to dynamically reconfigure according to the real-time manual selection. In most conventional percussive control algorithms, the configuration of the system does not alter based upon manual selection. The output of this PC function must then be passed to a PW function in order to transform the abstracted movement value into a weighted output. The Java code for the recursive logic for sequential movement is shown in the PC function, *FuzzyAggregated*, in Appendix B.11.

4.7.1.6 Performance Weightings

Having defined the PC functions that represent physical constraints in human percussive performance, the next stage is to define the PW function. This entails determining the methods for weighting the input value from the PC function and, since the primary concern of the performance reconstruction relates to creating timbral variation, the movement level values generated by the PC function will be directly related to sample selection. This section describes the construction and implementation of the PW function operator and the transformation of values from the PC function to weightings for sample selection and timbral output.

In section 4.6.2, a data reduction method reclassified the captured drum samples according to demarcated strike locations, adapted from Fletcher and Rossing's (2,2) mode of vibration (shown in Figure 4.26). Data reduction was undertaken in view of the impact of secondary timbral variations, and the timbral relevance of a given sample in a specific performance and compositional context. The chosen data reduction method restricts the timbral variation by grouping samples of similar strike locations together, which in the context of the vibrational characteristics of membranes also restricts the timbral variation across the zone. The data reduction method also allows for the inference of performance variation and performance accuracy through zone selection.

In the previous sections a method was defined for determining the level of movement based upon instrumental selection and context, with the movement levels *High*, *Medium*, and *Low*, the outputs from the PC function. This establishes the input argument for the PW function taken from the output of the PC function. The selection of zones via probabilistic methods is one potential method for zone selection. By applying different weightings for zone selection for each of the three movement levels, timbral variation becomes linked to performance context. It is therefore necessary to specify the relationship between strike location and movement level, in order to determine the probability difference between the movement levels.

Low levels of movement are those that are least affected by biomechanical effects caused by the movement itself, resulting in greater strike accuracy. For membranophones, this will result in the greater probability of striking a drum in the centre of the membrane (zone 5), as this is typically the primary target location (Sweeney, 2004a, p. 13). This does not, however, preclude other zones from being struck. The additional length of the drummer's reach increased, by the drumstick, does not proportionately disadvantage distal striking areas of the membrane (zones 1 and 4). Rather, proximal zones (2 and 3) are disadvantaged due to the oblique angle of the drum (the snare and tom-toms are usually at oblique angles) and the acute angle of the lower rim of these oblique drums. However, the most important consideration in the weightings between the two sets of proximal and distal zone is the posture of the drummer. The importance of posture and stability control in movements involving axial and sagittal movements of the upper body was described in section 3.2.4. Having the correct posture creates less biomechanical bias to striking the more proximal zones.

The probability of striking different zones of an idiophone requires a different approach than for membranophones, largely owing to the assignment of zone 5 to the "bell" area of the cymbal. Since the striking of the bell area is limited to specific performance contexts (demonstrated by the assignment of its own notational symbol), for the purposes of clarity, the striking of the bell will be considered an error in performance. The majority of cymbals (particularly the ride and crash cymbals) are positioned at a greater height, in distal locations, and tend to be positioned at an oblique angle. Consequently, there is a biomechanical bias towards striking these cymbals at the proximal zones (2 and 3) rather than the distal zones (1 and 4). Despite its position to the performer relative to the other cymbals, different biomechanical effects in playing the hi-hat produce similarities in the strike bias with the ride and crash cymbals. This bias is a direct result of striking the cymbal with the opposing hand (crossing the arms), which reduces the reach of the drummer, and creates a bias towards proximal zones 2 and 3. This biomechanical constraint has a similar effect on distal zones 1 and 4, with their selection being reduced. Consequently, the idiophone weightings are inverted from the zones on the membranophones. Six general rules for the probability of striking a zone on both membranophones and idiophones are:

Membranophones:

1. Zone 5, will be the most likely struck zone
2. Zones 1 and 4 will be the second most equally struck zones
3. Zones 2 and 3 will be the least struck zones

Idiophones:

1. Zones 2 and 3 will be the most likely struck zones
2. Zones 1 and 4, will be the second equally struck zones
3. Zone 5 will be the least struck zone

With these six rules defined, the next stage is to describe the effect of different levels of movement on these rules. For higher levels of movement, particularly where the drummer is playing at the limits of their ability, the overall stability of the performance reduces. Conceptually, a decrease in performance stability can be simply expressed in probabilistic weighting as a tendency for more equal probability of selection across the zones. Since the input of the PW function contains varying degrees of movement, these levels of movement can be assigned individual weightings. Furthermore, the differences in weightings for each zone across the skill levels can be scaled for weighting reduction and promotion depending on the implied movement level.

A set of weightings, which aim to reflect the factors described in this section regarding the movement and positioning of the various drums, will be discussed here. The probability weightings for each strike zone can be seen in Tables 4.4 (membranophones) and 4.5 (idiophones). Table 4.6 shows the overall percentage change in probability weightings between the unskilled and skilled drummer. These tables are represented graphically in Figures 4.42 (membranophones) and 4.43 (idiophones). Although the zones generally display inverse distributions of weighting between the two instrument types, there are subtle differences in the movement between the weightings, which can be characterised by the differences in the amount of scaling between the zones for each movement level. The most significant percentage change in probability between the two instrument types is in zone 5 where, as noted previously, this zone differs contextually between idiophones and membranophones.

Membranophones	Low	Medium	High
Zone 1	25%	23%	21%
Zone 2	5%	11%	17%
Zone 3	5%	11%	17%
Zone 4	25%	23%	21%
Zone 5	40%	32%	24%
Total	100%	100%	100%

Table 4.4: The PW function's membranophone zone weightings by movement level.

Idiophones	Low	Medium	High
Zone 1	13%	16%	19%
Zone 2	34%	29%	24%
Zone 3	34%	29%	24%
Zone 4	13%	16%	19%
Zone 5	6%	10%	14%
Total	100%	100%	100%

Table 4.5: The PW function's Idiophone zone weightings by movement level.

Zone	Membranophones	Idiophones
Zone 1	−4%	+6%
Zone 2	+12%	−10%
Zone 3	+12%	−10%
Zone 4	−4%	+6%
Zone 5	−16%	+8%
Total	100%	100%

Table 4.6: Percentage change in probability between skilled and unskilled weightings.

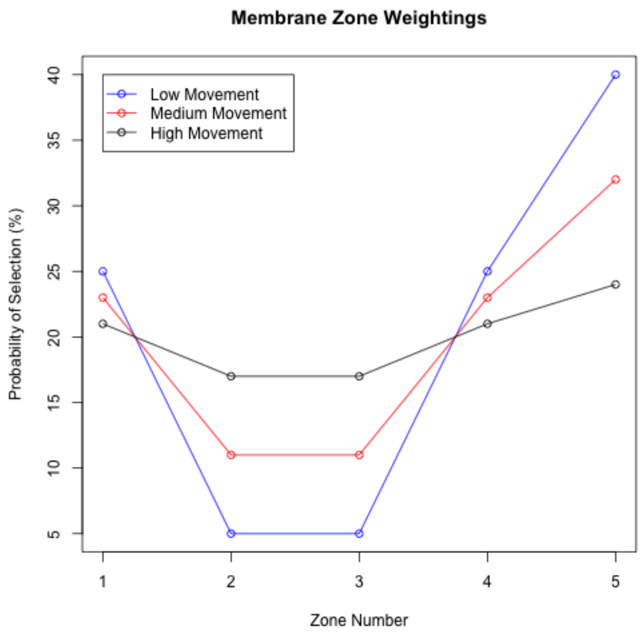


Figure 4.43: Membranophone mode weightings for different movement levels.

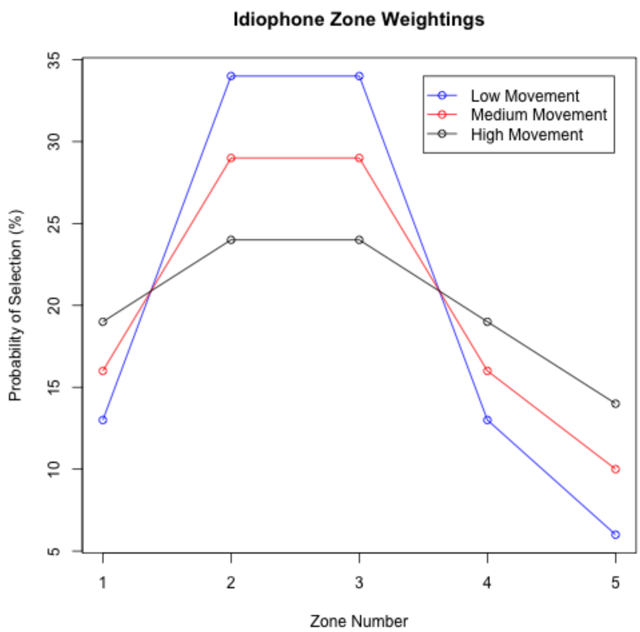


Figure 4.44: Idiophone mode weightings for different movement levels.

One way to ensure that the zone selection corresponds to the weightings is to represent each zone number in an array, n number of times (depending on the percentage value). Then, using a random number generator, performing a value lookup on the array position. The random number generator in Java code is as follows:

randomNumber = (int) (Math.*random*() * (max − min + 1)) + min;

Once the zone number has been returned from the array lookup, a sample from that particular sample bank can be selected and, subject to the local and global parameters, played back. The advantage of this method over existing systems is that repetition of identical samples becomes less likely and, with the probabilities based on performance context, the timbral variation inherent in the sample bank will be representative of the performance context.

4.7.1.7 Performance Data

In the previous sections, the LGP functions were defined as a function whose effects include the transformation of an instance or sequence of a performance reconstruction on a local or global level. Examples of this include temporal phenomenon, such as tempo, and temporal effects relating to metrical beat level conditions. Therefore, in order to define these functions, it is necessary to firstly describe the implementation of the timing fingerprints as performance data.

4.7.1.8 Timing Fingerprints

The values of the onsets in the captured audio performances were derived from the onset position relative to the nearest beat. This was determined by calculating the mid-point between the current and previous beats, and applying this value as the cut-off point, or magnetic area, for the beat assignment (see Figure 4.35). The onset deviation values were then given as negative (early) and positive (late) values. These were shown in relation to IOT in Figure 4.36. Due to the distribution of values and likelihood of onsets occurring at specific times, depending on the skill level chosen, the system will determine the onset times for each synthesised note according to the timing fingerprints and the calculated onset strike asynchrony.

In order to implement the timing fingerprints into the performance model, the complete set of onset values (positive and negative) for each performance was loaded into an array. Since the timing fingerprint is inherently weighted towards the participants' skill level, a similar approach to onset value selection can be taken, to that of the zone selection algorithm in the physical operator: a random number is generated and a value lookup is done on the randomly generated array position: a zeroth-order Markov model. This can be seen in the methods *Mode1DevLookup()*, *Mode2DevLookup()* and *Mode3DevLookup()* in *PerformanceWeightings.mxj* in Appendix B.12, where the "Mode" numbers in the methods' naming convention are a reference to the three participants.

Once an onset value has been obtained from the array, it is then passed as a variable to a method called *Timing()* (also found in *PerformanceWeightings.mxj*). Since the physical operator has already identified the correct sample bank in which to select a sample, the next stage is to control the flow of information in order to simulate temporal deviations. The control of information parsing was chosen as an alternative to using silence at the beginning of a sample, as previously done by Hellmer (2005), in order to reduce computational overhead caused by the extra buffering of the silence as audio. However, in order to control the flow of information, there are two considerations. The first consideration is the simulation of negative (early) onset deviations. A negative onset deviation theoretically occurs before the onset of a beat. Therefore, latency must be built into the system, with the reduction in latency relative to the negative onset deviation. For example, in the case of a 20ms negative onset deviation, a standard system latency of 100ms would be reduced to 80ms. In the case of positive (late) onset deviations, a 20ms value would need to be added to the 100ms system latency to produce 120ms latency. One of the problems in using system latency is the appropriateness of the size of the latency, with time-critical applications operating at different temporal levels. For example, an appropriate latency at 60BPM would be 1,000ms. At 120BPM (or double time), such latency would make an on-time onset appear late because the latency is larger than the beat interval.

The second consideration is independently controlling the latency of multiple onset deviation values. A computational approach that addresses this issue is multi-threading, where different threads of code run simultaneously and independently. In addition, using the *thread.sleep()* function, each thread can be simultaneously suspended for a specific time in milliseconds. Furthermore, the reporting of a thread's status can allow for time-critical events to be managed. One drawback to this method relates directly to the size of the latency, whereby the longer the thread sleeps, the more threads are required. Conversely, in the case of shorter thread sleeps, fewer threads are needed. This can be described in terms of concurrent sample selection, whereby the PC function defines a maximum of four instruments playable simultaneously. Where the latency level is higher than the beat interval to the following beat, more threads will be needed to accommodate the new selection while the previous selection is woken up and released.

There are two options that can overcome this. Firstly, a large number of threads can be used to accommodate all potential *thread.sleep()* values to beat interval overlaps. Secondly, the latency can be reduced to minimise the sleep time of each thread. This will subsequently reduce the beat overlap of the thread suspension. One critical difference between these two options relates to their suitability towards changes in beat interval or metrical level or, by extension, tempo. Since tempo and metrical level are important compositional parameters, this is a significant consideration in the addressing this issue. As a result, the system latency will be set to 100ms, with a maximum thread sleep time of 200ms, and minimum sleep time of 0ms. The implementation of the multi-threading can be seen in the method *Timing()* in *PerformanceWeightings.mxj*, in Appendix B.12. This method uses four threads.

4.7.1.9 Integrating Global Parameters with Micro-Temporality

Having established that timing fingerprints determine the temporal flow of the control data, attention must now turn to the implications of the local and global parameters. By definition, a parameter transforms the performance reconstruction on a global level. Since the PC and PW functions have defined timbral control, the local and global parameters must be concerned with the temporal transformations. Therefore, we can define these parameters as being concerned with transforming the asynchrony values from the tempo in the reference audio, to that of the newly specified tempo. However, since the flow of data in the fingerprint and the performance model operates on a sub-second temporal level, the transformation of the timing fingerprints not only occurs in a global context (e.g. tempo), but also at a local level (e.g. metrical levels), characterised by differences in localised variation.

There are, however, difficulties in implementing temporal transformations using global contexts such as tempo, particularly in the case of human performance variation. Previous research into timing and tempo have suggested links between structure and tempo deviations (Desain and Honing, 1993), and deviations in swing ratio being linearly linked with tempo (Friberg and Sundstrom, 2002). Since the focus of this investigation is on the creation of a compositional tool, there is no a priori structure in which transformations can be applied. The number of structural conditions that would need to be accounted for would entail significant analysis of multiple percussion pieces, which is outside the scope of this investigation.

One solution to this problem was suggested by Mazzola and Zahorka (1994), in which hierarchies of performance "fields" represent a structure of musical time starting from localised performance that extends upwards to affect the global nature of performance. Mazzola and Zahorka assigned local tempi to performance fields, with each performance field represented as a "subframe" of another (Mazzola and Zahorka, 1994, p. 47). This approach gives performance fields their own tempo within the constraints of a master tempo. Conceptually, this applies to the independent tempo differences in musical performance brought about by the complexity of musical structure. An example of this is the assignment of local tempi to each hand in piano performance.

The linear link between deviations in swing ratio and tempo, as suggested by Friberg and Sundstrom (2002), offers another conceptual way that global tempo can be implied from a static tempo. The calculation of IOI between the reference beat at different tempos, and the subsequent scaling of each value in the timing fingerprint, is one method of simulating tempo change. However, this method is not ideal since the timing fingerprint is relative to one metrical level. In order to simulate half, quarter, or eighth notes which are commonly found in percussive rhythm, the tempo would need to be increased. This would result in extremely high tempo values that could potentially impact system performance, and many of the smaller asynchronies would be imperceptible at extreme tempi.

One solution is to use a combined approach, applying the concept of tempo hierarchies to linear performance variation, where hierarchies representing different note levels contain independent fingerprint values. One benefit of this approach is a decrease in note duration between whole and half notes, which represents a tempo change equivalent to 60BPM to 120BPM. Consequently, the global tempo can remain static, with the perceived tempo being changed simultaneously. Compositionally, this approach allows for increases in density, and sequences of accented strikes. In this way, the local and global parameters can be considered a parameter whose function is the transformation of the timing fingerprint data into different hierarchical tempo levels in response to manual input.

4.7.1.10 Note, IOT, and Temporal Variation Hierarchies

In order to create note, tempo, and variation hierarchies from existing data, it is necessary to firstly determine the correct metrical hierarchy. The creation of hierarchies from the division of, and relational pattern of, metrical units in western music is not uncommon. In most cases, whole notes constitute the highest level of the metrical hierarchy (Cooper and Meyer, 1963; Lerdahl and Jackendoff, 1983), and allow subdivisions to several lower levels (half, quarter, eighth, sixteenth, and thirty-second notes etc.).

The initial onset asynchrony values for the timing fingerprints were analysed in relation to quarter notes, owing to the pulse of the reference audio, and the large IOIs between whole and half notes, which, in turn, include quarter note asynchronies. Consequently, there were difficulties in detecting asynchronies relative to whole and half notes. Furthermore, if assumptions were made concerning the linearity of asynchrony, relative to the tempo inferred by whole and half note durations, it would impact the system latency. Since whole and half notes both occur at the quarter note level, quarter notes constitute the highest analytical level in the system, and whole and half notes will derive their timing fingerprint data (e.g. the magnetic area) in the same way as quarter notes. Conceptually, having quarter note durations as the root of the hierarchy allows the hierarchy to be extended to lower metrical levels relevant to micro-timing. Extreme examples of subdivision at different metrical levels include asynchronies as low as 32nd notes. Compositionally, extending the metrical hierarchy to such a low level also allows for faster tempos. For example, the relative tempo of a sixteenth note at 100BPM is equivalent to 1,600BPM, which, as noted earlier in Figures 4.37 to 4.39, is similar to the IOTs of the participants' drum strikes at lower metrical levels. Furthermore, low metrical levels also provide greater scope for the compositional exploration of faster drum sequences, micro-timing, and the meso-periodic level, particularly the interaction between the lower and higher metrical levels.

In order to extrapolate the lower hierarchical levels of timing asynchronies from the timing fingerprint data, a similar method was used to that described in Figure 4.35. Since the subdivision of each lower metrical level is 50% of the IOI duration of the higher metrical level, the magnetic area is reduced by 50%. For each subdivision, the magnetic area is based on the previous note mid-point, and the values within that magnetic area are used for the timing asynchrony. This is illustrated in Figure 4.44.

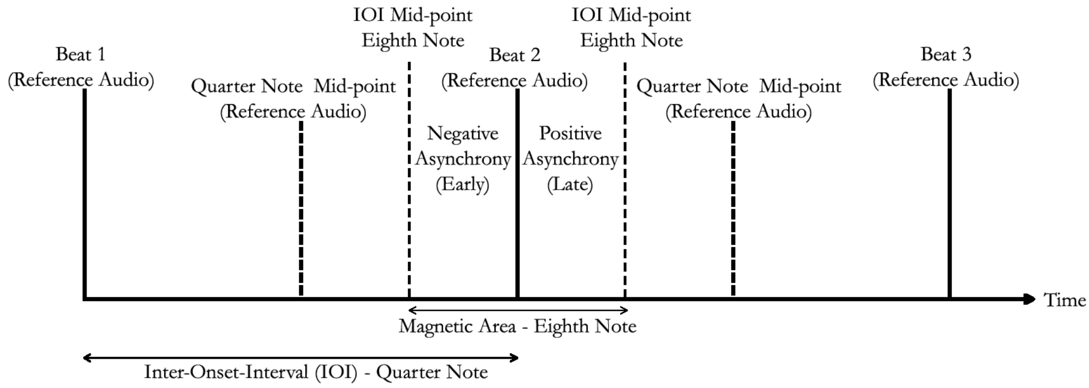


Figure 4.45: Membranophone mode weightings for different movement levels.

4.7.1.11 The Sample Bank

In order to ensure timely playback of the audio files, it is necessary to load the samples into RAM, and assign each sample to a MIDI velocity. As the audio data consists of approximately nine thousand samples covering nine different instruments, the implementation of the sample database is critical to minimising computational overhead. Since each instrument contains five demarcated zones, each zone can be assigned to an individual MIDI key number, which spans 41 keys. Since the bass drum has only one demarcated zone, there is only one MIDI key assigned to this instrument. An additional benefit of this approach is that the PD-103 can be used for live performances with a MIDI keyboard. Given the large dataset, compromises were necessary to ensure that playback was not affected by allocating too much memory into RAM. After several load tests, it was decided that 25 samples per MIDI key were sufficient, representing 100 samples per instrument. This allowed for significant timbral variation without affecting the performance of the software.

Each sample has four parameters: the designated sample number; a value for loudness; a value for spectral flatness; and a value for spectral centroid. Each of these parameters are stored in a separate array, and sorted in an ascending order based upon parametric value. The advantage of this approach is that a custom MIDI velocity curve can be created by converting a graphical representation of the MIDI velocity curve (constrained by the minimum and maximum parametric values), and performing an array lookup for the sample number corresponding to the parametric value. An example of this can be seen in the Velocity Curve Window, a screenshot of which is shown in Figure 4.45. The java code for the implementation of the velocity curves is shown in Appendix B.13, and is contained in the following files: *KickRMSarraymessage.java*, *RMSarraymessage.java*, *KickCentroidarraymessage.java*, *Centroidarraymessage.java*, *KickFlatnessarraymessage.java*, *Flatnessarraymessage.java*.

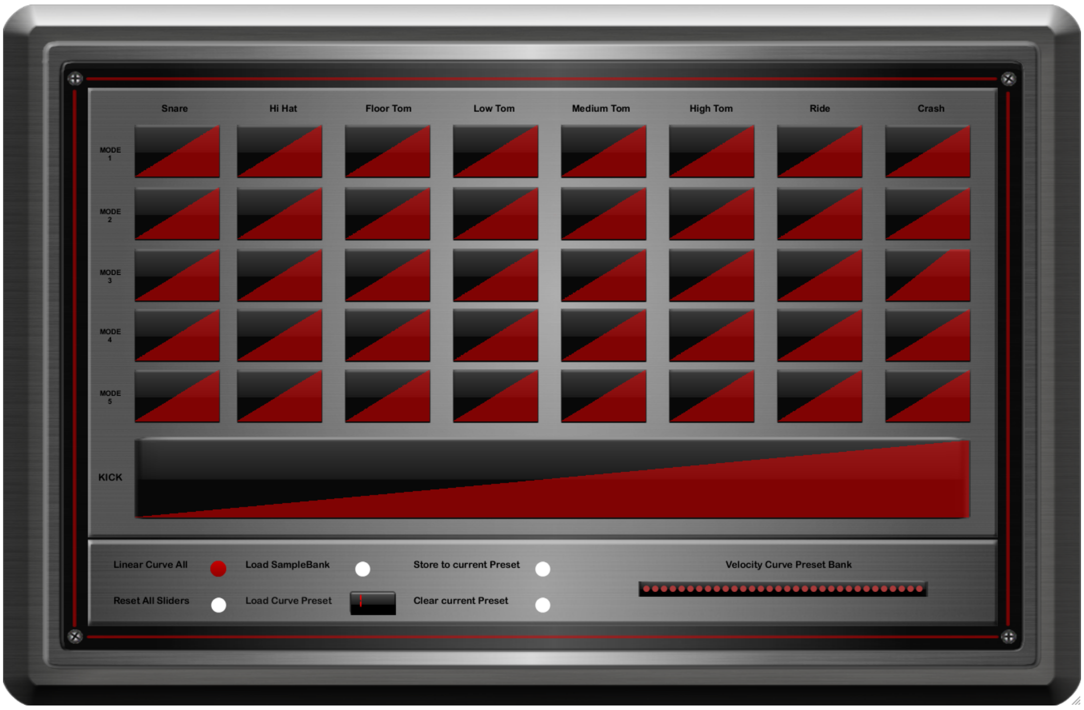


Figure 4.46: The Velocity Curve window of the PD-103.

4.8 Compositional Implementation

In section 3.3.3, the meso-period rhythms of Africa were discussed. From that discussion, parallels were drawn between the construction of the meso-period and the Western chromatic scale. In particular, Koetting (1970) identified a 12-note TUBS notation system as a suitable method for graphically representing African music, together with some key features of the meso-period. These key features (independency, flux, and broad bandwidth) were then expanded upon, and a compositional framework derived, for this thesis. Conceptually, these features are inherent in the performance model so do not require any specific computational implementation. Another part of the compositional framework is derived from the analysis of the sample banks themselves, and relates to the feature extraction methodology, which resulted in the timbral parameters of loudness, spectral flatness, and spectral centroid. This section will detail the functions of the software that relate to the compositional framework.

4.8.1 The 12-Step TUBS Representation

Since much of the compositional framework has been derived from the meso-periodic rhythms of Africa, it is important that the interface between the composer and computer program is conducive to its composition, including that of other genres. The link between the chromatic scale indicates that cycles of twelve notes are suitable for the generation of meso-periodic rhythms. From a Western perspective, cycles of sixteen (at a 4/4 time signature) are more common and, because twelve is divisible by four, it is still possible to create rhythms in 4/4. As a result, the interface will be based upon the 12-step TUBS model as defined by Koetting (1970). As mentioned earlier, three performance model representations are required in order to independently play all instruments. These are shown in the screenshot of the main PD-103 window in Figure 4.46.

4.8.2 Rhythmic Cycles

Referring to Figure 4.46, it is worth noting the red graphic sliders above the 12-step TUBS banks. These are called form banks, and refer to the cycling of the sequences for the TUBS bank below. Each sequence created by the TUBS system can be stored as a preset for live recall. The form banks allow the user to determine the number of cycles of a particular sequence in a preset, before moving to the next preset/sequence. Compositionally this is useful, as it allows for the graphic representation of cycle length for each bank, and allows the cycling for each bank to be independent. Furthermore, each form bank can be stored to a preset, allowing for the recall of macro-level compositional structures.

4.8.3 Compositional Timbre Parameters

It was noted that three representations were required in order to simulate independency within meso-periodic rhythm, as a result of the constraints in the performance model. In order to extend the compositional application of the model, in the context of the three timbral parameters discussed previously, it is necessary to create three representations for each parameter. This results in the creation of nine representations of the performance model, three representations each for loudness, spectral flatness, and spectral centroid. One of the benefits of this approach is that meso-periodic rhythms

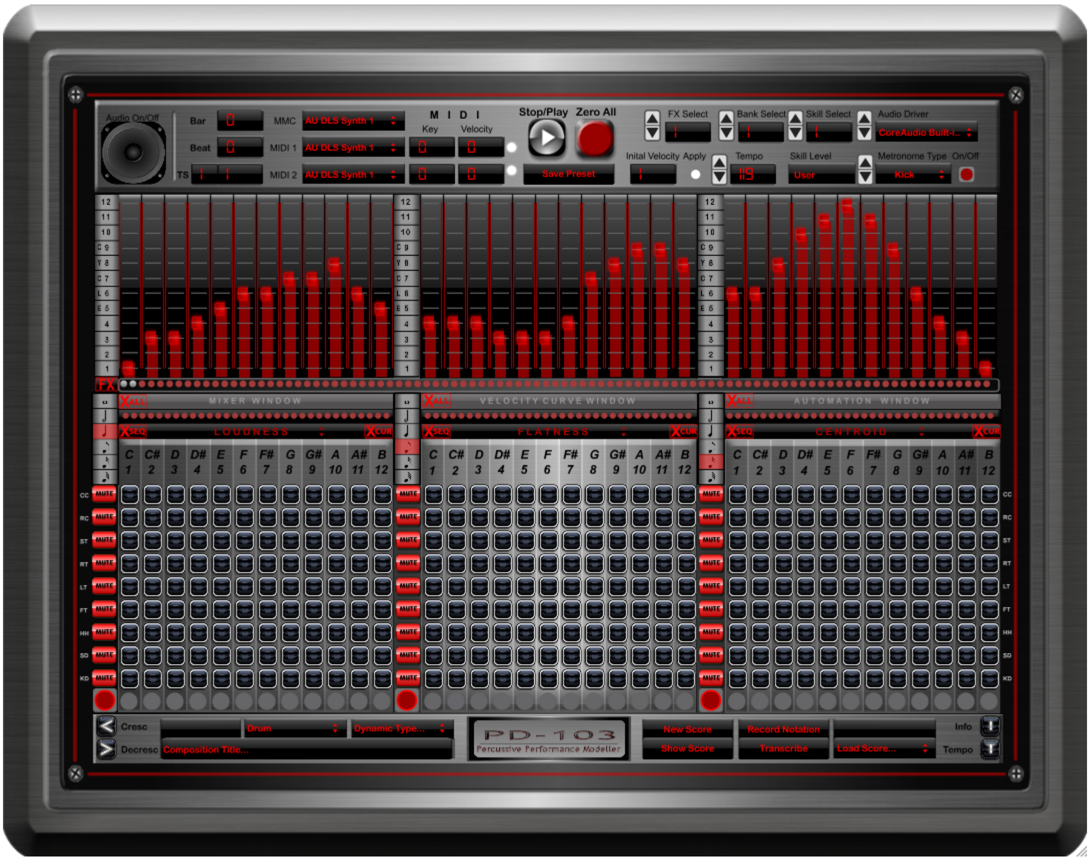


Figure 4.47: The main window of the PD-103, showing the three representations of the 12-step TUBS system, with each step labelled with an associated chromatic scale.

and high levels of independence can be achieved for non-conventional compositional parameters. Since the performance model seeks to simulate human performance, the use of flatness and spectral centroid also has the potential to generate interesting performance dynamics. Furthermore, with the ability to store sequences and form cycles as presets, it is possible for a composition to be split into three parts: one for each timbral parameter.

4.8.4 MIDI and Rewire Integration

The PD-103 was designed for composers and producers to create rhythms. The main method for operating the PD-103 is via the on-screen graphical user interface. However, modern home and commercial production and recording studios commonly use electronic instruments that employ different communications protocols. In order to make the PD-103 suitable for a variety of different musical applications, a design decision was made that would see the PD-103 capable of being controlled by a programmable MIDI controller, such as the Novation Remote 37SL, Novation Launchpad or similar. Such functionality would also enable the PD-103 to be used for live performance or improvisation. Other communications protocols such as Open Sound Control (OSC) are less limiting than MIDI, owing to their communications format. However, OSC is not as widely available as a communications protocol in many hardware controllers, whereas MIDI is still considered the universal standard. For this reason, it was decided that the PD-103 would be designed to include additional control via MIDI hardware devices. For example, the volume faders in the "Mixer" window are controllable using a MIDI controller. Although the software tool can be used with MIDI, it is not designed around the MIDI specification, which allows for future revisions of the software outside of the MIDI protocol.

A consequence of this decision was the limitation on velocity steps per MIDI note in during live performance or improvisation. Since the sample database was reclassified via demarcated zones, it was logical to assign each of the zones to an individual MIDI key. However, owing to computational overhead, it was not possible to preload all of the samples into memory. Consequently, a design decision was taken to limit the number of samples loaded to 25 samples per zone. The MIDI keyzone assignment (MIDI key numbers 0-41) is shown in Figure 4.48.

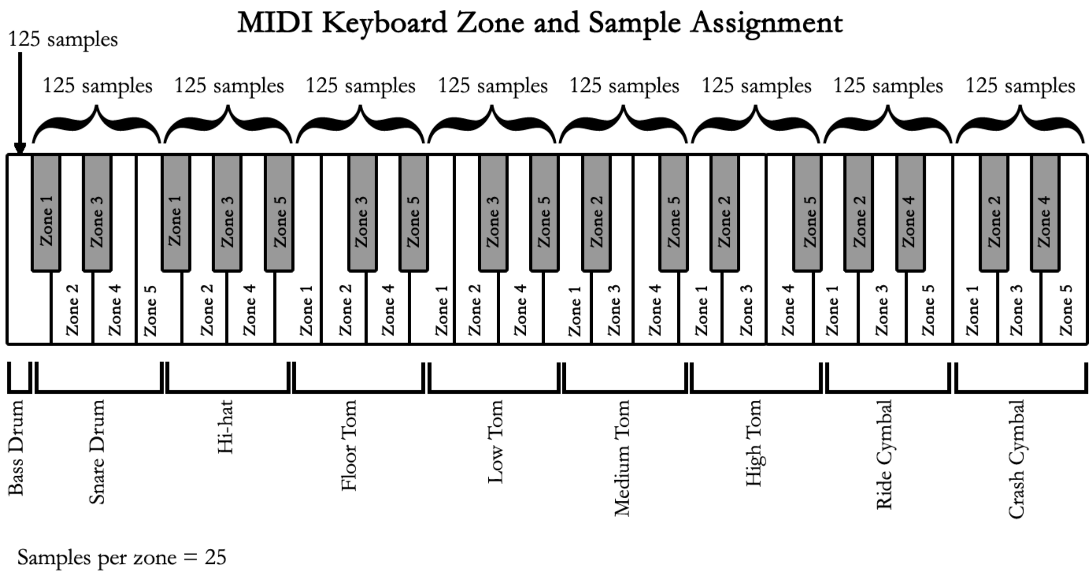


Figure 4.48: The PD-103 MIDI keyzone assignment.

One limitation of this implementation is that only one spectral parameter can be controlled at any given time. In order to change spectral parameter, the MIDI channel needs to be changed.[52] Operating the software via MIDI controller requires a MIDI channel change to trigger samples from the banks, where loudness, spectral flatness and spectral centroid are assigned to MIDI channels 1, 2 and 3 respectively. No such limitation exists when operating the software via the graphical user interface (see section 4.8.6 below for further information on bank polyphony). Any samples triggered via the MIDI keyboard are not subject to system generated timbral or timing asynchronies.

Because the performance model uses nine individual instruments, having the PD-103 output only a single stereo channel was seen as a limitation of the compositional workflow for mixing and mastering. Owing to the distribution and likelihood of timbral and timing variation selection in the software, replicating timbres in rhythmic sequences are unlikely. Using only a stereo output, requires multiple takes to output all nine instruments, and is not only time consuming, but may also hinder the creative process. This was overcome by extending the number of audio output channels to nine, and adding Rewire integration. The Rewire integration improves the compositional workflow of the PD-103 to allow it to record all nine instruments simultaneously in other compositional software such as Cubase, Logic or ProTools.

4.8.5 Presets

The PD-103 contains three main presets based upon performances of real drummers: unskilled, semi-skilled and skilled. However, there may be an instance where a composer may wish to have a "deadpan" performance (e.g. with no timing variations). For this reason, an additional preset entitled "User", bypasses all the system-determined onset times for the synthesised notes. However, the timbral selection algorithm is still operational, with the zone weightings being based upon the "skilled" preset.

4.8.6 Bank Polyphony

Bank polyphony is concerned with configuring the playback of the banks. There are three options for bank polyphony: "monophonic bank", "bank diversion", and "polyphonic bank". In monophonic bank mode, only the currently visible bank and parameter will be heard. Any sequence not visible is automatically muted. In bank diversion mode, a sequence that is not currently visible will be diverted to play according to the

[52] This being one of the limitations of MIDI compared with OSC.

currently visible parameter, irrespective of the original parameter of the sequence. In polyphonic bank mode, the parameters assigned to invisible sequences are preserved and played with those features. Polyphonic bank mode is the default mode. Operating the software via MIDI controller requires a MIDI channel change to trigger samples from the banks, where loudness, spectral flatness and spectral centroid are assigned to MIDI channels 1, 2 and 3 respectively.

4.8.7 Minimum System Requirements

Operating instructions are included in the disk image of the software, located in Appendix A.2. The PD-103 software was saved as a standalone Max/MSP application, which includes MAX Runtime 5.1.8. The application package includes the GUI components, the sample database and all relevant .jar files.

The minimum system requirements are:

- Mac with Intel processor
- Mac OSX version 10.6 or later
- 8GB RAM (16GB recommended)
- 7.5GB of Hard Disk Space
- 1280 x 1024 screen resolution (2880 x 1800 recommended)

Some additional features in the software may require third-party applications for full functionality. These, however, are not essential. Additional components requiring manual installation (see Appendix A.3) are:

- Maxscore (free). Maxscore is the notation interface for Max/MSP.
- JSML License (US $50 students, US $120 others). This license allows you to use Maxscore. Licenses can be obtained from: `http://www.algomusic.com/jmsl/purchase.html`.
- Finale Notepad (free). The notation component of the software uses the font "Maestro 24", and is available to download from: `http://www.finalemusic.com/products/finale-notepad/`.
- Rewire (free). The PD-103, when used in conjunction with Rewire, enables the user to route the audio output directly into a Rewire-compatible DAW, such as

Cubase. This also enables the user to record audio without the need for additional hardware. It is recommended that 16Gb RAM is used when the PD-103 is used in Rewire mode. Rewire is available from: http://www.propellerheads.se/download/updates_rewire/.

- "Digital 7" font. This font is used for the LCD display throughout the software. If this font is not installed, the font will revert to the default Max/MSP system font.

It is recommended that the sample bank be loaded in a sequential two-step process to ensure that all samples are loaded, and ensure the smooth operation of the software. Firstly, it is recommended that the user load the parametric values into memory. Secondly, it is recommended that the user load the samples into the memory. Failure to follow this procedure may result in the sample bank being incompletely loaded into the computer's memory. The recommended process is as follows:

1. Load the velocity curve preset in the Velocity Curve Window and wait for the velocity curve presets to be loaded. Once the load bar indicator in the Velocity Curve panel on the main screen has turned completely grey, move on to the next step.

2. Load the samples into memory and wait for the three load bar indicators on the parameter selection menus on the main screen to turn red before pressing play.

When using the PD-103, it recommended that the tempo remain between 100BPM and 120BPM. This is because the multi-threading for the timing variations through the *thread.sleep()* function in the java code, suspends the thread. At higher tempi, with multiple simultaneous events, it is more likely that all of the threads will be in sleep mode, resulting in the dropping of events or queues for events to be parsed through the thread. This will result in unpredictable behaviour, which can be remedied in future versions by adding more threads.

Chapter 5

Composition Portfolio and Analytical Notes

"Numbers don't make art. People do. It's all in the choice of the tools, the colors, the textures and the rendering"
– Gary Lee Nelson, 1990

5.1 Study No 1: African Meso-Periodic (I)

- **Length: 04:27**
- **Parameters: Loudness**
- **Skill Level: Skilled**
- **Number of TUBS Banks: One primary, two supporting**

Study No. 1 was written solely for percussion, using the complete range of instruments contained in the PD-103, and augmented with orchestral gongs. This study was composed using one primary bank, in conjunction with two supporting TUBS banks. It used "loudness" as the main feature-based parameter. The skill level of the drummer, from which the human performance data was initially drawn, was "skilled". Notably, the rhythmic construction of this study was based upon Walter Schloss' dichotomy between relative duration and the chromatic scale intervals, which was described in section 3.3.3.

As previously mentioned, the theory of Participatory Discrepancies, and its relationship to African music, was a particular inspiration for research into the area of percussive performance variation and composition. This inspiration is reflected in the creation of Study No. 1, which embraces African meso-periodic rhythms. Meso-periodic music is typically characterised by the repetition of rhythmic patterns, using a variety of different instruments. While it was theoretically possible to apply this type of repetition to the pattern described in Figure 3.14, which showed the graphical representation of the meso-period in relation to the chromatic scale, such an approach was aesthetically limited owing to the restricted number of instruments in the PD-103. Aesthetically, it was more pleasing to consider the instruments as separate streams with different aesthetic and structural functions, within a meso-periodic inspired contemporary perspective. Thus, the motivic rhythmic pattern was composed with a focus on four main instruments: ride cymbal; low tom; medium tom; and high tom. This is shown in TUBS notation in Figure 5.1.

Following typical African musical convention, the ride cymbal in this piece was assigned the role of the fundamental bell pattern, owing to the similarities with the ride cymbal's "bell". Chromatically, the ride cymbal represents a C major scale. The high tom pattern in this piece was based upon the standard pentatonic scale, with the exception of the extra note at the beginning of the pattern (C/1), which was used to

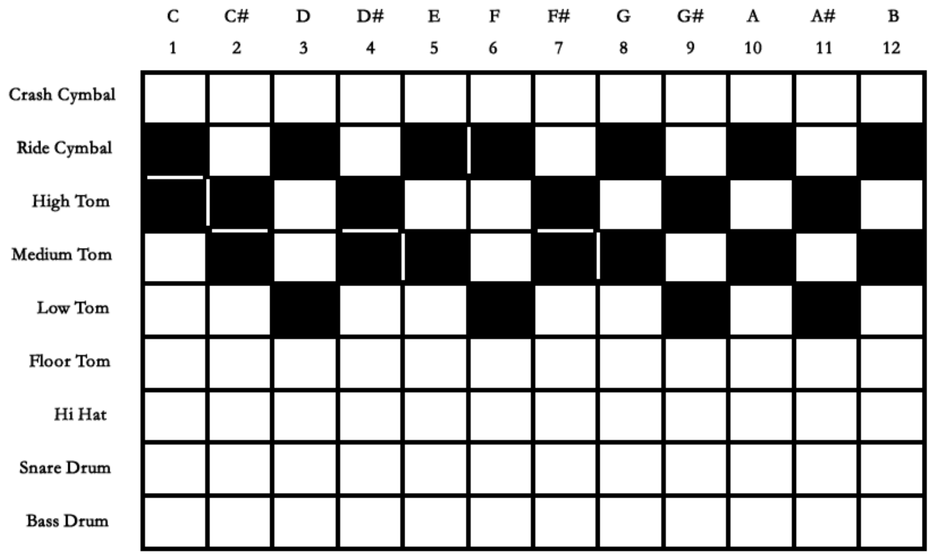


Figure 5.1: The TUBS notation for the fundamental rhythm in Study No. 1. Adapted from Schloss (1985).

accentuate the start point of the cycle. The medium tom was represented by a C# locrian 2 scale. This was chosen for two reasons. Firstly, the software was created using specific performance rules, one of which was derived from the physical constraints of the performer whereby, at any given time, the performer can only strike two instruments with the hands simultaneously. This prevented the phasing of scales with the same tonic. Secondly, by dividing the twelve periods in half, the C# locrian scale creates tension within the cycle by complementing the ride cymbal (appearing between the ride cymbal) for most of the first six periods, followed by a complementary inversion towards the high tom.

As a compositional tool, the software presented some very interesting compositional challenges. As mentioned above, the software was created using specific performance rules, one of which was derived from the actual physical constraints of the performer. As a result, the density referent, which was limited to the simultaneous selection of only two instruments, made applying additional phased modes to the pattern impossible. Thus, the low tom was based upon the C natural minor scale, with several scale notes omitted (C, D#, and G). By using this constrained partial scale on the rhythm, the rhythm density increased on the ride cymbal at intervals 9 and 11. Density was then shifted to the high tom at intervals 9 and 11. Ultimately, the shift in density produced changes to the movement level in the current selection of the sample selection algorithm. Despite the change in movement levels, however, once the previous selection was taken into consideration the movement levels became more consistent, with *low* movement levels occurring at intervals 1 and 6. Consequently, the timbral stability of events at intervals 1 and 6 were greater than at other intervals in the pattern. Those intervals corresponded to the second ride cymbal strike, in both semitone intervals, in the pattern. The movement levels for the fundamental pattern for this study are shown in Figure 5.2.

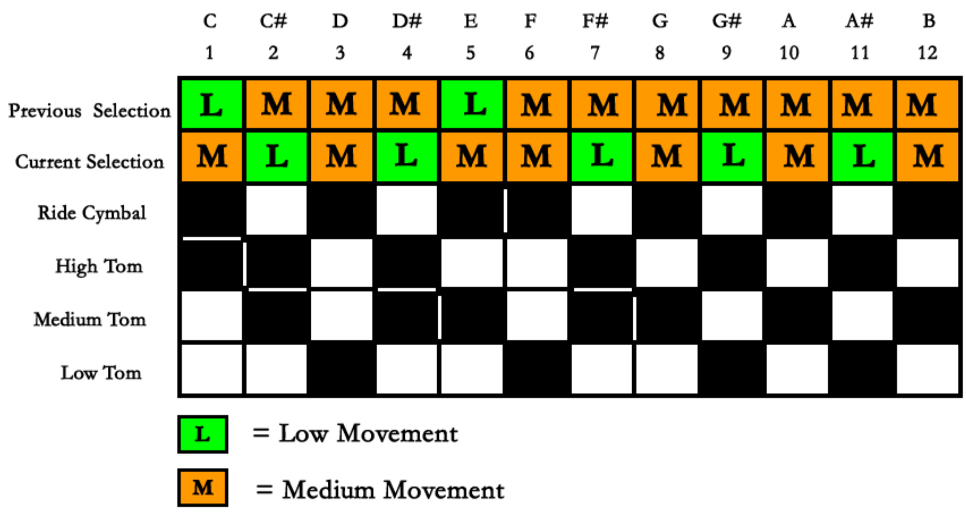


Figure 5.2: Movement levels of the fundamental pattern for Study No. 1.

While the rhythm pattern of Study No. 1 extends the complexity, and complementary nature, of the period sub-samplings, it presents difficulties in parsing the patterns individually. This is due to algorithmic constraints, and the instrumental limitations of the software. African musical form consists of rhythmic fragments and accompaniment patterns. This study deconstructs the fundamental rhythmic pattern and augments it with additional instruments. The independent evolution of the additional instruments, in relation to their usefulness over time, indicated by numbers in the sequence, is reflected in Figure 5.3. The instruments are, for the most part, Essential/Constant, with the exception of the ride cymbal, whose evolution though different levels of usefulness moves from Essential/Variable, to Non-Essential/Variable, and back again. This demonstrated a clear correlation between the idiophones and the increase in broad bandwidth. The evolution of the ride cymbal, therefore, allowed a greater exploration in the interplay between broad bandwidth, rhythm, and melody. This exploration was facilitated by strategically ensuring that other idiophones were either Essential/Variable or Non-Essential/Constant. By reducing the impact of the ride cymbal, from an Essential/Constant role, the underlying rhythmical patterns in the membranophones were exposed.

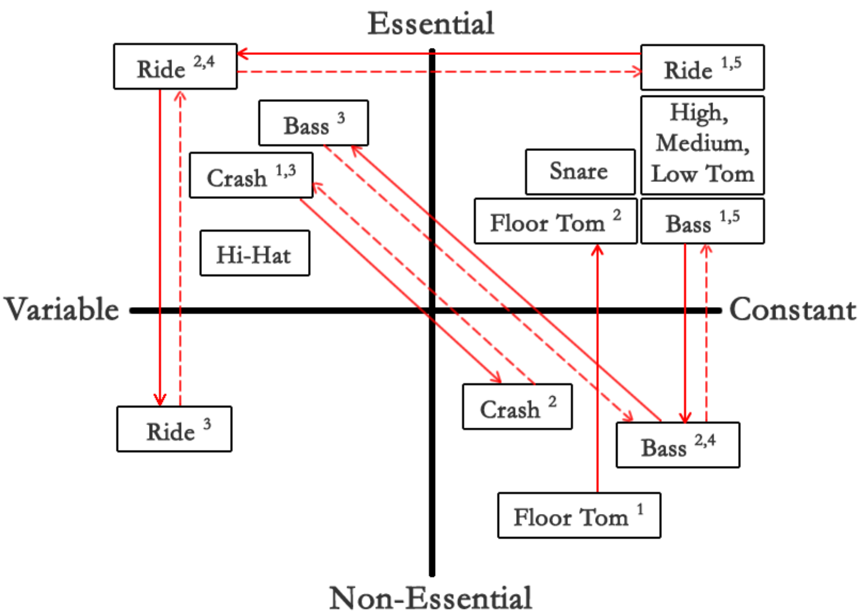


Figure 5.3: The hierarchy of usefulness for the instruments in Study No. 1.

This study can be separated into seven passages:

1. Introduction and rising tension of the fundamental pattern;
2. Extension of the pattern to include the bass drum;[52]
3. Exposure of the tom-toms in the fundamental pattern, with the ride cymbal becoming Non-Essential;

[52] In this part of the study, the bass drum is intended to be representative of the pulsating sound of the dancers' footsteps during an African music performance. At this point, the ride cymbal becomes variable, and the tension from the previous section is released.

4. Return of the ride cymbal with accompaniment fragments of Essential/Variable instruments;
5. Low complexity rhythmic pattern using constant instruments not included in the fundamental pattern;
6. Membrane pattern, increasing in complexity from basic rhythmic fragment; and
7. Rhythmic complexity increases, all instruments audible, leading to a resolution.

Study No. 1 drew on the broad bandwidth, melody, and rhythm paradigm for compositional purposes. This provided a very useful method for visualising the parameters over time and, specifically, mapping the evolution of the parameters across different sections of a piece. The first section of the study emphasises rhythm, for example, with less emphasis on bandwidth and melody. The subsequent two sections see an increase in bandwidth. This is followed by section four, which shows an increase in rhythm while maintaining the bandwidth. Sections five and six gradually decrease in speed and volume. The final section finishes the piece with high broad bandwidth and high rhythmic complexity. Figure 5.4 is a visual representation of the changes in broad bandwidth, melody and rhythm across the different sections for Study No. 1.

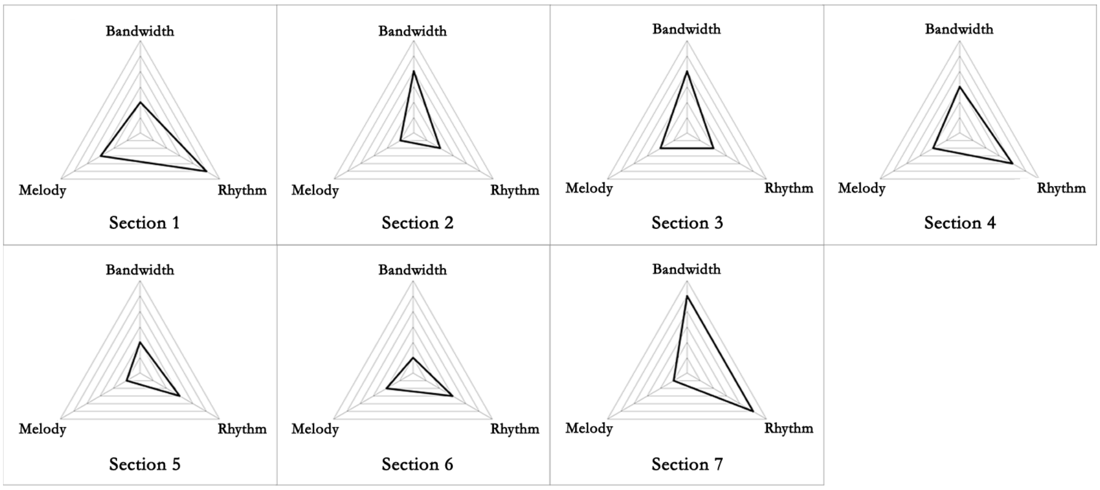


Figure 5.4: Broad bandwidth for the sections in Study No. 1.

5.2 Study No 2: African Meso-Periodic (II)

- **Length: 04:50**
- **Parameters: Loudness**
- **Skill Level: Unskilled**
- **Number of TUBS Banks: One primary, two supporting**

Study No. 2, like the previous study, was composed using one primary bank, in conjunction with two supporting TUBS banks. It also used "loudness" as the main feature-based parameter. However, in this study the skill level of the drummer, from which the human performance data was initially drawn, was "unskilled". Following on from Study No. 1, this piece makes a deeper exploration of meso-periodic rhythms in an attempt to extend the reach of the compositional tool. The bell pattern was replaced with a floor tom in a C major scale pattern. A timpani was used to accentuate various rhythmical elements, and to convey a sense of both "motion" and "tension". The pattern was spread across three different TUBS banks, in order to produce lower movements levels, compared to those in Study No. 1. Notably, only the first bank (with the fundamental pattern) contained a current sequence with more than one instrument being played simultaneously.

The first section of this study used a TUBS notation for the fundamental rhythm and augmenting patterns. This is shown in Figure 5.5, and includes an overlay of the patterns in the second and third TUBS banks. Where an instrument in TUBS banks two or three are played at the same time as an instrument in the first bank, the bank number is indicated in the period As the pattern progressed, the low, medium, and high toms were introduced. This rhythmical fusion produced an interesting compound meso-periodic rhythm, which contained different pitches that conveyed melodic undertones. The consistent pitch stability saw variations in timbre and dynamic strike levels. This can be attributed to the sample selection algorithm, which gives the impression of multiple performers improvising a similar pattern, as seen in African music. In this piece, the introduction of the bass drum evoked the pulsating sound of dancers' footsteps, which is an inherent part of African music, in much the same way it did in Study No. 1. At the same time, the regularity of the bass drum created a tension between the genre and rhythm of African music and Western electronic dance music. The increased pattern density, augmented by the bass drum, also increased the perceived movement

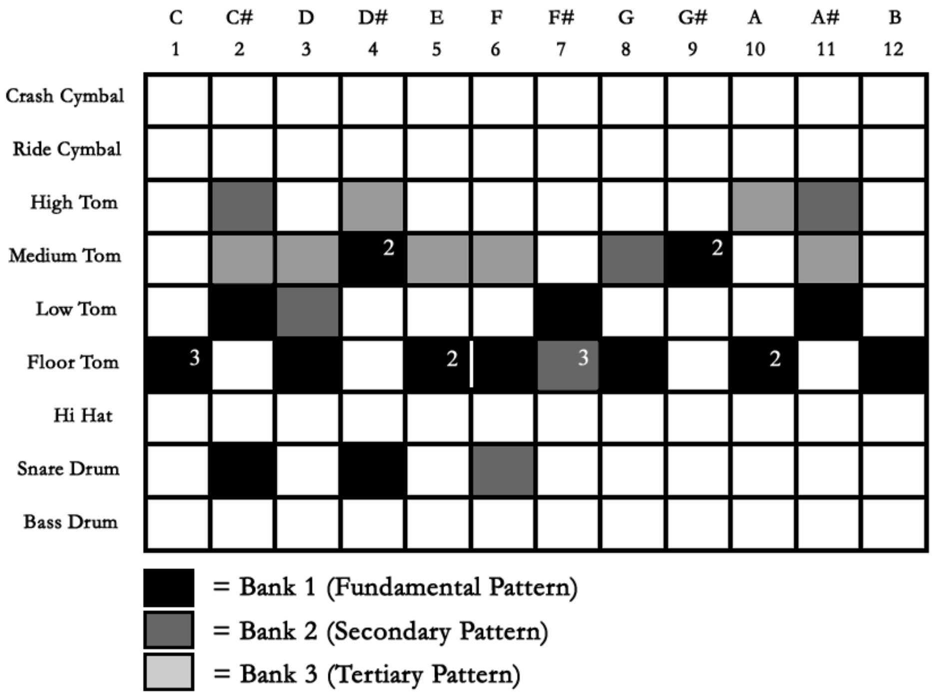


Figure 5.5: The TUBS notation for the fundamental rhythm and augmenting patterns on the secondary and tertiary TUBS banks, for the first section in Study No. 2. Adapted from Schloss (1985).

complexity in the performance model. This presented some rhythmic irregularities, whereby the temporal location, and irregular dynamic levels of the strikes, made the rhythm feel "rushed". It also made the piece sound like the performers had reached their playing capacity.

In the previous study, it was noted that the sub-samplings of the periods between the ride cymbal and the high tom were complementary, thus increasing the rhythmic instability for parsing the rhythm. In contrast, the pattern in the first section of this study saw the ride cymbal replaced with the floor tom, as an Essential/Constant instrument in the rhythm. While the other tom-toms in this pattern are complementary, there is greater stability in the second half of the rhythm. This can be attributed to the spectral similarity of the instruments in the pattern, which is indicative of the pattern having a lower broad bandwidth, thereby producing an increase in melodic and rhythmic elements. Spreading the patterns across three TUBS banks increased the density referent, while maintaining lower overall movement. This resulted in greater timbral stability, and greater rhythmic complexities.

The second section of Study No. 2 saw the instrumental focus move towards the idiophones. The Essential/Constant nature of the floor tom was replaced by the ride cymbal, with the hi-hat and crash cymbal supporting this transition. The low-tom was reintroduced, with an increased level of usefulness. The bass drum added pulse at a lower metrical level. The TUBS notation for this part of the study is shown in Figure 5.6.

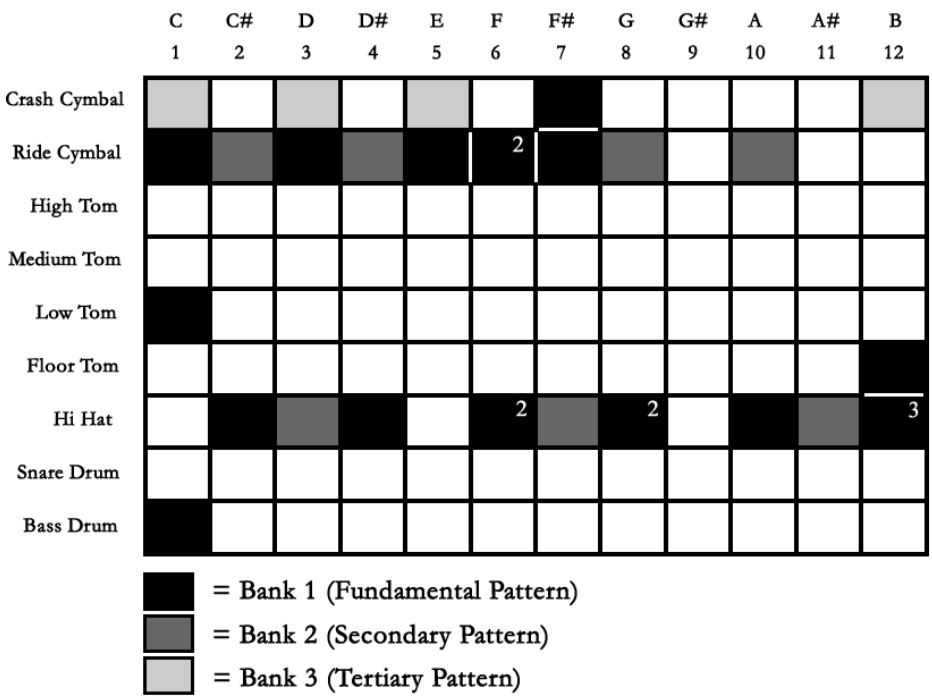


Figure 5.6: The TUBS notation for the fundamental rhythm and augmenting patterns on the secondary and tertiary TUBS banks, for the second section in Study No. 2. Adapted from Schloss (1985).

The second section of this study has similar characteristics to the first. For example, the broad bandwidth is lower, due to the focus on idiophones. This resulted in a less melodic rhythm, compared to the first section where the tom-toms were the dominant rhythmical component. This can be attributed to the mechanical properties of idiophones, and the production of nonlinear chaotic frequencies. Notably, by spreading the patterns across three TUBS banks, timbral variation was reduced while maintaining a higher density referent. This is demonstrated by the timbral stability in ride cymbal strikes across the patterns.

Interestingly, for the most part, the first half of the fundamental pattern in section one had *medium* levels of movement. Combined with the density referent caused by the number of simultaneous strikes in the other TUBS banks, at lower levels of movement, there was a mixture of stable and varied timbral variation. However, this effect is less pronounced in the second section. A comparative diagram illustrating the movement levels for the two fundamental patterns for both sections is shown in Figure 5.7.
Study No. 2 is structured in rounded binary form, based upon the repetition of rhythmic patterns in the low tom. While this is apparent both the first and second sections of the study, it is not overtly noticeable because the density referent, and the spectral similarity of multiple patterns on the same instrument, presents problems perceiving the isolated pattern. The second occurrence of the primary low tom pattern appears in isolation, stripped down to a skeleton pattern whilst still being both Essential/Constant in the hierarchy of usefulness. This is shown in Figure 5.8.
In this study, the transitions in the different degrees of usefulness are more "one-way" than in Study No. 1, that is to say, the instruments do not return to their original levels of usefuleness but remain in the new usefulness context. Therefore, the changes in usefulness are more pronounced in this piece compared to the previous piece. The first study demonstrated the use of movement across quadrants, at different times, and particularly involving return movements (e.g. the bass drum, and the crash and ride cymbals). In this study, transitions between usefulness stemmed from intermittent instruments. In the first section, intermittent instruments were used to accentuate tension. In the second section, these intermittent instruments constituted the fundamental rhythm.

The most significant difference between the first and second sections of Study No. 2 is the change in the Variable/Constant nature of the idiophones. Because the tom-toms were the most dominant instruments in the first section of this piece, albeit augmented by idiophones at various points, the overall broad bandwidth was reduced. Moreover, because tom-toms tend to convey a sense of pitch, the overall structure of this pattern conveys a greater sense of melody. With a greater sense of melody, is an increased per-

Figure 5.7: A comparative diagram showing the movement levels of the fundamental patterns for sections one and two in Study No. 2.

ception of rhythm. In the second section of the piece, the transition to idiophones as the principal instrument type reduced the melodic feel of the rhythm. The reintroduction of the tom-toms, bass drum, and the snare drum increased the overall bandwidth and, at the same time, reduced the sense of rhythm. A comparison of the bandwidth in both sections of this study is show in Figure 5.9.

Study No. 2 demonstrates the ways in which the compositional software tool, can extend the fundamental concepts of meso-periodic rhythmic construction. By extending these concepts, composers can create new compositions that explore the synergies between rhythm and melody, particularly in the context of an instrument's broad bandwidth and rhythmic context. Furthermore, this study demonstrates the suitability of

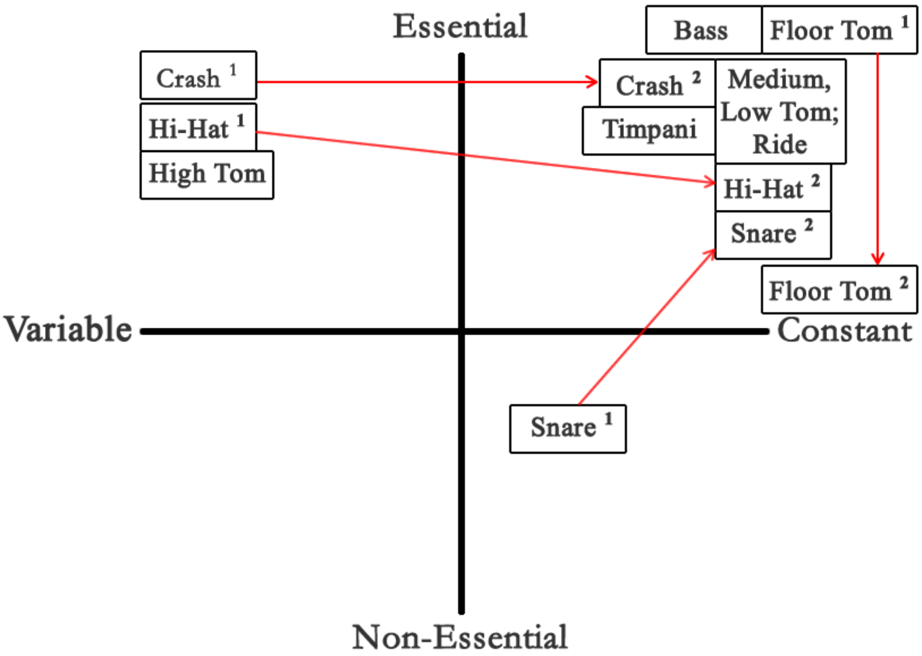


Figure 5.8: The hierarchy of usefulness for the instruments in Study No. 2: African Meso-Periodic (II).

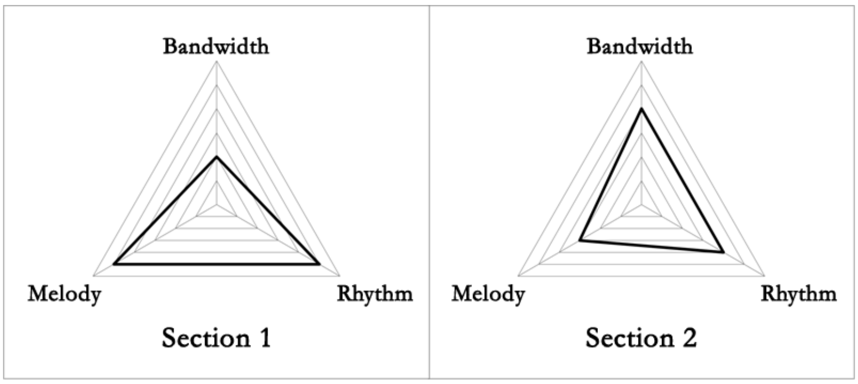


Figure 5.9: Movement levels of the fundamental pattern for Study No. 2.

the compositional tool in creating temporal and timbral variations, which are inherent in meso-periodic African music, and in creating different rhythmic effects for differing rhythmic contexts.

5.3 Study No 3: Exploration of Pitch Variation in the Tom-Toms

- **Length: 03:37**
- **Parameters: Spectral flatness**
- **Skill Level: Semi-skilled**
- **Number of TUBS Banks: One primary, one supporting**

Study No. 3 was composed using one primary bank and one supporting bank. It used "spectral flatness" as the main feature-based parameter. The skill level of the drummer, from which the human performance data was initially drawn, was "semi-skilled".

Tom-toms convey a greater sense of pitch than other membranophones, and each tom is capable of conveying a different pitch. In fact, given the potential for dis-uniform tuning, it is possible for a tom-tom to convey different pitches across the surface of the membrane. This study begins with the floor tom, and explores how the pitch varies across the instrument, with a view to investigating the relationship between pitch and timbre. From 17-28 seconds, the floor tom pattern presents interesting varieties of pitch and timbre, including some intra-instrument pitch discordancy, and some dominantly resonating high-low pitches, reminiscent of other pitched instrument types. This distinct high-low pitch combination produced a sense of tension, which was compositionally explored by accentuating the lower pitches with additional instruments. This both heightened and temporarily resolved the tension. At the same time, random dynamic amplitude variations (including the inherent conveyance of strike strength) reinforced the tension created by the pitch variations. This tension was then dispersed by the introduction of the medium tom, whose dynamic amplitude variations were more stable, but whose pitch was discordant with the floor tom, thereby producing a "drone" effect. The modulatory strike pattern of the medium tom caused the discordancy to subside, returning temporarily to the original floor tom, which, in those few short strikes, continued to generate intra-instrumental pitch discordance.

Following the investigation of joint pitches and timbres between the floor and medium tom, was an exploration of the pitch interaction in the floor and low tom. The low tom pattern was more irregular than that of the medium tom. The result was that, with an increase in density, a complex rhythm was produced which had no regularity or discernable pulse, but retained a sense of tension through intra- and inter-instrument pitch differences (00:58-01:06). The next stage of the exploration saw the addition of the high tom, combined with the floor tom and the low tom, as well as the reintroduction of the medium tom (01:28-01:45). The high tom assumed a role similar to that of the snare jazz drum in order to augment the addition of the bass drum at regular intervals relative to the low tom. The addition of the high tom and the bass drum enhanced the pulse, and added a beat to the floor and low tom combination. However, despite the addition of this beat, deriving a pulse from the irregularity of the floor and low tom patterns produced a rhythmic dissonance, increasing the sense of rhythmic complexity and irregularity.

The complexity and irregularity of the floor and low tom patterns produced non-repeated, small, melodic patterns as the PD-103 cycled through the pattern and selected the samples. At the same time, the speed and density referent of the patterns made it difficult to parse individual pitches. Consequently, regularity in the sub-patterns was inferred through the presence of individual pitches and timbres with high dynamic amplitudes. In the fourth section of this study, a change to the floor, high, and low tom patterns continued the irregular rhythm, with the slower tempo creating more space for the different timbres and pitches to be recognised. The first floor tom pitch in this section provided a reference pitch, which can be recognised throughout the section, and which created tension and contrast against other pitches. The droning medium tom pattern, from the previous sections, was reintroduced in the second part of this section to provide rhythmic stability. This rhythmic stability, however, was reduced when the floor tom pattern was played at double time. The louder strikes and dominant pitches of the floor tom masked the medium tom.

Inherent in this study is timbral discordancy, which can be attributed to the use of "spectral centroid" as a compositional parameter. In particular, this is due to the similarity in the brightness of strikes across the surface of a tom-tom, being distinct from the perceived pitch differences. This secondary effect was described in section 4.6.1. Nevertheless, a sense of tension was evoked by the discordancy in the tom-tom pitches that was heightened by the irregularity and complexity of the rhythm. This relationship is worthy of future investigation.

5.4 Study No 4: Spectral Centroid and Isomorphic Rhythm

- **Length: 05:03**
- **Parameters: Spectral centroid**
- **Skill Level: Semi-skilled**
- **Number of TUBS Banks: Three primary**

Study No. 3 was written solely for percussion, using the complete range of instruments contained in the PD-103, and augmented with synthesiser, brass ensemble, timpani, and crash cymbal. The study used “spectral centroid” as the main feature-based parameter. The skill level of the drummer, from which the human performance data was initially drawn, was “semi-skilled”. The study was composed using all three TUBS banks. Each bank represents a different level in the metrical hierarchy, with differences in velocity across the levels. These are:

- Bank 1: 1/4 note. The velocities of all events in this bank are at maximum;
- Bank 2: 1/8 note. The velocities of all events in this bank are at 50
- Bank 3: 1/16 note. The velocities of all events in this bank are at minimum.

In the General Theory of Tonal Music put forward by Lerdahl and Jackendoff (1983), hierarchical meter is based upon amplitude. This study expands on this concept, in the context of spectral centroid. In particular, the aim of this study was to examine the effect of spectral centroid on the organisation and perception of different metrical levels. Continuing the TUBS notation, and construction of African meso-periodic music, the high tom pattern in this piece was expressed as the diatonic scale. The medium tom was expressed as a pentatonic scale, with an added E(5). The low tom had strikes at D(3), F(6), and A(10). The floor tom had strikes at C(1) and F(7). While the pattern of this study is similar to those of Study No. 1 and No. 2, it makes use of different instruments. This study differs in its use of the high and medium tom, as well as the recursive hierarchical meter with scaled spectral centroid. A rudimentary illustration of the patterns, and their relative spectral centroid amounts, at different metrical levels, is show in Figure 5.10.

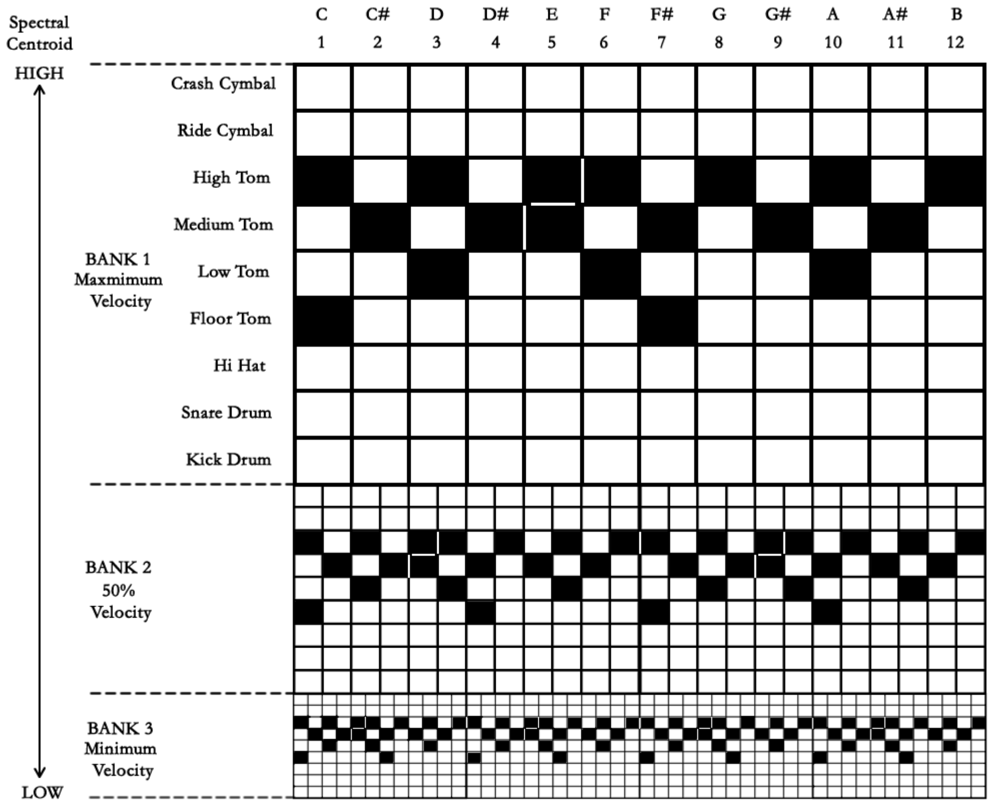


Figure 5.10: An illustration of the isomorphic rhythmic patterns in Study No. 4.

This study, using spectral centroid as the main feature-based parameter, produced some interesting compositional results. Most noticeably, dull sounds from the lowest metrical level (Bank 3). In general, notes generated from the highest metrical level (Bank 1) had considerably more amplitude and brightness. This observation is commensurate with greater strike strength, and the excitation of more modes, thereby producing an effect similar to that of a metrical hierarchy by amplitude.

In addition, the use of the tom-toms provided an opportunity to explore the timbral and melodic potential of the different metrical levels. This is particularly noticeable at the beginning of the study (15-40 seconds). In this section, the first Bank 1 cycle was repeated to produce a familiar melodic and rhythmic line, which added coherence.

The start of section two (01:20-01:40) was composed using TUBS banks 1 and 2, whilst Bank 3 was omitted. This resulted in a less complex, and less dense, pattern, with an increase in brightness over the cycle. As the piece began to slowly deconstruct, layers of brightness and complexity were removed. Notably, many of the strikes in this piece consisted of rim shots that were coincidentally captured in the sample database. The spectral centroid parameterisation made these strikes more prominent because the rim shots themselves contained wide bandwidth frequencies and brightness. A secondary reduction in rhythmic complexity occurred at 03:30-03:50, which contained no distinct rhythmic patterns. However, this did not necessarily reduce the complexity of the pattern in this piece because consecutive notes used different instruments. In particular, this piece contains significant timbral variation, metrical layers, and depth. This can be attributed to the range of velocities available to the composer, with each velocity occupying a different subset of samples.

The final part of this piece began with a slow crescendo of simultaneous rhythms, some of which are reminiscent of patterns in the previous studies. Because each of the different metrical levels employ different samples from the same instruments, it is difficult to identify a pulse. Firstly, this makes it difficult to isolate samples from a single bank. Irregularities in the samples, within the TUBS banks, make it difficult to identify regularly recurring patterns, even though the representations of the cycled patterns are regular. Secondly, there is a large IOI time of 5.5 seconds with the floor tom in Bank 1. Such a large IOI makes identifying a pulse, particularly with the density and rhythmic complexity in between, difficult.

5.5 Study No 5: "Quark"

- **Length: 01:00**
- **Parameters: Loudness**
- **Skill Level: Skilled**
- **Number of TUBS Banks: One primary, one supporting**

Quark was written for Vox Novus 60x60 Radio Request Extravaganza, and aired on August 24, 2012, on WDGR 91.1 FM, Plainfield, Vermont U.S.A.[53] It was based on the premise that, since the beginning of civilisation, humans have excited different objects in order to produce sound. In addition, rhythm, as a fundamental aspect of music, has undergone dramatic change. Quark represents the evolution of percussive rhythm, over human development, from the distant clanging of early man, to the techno-centric rhythms in modern popular music.

Quark used the complete range of instruments contained in the PD-103. It was composed using one primary TUBS bank, and one supporting TUBS bank. It used "loudness" as the main feature-based parameter. The skill level of the drummer, from which the human performance data was initially drawn, was "skilled". Timpani and synthesiser augment this piece.

One of the challenges in composing this piece was how to communicate the concept of rhythmic evolution within a one-minute timeframe. This was achieved by breaking the piece into two parts: the "old" and the "new".

The piece opens dramatically with timpani (at 0-24 seconds) and marks the beginning, and early growth of, human civilisation (the "old"). The synthesiser adds a layer of motion, and represents the fast-forwarding of clock hands as time moves forwards through the centuries. The initial tom-tom patterns are filtered, with the early filtering representative of early percussion. The irregularity of the rhythm represents early musical formation, which, coupled with low-pass filtering, is representative of the primitive construction of percussion, with all its inherent vibrational flaws. As the filter evolves, so too does the rhythmic pattern, which incorporates a snare drum and

[53] The full radio programme is available to download from: https://v2-staging-nickf.soundcloud.com/wgdr/60x60-radio-extravaganza-part.

additional tom-toms. The addition of these instruments is representative of the growth of human civilisation, as generations expand and build upon previous knowledge. The second timpani (at 12 seconds) signals a change in the speed of evolution. This is intended to depict evolution moving forward thousands, not hundreds, of years. The conceptual speed shows empires rising and falling. Three things heighten the tension here. Firstly, the evolution of the low-pass filter, which, at this point, does not feel resolved. Secondly, the synthesiser increases in frequency and amplitude to form the start of a crescendo. Thirdly, the gradual introduction of idiophones represent technological progress.

At 24 seconds, just prior to the third timpani, a small rhythmic stutter, created by an unusual sample selection of the snare drum, coupled with an unexpected crash cymbal strike, represents the transition from “old” to “new”. It is this crash cymbal that marks the end of the first section of this piece.

The second section of this piece begins at 24 seconds with the bass drum, which is reminiscent of electronic dance music, and signifies the early proliferation of technology in music, and the beginnings of the interconnectedness of human civilisation from both a cultural and musical perspective (the “new”). The irregular metrical locations of the crash cymbal, as well as the sudden and short appearance of the snare drum, provides unexpected accentuations, similar to unexpected global political and historical events.

A final challenge of this piece was to present the concepts of “old” and “new” in the context of the compositional software tool, and by simulating real human percussive performance.

5.6 Study No 6: Demonstration of Improvisational Application

- **Length: 09:18**
- **Parameters: Loudness**
- **Skill Level: Unskilled**
- **Number of TUBS Banks: One primary, one supporting**
- **Additional equipment used: Novation Remote 37SL**

Study No. 6 was composed using one primary bank, in conjunction with one supporting bank. It used "loudness" as the main feature-based parameter. The skill level of the drummer, from which the human performance data was initially drawn, was "unskilled".

When designing a creative software application, it is important to ensure that the user can maximise the potential for creative application by producing the most relevant and versatile music tool possible. The improvisational capabilities of the PD-103 allows for the accurate creation of dynamic phrases and embellishments, which, due to the stochastic nature of the sample selection and temporal variation algorithm, is more difficult under normal operating conditions. The PD-103 sample database was mapped from MIDI keys 0-40. Each key triggers samples from a demarcated performance zone (one zone for the bass drum, five zones for all others). In addition, the MIDI keys were not subject to timing variations. Thus, this study should only be considered in relation to the improvisational application of the software, when combined with composed patterns.

There are two key challenges in composing a piece that uses composed and improvised material. The first challenge, which is unique to this compositional tool, concerns the disconnect between the user interface and the sound material. Mapping the samples to different keys on a MIDI controller, for example, makes it difficult to adequately improvise, particularly in the way a real drummer might improvise during a performance. The second challenge lies in determining the function of the improvisation, and whether the improvisations are intended to augment the composed rhythms, or to act as the main aspect of the composition, with the composed rhythms assuming a supportive role.

Limitations in the implementation of the software, notably the way the samples were loaded and played back, and the disconnect between the interface and sound of the instruments, made the improvisational implementation more suited to an exploration of timbre by simulating ornamental free-form phrases. In this piece, these free-form phrases were used as structural devices to develop tension and release tension between the composed rhythms, and support the increase in tension created by the harp.

One notable part of the improvisation are the snare drum rolls. which had to be manually edited. While it is possible to perform snare rolls in a real drum performance, it is not easy to reproduce snare rolls on a keyboard using samples. One particular challenge of manually editing a snare roll is how to manage the timbral variety of the roll. This was achieved by ensuring that the samples were not repeated. It was also achieved by changing the temporal compression and expansion of the roll, so that the drum roll had enough dynamic variation to convey a sense of improvisation by a real drummer. In addition, small variations in timbre and dynamic accents were added to also convey a sense of improvisation.

The composed rhythms gradually increased in tempo towards the end of the piece. To achieve this, it was necessary to increase the speed of the improvisation. This is apparent at approximately 3 minutes, where fast, cross-instrument strikes are played. Although it is possible for a real drummer to achieve these speeds, the instrumental variation in the sequence combined with the speed of the strikes, may be more conducive to producing perceptual effects such as temporal masking. This may present difficulties in parsing the improvisation as that of a real drummer.

The improvisations contained in this piece provide a convincing simulation of real human improvisation on drums. Further work could address the challenges discussed above in order to explore improvements to the implementation. For example, mapping the samples to real drum pads for MIDI performance by a real drummer, or by using instruments other than the harp.

5.7 Study No 7: Electroacoustic Study

- **Length: 10:00**
- **Format: Surround Sound (5.1)**

One important goal of this research was to extend the PD-103 beyond existing genres of music. To demonstrate the applicability of the compositional tool for composers of other musical genres, a surround-sound electroacoustic piece was created using individual samples from the PD-103. This study combined rhythmical elements from Study No. 3, with samples with different timbres, in order to create subtle variances in sonic effects and textures. This study also comprised micro-timbral and micro-temporal variations, which were inherent in the rhythmical elements in Study No. 3.

This piece has two sections. The first section is an exploration of micro-timbre. Using different effects-processing algorithms and spatialisation techniques on different membranophone and idiophone samples produced an intensification of micro-timbral differences. This resulted in consonant and dissonant spectra and artefacts. Much of the consonance and dissonance encountered in the first section of this study was caused by the physical properties of the instruments, such as their vibrational characteristics and resonance and decay properties. These physical properties were manipulated to conceptually represent the flowing effect of vibration through different molecules in order to produce different sounds.

The second section of this study also used effects-processing algorithms and spatialisation techniques in order to magnify the micro-timbral and micro-temporal variations that were inherent in Study No. 3. In some instances, micro-variations in the rhythm were in conflict with different effect-processing algorithms. In nearly all instances, the micro-variations were magnified. This produced unexpected results, which cascaded through various rhythmic and textural elements in order to create an entirely new rhythmic experience. The spatialisation of the micro-timbral and micro-temporal variations conceptually represent the listener existing microscopically, observing the molecule interaction caused by the drum strikes.

In this piece, spatialisation was approached from several different perspectives. Firstly, the layered background textures contain high pitch noises reminiscent of distant screams. To provide depth and a sense of chaos, these textures were generated by short, delayed, filtered, bass drums. This background texture also has a circular panning trajectory, which accentuates a feeling of motion, space, and urgency.

The foreground textures were created from individual, processed, membranophone strikes of the bass drum and different toms. The bass drum produced small rhythmic fragments, which were interrupted by a powerful floor tom strike that was accentuated by a prior escalation of background textures. The floor tom was filtered and delayed, creating a pulse that continued beyond the second strike. Beyond the second strike, the floor tom textures were explored, evolving from a deep pulse to a single-pitched rhythmic texture. Finally, the gradual mutation into noise (via bit-crushing) was subjected to a low-pass filter, which maintained a steady, background pulse.

The textured background, combined with the pulse, conveys a sense of urgency and tension, which slowly concludes after the ride cymbal (at 2:50). The finality of the ride cymbal in this section is critical. It marks the transition from a timbral exploration of membranophones, to a timbral exploration of idiophones. At the same time, the processing and increased resonance of the ride cymbal changes the sonic qualities of the strike in such a way that its sound conveys a different excitation characteristic, from a strike, to a scraping texture. Moreover, the frequencies in the ride cymbal, which are usually chaotic and non-linear, are much smoother, with bell undertones.

The spatialisation approach to the cymbals was also chaotic in nature, with almost tumultuous panning. The prominent ride cymbal strike (at 3:34) demonstrates an increase in higher frequencies, and produces a more crisp, scraping texture. In addition, the decay of the ride cymbal, which continues for almost sixty seconds, was gradually modulated with an increase in resonance. This produced an interesting spectromorphological effect whereby the final parts of the decay resemble a single tone.

A reverb and equaliser was applied to one floor tom (at 4:10), signalling the end of the timbral exploration of the idiophones. The floor tom itself symbolically represents large planks of wood being dropped on the floor of an empty warehouse. At this point, background textures and the floor tom pulse were reintroduced in order to create a sense of tension, which was heightened by an increase in volume. The regularity of the pulse was then contrasted with a delayed, pitch-shifted, snare drum whose spatialisation followed the delayed drums. These effects evoke the sound of a snare drum rolling down a flight of stairs, bouncing on the steps.

The second section began with a sustained buzzing texture (at 5:16), which increased in pitch. This was followed by the introduction of a rhythmic pattern from Study No. 3. Separating the individual components of the rhythm, applying individual delays, and assigning them to opposing spatial locations increased the complexity of this rhythmic pattern. A bass tone added a sense of metrical regularity and tension to the pattern. This pattern was then gradually deconstructed. The removal of the floor, left, and right toms produced a scarce rhythm which, in combination with delays and reverbs, conveys a sense of irregularity and space.

The final part of section two began with a reversed, delayed, ride cymbal (at 7:47). A second pattern, comprising the medium and high tom, was then introduced. The pitches and timbres in the rhythmic pattern were intended to be reminiscent of African percussive music. The delays applied to this rhythm created small repetitions, and include some minor phasing from applied modulation. This evoked a sense of release from the tension in the previous sections, which is accentuated by the interspersed bass drum at different metrical points. The piece finishes with a slow decay of the delayed rhythmic pattern, overlayed by the background texture from the beginning of the piece, which crescendos to completion of the piece.

5.8 Study No 8: Demonstration of Skill Levels

- **Total Length: 02:40**
- **Part 1: Solo Unskilled (length: 00:25)**
- **Part 2: Solo Semi-skilled (length: 00:28)**
- **Part 3: Solo Skilled (length: 00:27)**
- **Part 4: Accompanied Unskilled (length: 00:25)**
- **Part 5: Accompanied Semi-skilled (length: 00:28)**
- **Part 6: Accompanied Skilled (length: 00:27)**
- **Parameters: Loudness**

Study No. 8 is intended as a demonstration of performer skill levels. Divided into six sections, this study comprises of solo "unskilled", "semi-skilled", and "skilled" performances, and accompanied "unskilled", "semi-skilled", and "skilled" performances. It was composed using "loudness" as the main feature-based parameter.

When forming a musical group, performer skill is often a key factor in the choice of a musician. However, skill is a highly subjective quality and its perception is dependent on a number of factors. Two ways in which skill can manifest itself in a drumming performance is through consistent performance timing, and stable timbre production.

The extraction of performance data, from three differently skilled drummers, was a significant component of this work. In section 4.6, a comparison of participants' timing fingerprint data found differences in the consistency of onset deviations, and inter-onset tempos, across participant skill levels. A subsequent evaluation of performance movement led to the formulation of a hypothesis, which linked levels of movement and timbre production to higher levels of movement, which led to higher variations in strike location and, consequently, more variations in timbre production. This hypothesis was tested by way of zone weightings, as described in section 4.7.

This study, therefore, seeks to demonstrate performer skill by providing a direction comparison to each skill level. It is comprised of six parts, each part using an identical rhythm. Parts one to three contain an identical solo drum rhythm using the unskilled, semi-skilled, and skilled presets respectively. Parts four to six contain the same three excerpts, augmented by additional jazz VST instruments. Because the jazz instruments are MIDI-based, their playback is consistent. This is particularly useful in providing a reference point with which to listen to the drums.

From a temporal perspective the differences between the three skill levels are subtle. The unskilled preset appears more rigidly timed, whilst the semi-skilled preset had more "feel". The skilled drummer had a distinct swing, particularly on the ride cymbal. From a timbral perspective, the unskilled preset had much less timbral stability. During the demonstration some of the instruments were unexpectedly accented (e.g. the ride cymbal and the hi-hat), while other instruments appeared to have softer strikes in some places, and stronger strikes in others, with little regard for structural appropriateness (e.g. the snare drum). The semi-skilled preset had similar timbral characteristics, but was more timbrally consistent with the snare drum, hi-hat, and ride and crash cymbals. In contrast, the skilled drummer was more consistent with timbre, especially in the case of the snare drum, and the accents on the ride cymbals appear to be more consistent. One primary observation was that the results of the timbral selection algorithm worked as expected, although it is not easy to differentiate the temporal variations across the skill levels.

In the context of the jazz accompaniment, the performance from the unskilled preset appears largely uninspiring. The variations highlighted above were more prominent, and had a much greater effect on the overall excerpt. For the semi-skilled drummer, the differences in the snare drum, most notably the greater consistency, contribute to a more "bouncy" feel. The ride cymbal in the skilled preset produced a different feel compared to the unskilled and semi-skilled presents, producing more "swing" than "bounce". Further listening tests, however, would provide more insight into the implementation in order to make refinements to the algorithms.

5.9 Study No 9: Demonstration of Parametric Variation

- **Total Length: 05:15**
- **Part 1: Loudness (length: 01:45)**
- **Part 2: Spectral Flatness (length: 01:45)**
- **Part 3: Spectral Centroid (length: 01:45)**
- **Skill Level: Skilled**

The compositional approach in Study No. 9 used all three feature-based parameters: loudness; spectral flatness, and spectral centroid. In order to understand the differences and compositional effects of these parameters, this study comprises three short, identical, rhythmic patterns using each of the three parameters. One of the intentions of this study was to create a simple, repetitive, rhythmic pattern, using the loudness parameter, and to transpose this onto the other two parameters (spectral flatness and spectral centroid). This piece is augmented with Djembe and Krin in order to provide a rhythmical reference point. In order to assist in the understanding of the effect of the feature-based parameters, Figure 5.11 (overleaf) shows a spectrogram of the first 34 seconds of the three parts of this study, with loudness (top), spectral flatness (middle), and spectral centroid (bottom).

In part one, loudness, there is an identifiable motif towards the end of each cycle, which emanates from the right tom. Although the pattern is temporally stable, the ride cymbal has periods of both higher and lower amplitude. In contrast, the rhythm, using spectral flatness, had no discernable motif. In general, the sounds were indicative of greater strike strength. This is supported in the spectrogram, with a greater number of high-amplitude frequencies across this excerpt. In this example, the ride cymbals appear to be more consistently struck, with fewer weak sections. Interestingly, the low and medium toms in the spectral centroid example are less pronounced, with the high tom having become more prominent. This is supported in Figure 5.11, which shows a general reduction in amplitude of frequencies below 7019Hz. This overall frequency reduction can be attributed to the default frequency threshold in the MIRToolbox spectral centroid analysis algorithm (1500Hz). Notably, an alteration of this frequency threshold might produce vastly different results and, thus, constitutes an area for further research.

The three parameters shared little similarity in the temporal locations of events with

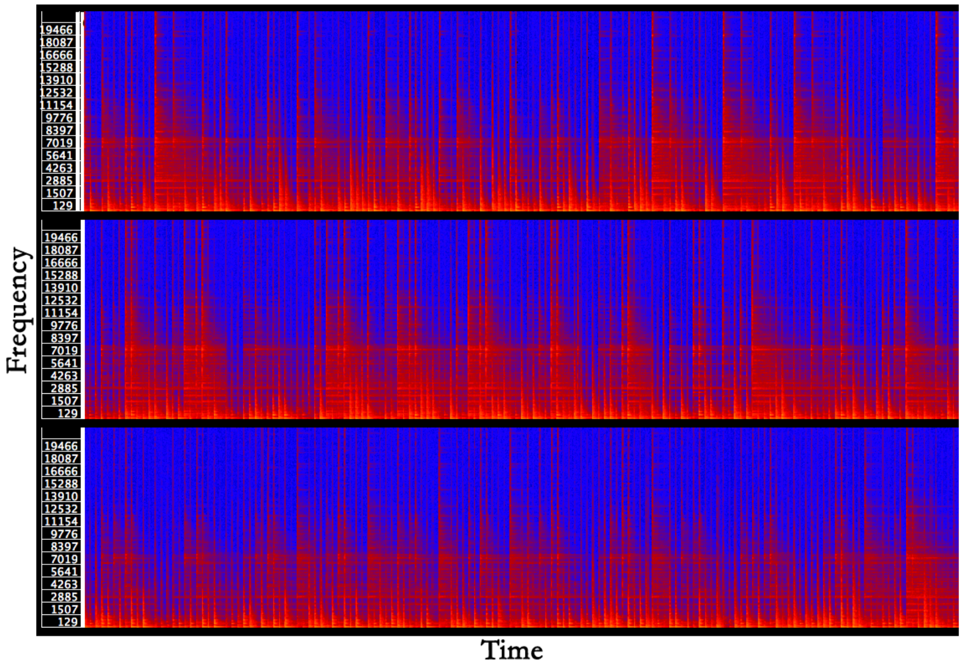


Figure 5.11: A comparative spectrogram of the first 34 seconds of each parameter in Study No. 9, showing loudness (top), spectral flatness (middle), and spectral centroid (bottom).

wide bandwidth. In typical drumming performances it is common for events with similar spectral characteristic to positively correlate to a dynamic phrase or accent. The arbitrary nature of those events, particularly in the flatness excerpt, produced an interesting sequence of variations, which bore little resemblance to real human performance.

Chapter 6

Conclusion

"To emphasise only the beautiful seems to me to be like a mathematical system that only concerns itself with positive numbers"
– Paul Klee, 1906

As discussed in Chapter One, the main aim of this research was to create a compositional software tool that simulates human performance variation in percussion, using a nine-piece jazz drum set, in order to generate new and varied musical works. Due to the interdisciplinary nature of computer music, this research took a multifaceted approach, taking into consideration: the effect of instrumental mechanics on timbre; synthesis techniques for modelling timbre; human performance; computational representation of human performance; and the subsequent implementation and compositional application of the instrumental and performance model.

Firstly, the theoretical framework upon which the performance software model was based comprised of three parts: a nine-piece jazz drum set; human performance; and compositional approaches. Part one involved an investigation into the acoustical and mechanical behaviour of key instruments of a nine-piece jazz drum set, and their effect on the production of timbre. In part two the biomechanics of human percussive performance on a nine-piece jazz drum set were examined. Part three reviewed relevant compositional approaches, including computer-assisted composition, complex rhythms, composing using spectral features, and electroacoustic composition. Secondly, the research took a two-part experimental approach. The first part of the methodological approach was concerned with the collection of timbral data from the instruments of a nine-piece jazz drum set. The main aim of which, was to assess the effects of strike strength and location on drum timbre, in order to develop a methodology with which to create an effective micro-timbre model. The second part was concerned with the collection of human percussive performance information captured from three differently skilled drummers. The aim was to assess the relationship between performer skill level, movement, performance variation, and timing variation during a performance, in order to develop a methodology with which to create an effective micro-timing model. Finally, the composition portfolio and analytical notes demonstrate the compositional applications of the software tool.

This chapter summarises and draws conclusions from the work presented in this thesis, and provides directions for future research.

6.1 Sample Collection, Analysis, and Classification

In order to create a compositional software tool that simulates human performance variation in percussion, the first part of the approach comprised of sample collection, analysis, and classification.

The sample collection comprised nine-thousand individual hits taken from instruments of a nine-piece jazz drum set. It should be noted that the drum set used in the experiment was a typical, and relatively inexpensive, jazz drum set. It had seen considerable use in different environments and, therefore, was not in the best condition. This resulted in lower quality samples, and the potential for vibrational inconsistencies owing to the drum sets history. In addition, at the time the samples were collected the performance classification concept was still in its infancy. This meant that some instruments had greater variations in amplitude and timbre across performance modes at similar parametric values. Although, large timbral variations might also be attributed to the size of the strike surface area in each of the demarcated performance zones, encompassing the area close to the centre and, at the other extreme, the rim. Tension in the drum is higher towards the rim, thus the timbral variation is inherently different.

One of the most important factors to consider, when attempting to sonically represent a nine-piece jazz drum set for sound synthesis in computer modelling, is the inherent micro-timbral variety in the instruments. In this thesis, the sampling synthesis paradigm served to identify the protocols and procedures necessary for capturing timbres for each instrument. The sampling synthesis paradigm included a set of rules that governed the performance of strikes for data capture (see section 4.3.1). The levels of micro-timbral variation, across the different parameters, attest to the robustness of the protocols and procedures. The methods used to capture the strikes, including the microphone type and the positioning of each instrument, produced clear samples with minimal tonal coloration. In addition, the neutral-reproduction characteristics of the microphones effectively facilitated the capture of micro-timbral variations, which were audible across the sample database, and were visually represented in the graphical analysis of the samples.

The sample database had to be prepared for analysis. The beat detection points were manually checked and verified using Pro Tools “Beat Detective”. This prevented truncation errors that could have negatively impacted the sample analysis. However, this process was time-consuming and labour-intensive. One area for process improvement might be to consider alternative methods, requiring less manual intervention, for automatic beat detection. This would allow for the faster creation of new sample databases and timing fingerprints. Once the truncation points were manually checked and verified, MIRToolbox was used to analyse the sample database. This process was relatively straightforward, and the sample database yielded minimal calculation errors (approximately 10% for each instrument, presented as NaN values in the Matlab output files).

It is important to note that it was not the intention to appraise the applicability of the analysis functions to the dataset. This is owing to the potential for idiosyncratic behaviours in the construction of these functions when applied to percussion instruments. One example of this was in the difficulty of correctly analysing cymbals using the autocorrelation function in MIRToolbox. One area for further investigation would be the exploration of the use of the MIRToolbox as an analytical tool for the compositional parameterisation of percussion by producing another, larger, dataset for each timbral parameter. It would also create greater stability in the sample selection by mitigating timbral and amplitude disparities caused by having unequal numbers of samples in each classification. Increasing the number of samples would also increase the number of performance modes. This would enhance the resolution of the instrumental sonification. However, increasing the number of demarcated performance zones would increase the resolution of the performance context, thus requiring more parametric datasets, contextual rules, and performance weightings.

Once the sample database had been analysed in MIRToolbox, it was reclassified using three feature-based parameters (loudness, spectral flatness, and spectral centroid), and reordered based on those parameters (low to high). After listening to the database, it became clear that the process of sample reclassification and reordering would produce some very inspiring aural and compositional possibilities. At the same time, however, the linear ordering of samples by feature had one unexpected consequence. This was due, primarily, to the large number of samples in the database. The linear ordering of samples, using each of the three feature-based parameters, rendered other timbral features non-linear. For example, when the sample database was manipulated to increase loudness levels, from low to high, it produced random pitch variations. This can be attributed to the wide timbral variations in the instruments. This issue was addressed by classifying the drum strikes by location, taking into account the vibrational characteristics of drums (see Chapter Three). Reducing the surface area of each group of samples ensured that each sample, within each group, was subject to similar vibrational behaviours. This resulted in a reduction in differences in timbral characteristics.

It was necessary to demarcate the surface of each instrument of the nine-piece jazz drum set in order to maintain the timbral consistency of the sample database. As discussed in section 4.6.1, one unexpected consequence of arbitrarily tuning the membranes was the presence of timbral inconsistencies in the sample database. Consistent tuning would have made timbral variations less prominent and negated the need to demarcate the surface of each drum component into smaller strike zones. Nevertheless, from a performance-modelling perspective, the decision to demarcate strike zones across the surface of each drum instrument marked an entirely new approach to generating micro-timbre through performance weightings, as well as zonal demarcations based upon vibrational characteristics. Having said that, further research might assess the viability of this approach on uniformly tuned drums.

While the samples in the database displayed other, secondary, timbral characteristics, each of the feature-based parameters (loudness, spectral flatness, and spectral centroid) had their own, unique, characteristics. The "loudness" parameter, for example, worked as expected across all of the instruments. When assigned linearly to MIDI velocity, the changes in amplitude were commensurate with the velocity value. The timbral variations across the loudness curve were more pronounced than expected, however this can be attributed to the strike area of the demarcated performance zones.

Likewise, using a similar MIDI velocity-mapping process, the "spectral flatness" parameter produced some interesting timbral, and dynamic results, which differed across the instruments. A higher spectral flatness value produced a flatter sound in the snare drum. Notably, the snare drum was consistently lower in amplitude at higher spectral flatness values across the demarcated zones. Similarly, the amplitude of the hi-hat changed with increased spectral flatness. Samples with the greatest spectral flatness had a characteristic "closed hi-hat" sound, indicating that the vibrational interaction with the bottom cymbal caused greater spectral flatness. The decrease in amplitude, with increased spectral flatness, indicated a peak strike-strength equivalent to the maximum amplitude of the vibrational interaction. This closely resembled white noise compared to louder strikes and suggests that strikes with attack transients, greater in amplitude than the vibrational interaction of the two cymbal's decay, produce lower levels of spectral flatness.

The amplitude of the tom-toms behaved in much the same way as the hi-hat, with spectrally flatter samples having lower amplitudes. This is consistent with changes in spectral slope, which reduced the spectral flatness for increased strike strengths.

Conversely, the ride and crash cymbals were louder at higher spectral flatness levels. Stronger strikes excite a large number of non-linear chaotic frequencies very quickly. Therefore, in the case of the cymbals, this is most likely attributed to the chosen method of calculating averages in the MIRToolbox *mirflatness* analysis algorithm, whereby frequencies are measured as an average over the duration of the sample. Further research might investigate the suitability of this algorithm, on percussion instruments, for the purposes of producing more accurate representations of spectral flatness at different strike strengths. Further research might also investigate the compositional possibilities of combining the inverse amplitude mappings (of the ride and crash cymbals) with the inherent secondary timbral features.

The spectral centroid parameter, with similar MIDI velocity mappings, produced some compelling compositional possibilities. At the same time, it produced unsurprising acoustical results. In certain instances, the spectral centroid parameter behaved in much the same way as the spectral flatness parameter, particularly in the case of the snare drum. In other instances, however, the spectral centroid parameter produced results that differed to spectral flatness. For example, the hi-hat tended to produce more "open" sounds at maximum flatness. Nevertheless, this was expected and was most likely caused by the increase in the number of frequencies excited, and the longer decay times, which corresponded to an increase in strike strength, resulting in a higher proportion of spectral energy in frequencies higher than the spectral centroid cut-off. This is also true of the crash and ride cymbals. The tom-toms displayed similar amplitude characteristics, to that of spectral flatness, although the samples had a brighter decay, with a higher spectral centre of gravity. Moreover, as the spectral centre of gravity increased, there was a small, but noticeable, change in pitch. A further area of work might repeat the sample database analysis, using a different threshold level than the 1500Hz default setting in MIRToolbox.

6.2 Performance Capture and Analysis

The second part of the approach involved performance capture and analysis. Human percussive performance information was captured from three drummers. This information was then analysed in order to assess the relationship between performer skill level, movement, performance variation, and timing variation during a performance, in order to develop a methodology with which to create an effective micro-timing model.

The recruitment and selection of participants took place over a six-week period. The selection process was conducted at The University of Sydney, and was open to both staff and students of the university.

Skill level is an important part of percussive performance. Therefore, three differently skilled drummers (unskilled, semi-skilled, and skilled) were chosen to participate in this study. In addition, the differently skilled drummers were selected with a view to investigating the ways in which skill development manifests empirically. However, this approach was limited given the small number of participants in the study, and given the subjective nature of judging performance skill. Further research might employ a larger number of participant performances, within each skill level, in order to fully explore characteristic traits. This thesis examined the behavioural and cognitive aspects of carrying out a physical movement, and the physical aspects of carrying out and coordinating a movement (procedural implementation) in Table 3.2. However, further research could be augmented with studies in the areas of human movement, and performance learning and development, not covered in Chapter Three.

The next phase of the approach involved the collection of data from real human performance in order to construct timing fingerprints, and to generate rules for the implementation of the model. This was done using captured audio, video, and accelerometer data from the performances of each of the three differently skilled drummers. Notably, the audio, video, and accelerometer data was crucial in understanding the effect of human movement as a contributory factor in timing variations.

In the first part of the methodological approach, timbral data was collected from the instruments of a nine-piece jazz drum set and incorporated into a sample database. As a result, it was not necessary to do an individual instrument capture during the real performances because only the timing values relative to the reference audio were required. The audio was collected using the same technical set-up used in the collection of the initial sample database, which minimised the amount of data collected. Whilst this method captured all of the performance strikes, the approach was limited. Capturing each specific instrument separately would enhance the contextual timing variations. Further research, therefore, might focus on improving the timing fingerprints.

In order to identify the effect of movement on performance timing, the performances were captured using overhead and side video cameras. Accelerometers were used to capture the roll and tilt of the participants' hands during performance. The video was then analysed using a Jitter algorithm, which compared inter-frame differences in movement. Taking video recordings of the performances was beneficial in two ways. Firstly, it enabled the empirical analysis of movement. Secondly, it afforded an opportunity to make observations on the participants' performances and, more generally, on percussive performance, in order to develop performance rules for the model.

The video analysis algorithm worked as expected, albeit with some limitations. Firstly, the algorithm did not discriminate between participant movement and other movement in the frame, because it was designed to analyse frame-by-frame video movement. This limitation was particularly important in the case of the semi-skilled drummer, whose drum strikes were generally stronger than those of the other two drummers. The stronger drum strikes caused the drum kit to move, on occasion, which led to an increase in movement levels in the video. Secondly, each of the participants struck different instruments of the drums a different number of times. For example, the crash cymbal moves more when struck, compared to a snare drum, resulting in greater movement per strike. At the same time, more participant movement was required to strike a crash cymbal. Thus, cymbal movement was more exaggerated in certain instances. Thirdly, additional factors, such as clothing and jewellery, contributed towards movement levels. These factors might be considered "sensor noise", however they had a minimal effect on the results. Further research might attempt to address each of these three limitations.

Accelerometer data were collected from each performance. Importantly, the empirical sensor data provided information from which a series of rules were created to infer performance context. In addition, the sensor data was particularly useful in identifying the tilt and roll of each of the participants' hands. The level of biomechanical movement captured in this investigation presents insights into the differing stick techniques and control of differently skilled performers. However, a greater understanding of the biomechanical aspects of performance is required to create rules that take into account intra-instrument level performance variations, compared to the inter-instrument level variations in this model. Further research might analyse percussive performance using analysis techniques available in the sports sciences, such as telemetric EMG systems, and motion analysis systems. This could lead to an improved understanding of performance injury to percussionists.

In the experiment, each of the drummers was asked to play along to a single reference track, *Bird's Lament* by Moondog. Once the fingerprint data was captured from the three performances, it was analysed in order to assess the relationship between performer skill level, movement, performance variation, and timing variation during a performance. For the purposes of this research, the decision to use a single track provided sufficient timing variations upon which to base a compositional tool that could be applied to different compositional styles, such as the one presented in this thesis. Notably, Sonic Visualiser was a particularly useful and accurate tool for analysing the performance data, and extracting and generating the timing fingerprints to be used in the model. However, this part of the methodology was very time-consuming. Further research might investigate ways of creating a more automated and efficient analysis methodology. Nevertheless, the timing fingerprint data did produce valuable information, which the drummers could use to critique performance with a view to making improvements.

The decision to use data based on a single performance, in order generate timing fingerprints representative of each individual drummer, presented an interesting paradox. On one hand, the re-contextualisation of raw performance data across different musical genres extended the performance information into new conceptual realms for compositional purposes. On the other hand, performance data based on a single jazz piece had specific genre constraints, owing to differences in performance technique and style across genres. Further research might investigate whether multiple cross-genre timing fingerprints can produce a more generic, stylistic performance fingerprint, and identify genre-neutral temporal features. This, however, would require an increase in sensor resolution, which, in itself, constitutes another area of research.

6.3 Performance Model

The third part of the approach comprised the computational design of the performance model, which included five methodological features (performance rules, physical constraints, instrumental configuration and movement levels, performance weightings, and timing fingerprints), and the implementation of the performance model.

The main aim of this thesis was to create a compositional software tool that simulates human percussive performance variation. As part of this investigation, a unique methodological framework was developed which linked disparate timing and timbral information with performance-driven rules. The compositional approach relied primarily on three feature-based parameters, one of which was "loudness". The loudness parameter was particularly useful for recreating performance and evaluating the model. With that in mind, the stochastic approach to sample selection, that is, the demarcation of performance zones in the context of movement in human percussive performance, produced timbral variations, which, in many cases, were consistent with performance complexity. However, in many instances these timbral variations were inconsistent with typical motivic dynamic changes. This can be attributed to the unequal distribution of samples within each demarcated zone. Notably, the sample selection algorithm, based upon current and first-order instrument selection, worked as expected, with greater inferred movement complexity producing greater timbral variations across the instruments. This was particularly relevant in the case of the membranophones because timbral variation in the idiophones was limited across the demarcation zones.

Performance rules, based on the physical aspects of human performance, and the physical nature of the interaction, were the first of five methodological features of computational design. These rules were generated from the analysis of three different types of data (the physical constraints of the performer, the physical constraints of the instrument configuration, and the context of simultaneous movement), which was extracted from participant performances. Notably, the analysis of these three types of data made a significant contribution to timbral and temporal variation.

Modelling the physical constraints of the performer benefited and limited the performance model to different degrees. For example, to replicate human performance and, in particular, the physical constraints of the performer, the number of strikes was limited to four (two each for the hands and feet). As a result, when it came to programming, having a smaller number of instrument combinations made it easier to implement the model, and it decreased the computational overhead by reducing the control stream. However, this approach was limited compositionally because the choice of simultaneously playable instruments for new music was significantly reduced. This was particularly relevant for meso-periodic music, where each instrument is considered a separate entity, played by multiple hands. To overcome this limitation in the study, three banks were created. This allowed for more instruments to be played simultaneously. It also facilitated the selection of instrumental combinations, from each bank, which allowed the user to consider the effect of movement level determined by instrumental configuration.

Modelling the physical constraints of the instrumental configuration was based upon relative height from the floor (high drums, medium drums, and low drums). The physical constraints also considered biomechanics and posture control. This approach made allowances for different drum configurations by assuming generic properties for each instrument. Assigning movement levels to different combinations of instruments made it possible to simulate strike accuracy through modal demarcations. Moreover, abstracting the combinations of drums being played also allowed for sequences of combinations to be evaluated in real-time for overall first-order movement. This dynamic, performance-context driven approach minimised timbral variation between sequences by modifying the probability of timbral variation. By varying the probability of timbral variation between the representations of skill levels, the perception of performance error was manipulated through timbre.

Different performance weightings, for zone selection, were applied to each of the three drum configurations, or movement levels discussed above (high drum, medium drum, and low drum). This approach directly linked timbral variation to performance context. One example of this is Study No. 6, which contained fewer timbral variations due to the improved consistency of the performance (skilled drummer). However, further research might undertake a series of listening tests to identify the efficacy of timbral and temporal variations (timing fingerprints) inherent in the model, and validate the performance model.

Timing fingerprints were another important methodological feature of the computational design. However, one disadvantage of this approach was the choice of temporal variation values. The model selects values without differentiating between instruments (e.g. a crash cymbal or a snare drum), even though values with a high inter-onset tempo are, statistically, more likely to be the direct result of a faster sequence, or an accenting fill at the end of a bar, which often contains a crash. Consequently, the timing fingerprint fails to contextualise the role of the instrument with the timing variations. Further research might investigate the ways in which timing fingerprint data can be more relevant for real instrument variations.

From a temporal perspective, the timing fingerprints were stochastically weighted towards specific temporal locations, relative to the beat, in the reference track. This was particularly the case for the unskilled and skilled drummers, who were more consistent in their temporal accuracy. The result was a larger percentage of sub $+/-20$ms asynchronies, which are difficult for a listener to detect in light of the complexity added by discrete timbral events. Further research might undertake a series of listening tests in which participants are asked to identify temporal asynchronies in a short excerpt produced using the compositional tool.

Further research might also investigate the lack of structural music context in larger temporal values. Larger temporal variations were more noticeable because the temporal variation values were selected stochastically. Additional rules, which infer the effect of structural elements of percussive performance variation, similar to that of the KTH rule systems and GERM model rules, might address this issue.

6.4 Software Implementation

The fourth part of the approach comprised the implementation of the software. One of the main aims of this thesis was to create a piece of software that ensured a level of human interaction that facilitates compositional applicability across music genres.

To do so, one important consideration was in the design philosophy. Performance variation was assigned to the computer in order to ensure the user maintains compositional control. In addition, other features were added to the software in order to provide the user with more compositional control. This included a mixer, rewire integration, automation, velocity curve control, and reverb and delay effects. All of these features combined to provide a composer with a stable, useable, aesthetically pleasing compositional tool that successfully simulates human percussive performance.

The most effective way to link the sample paradigm, with compositional paradigm was by using the TUBS notation system, particularly in the first and zero order selection, and the calculation of the movement levels. Although the TUBS notation is not unique to percussion-based software- and hardware-tools, this approach was unique insofar as the maximum cycle length was twelve, rather than the standard sixteen. This cycle length contributed to the creation of complex rhythms such as meso-periodic rhythm. Further research might investigate different graphical methods for representing sound events and samples, and might include the auto-generation of percussive material from additional sources.

The time-sensitive nature of the software, together with the need for user-stability, meant that it was necessary to ensure sufficient computational overhead at runtime. One of the main constraints, in using samples, was the need to load the samples into memory for playback. This constraint reduced the number of samples loaded in each demarcated zone. Despite this reduction, each instrument was represented by 125 samples. This resulted in significant timbral variety. Increased computer memory would allow this number to be increased, to a maximum of 127 samples per demarcated zone, which is a total of 635 samples per instrument. Further increases would allow for an increase in the number of demarcated zones, and in the number of instrument samples.

The time-sensitive nature of the software constrained the implementation of temporal variation, where timing variations were created by adding/subtracting the selected timing value in the fingerprint from a default latency of 100ms built into the flow of the sample trigger messages, in two ways. Firstly, from a performance perspective, the effect of a temporal variation, greater than an IOI of the beat, resulted in a "skipped" beat. Secondly, large variations tended to queue messages, and resulted in the delayed performance of simultaneous and concurrent samples.

The first constraint was addressed by scaling the selected variations from the fingerprint to different beat levels. Whilst this approach made assumptions concerning the relationship between variations and tempo, the effect was not noticeable when changing to adjacent tempi, for example from 119BPM to 120BPM. The second constraint was addressed by multithreading the sample trigger messages. However, as the system tempo and complexity increases, messages were still queued, causing late and missed notes. Further research might address this issue by increasing the number of threads in the system in order to allow for greater thread redundancy, as well as increased system performance at higher tempi and with greater complexity. Further work might also identify the impact of small asynchronies, in the context of computational software performance within time-critical applications, by benchmarking system performance.

6.5 Compositional Implementation

The fifth part of the approach comprised the compositional implementation of the software tool, which considered 12-step TUBS representation, rhythmic cycles, compositional timbre parameters, and minimum system requirements. The composition portfolio and accompanying analytical notes are intended to demonstrate the compositional implementation and musical applications of the human performance variation software tool.

The compositional framework of this study was, for the most part, derived from the meso-periodic rhythms of Africa. It was important, therefore, that the interface between the composer and the computer program was conducive to composition, particularly of other musical genres. The compositional interface used in this study was based upon the 12-step TUBS model, discussed in Chapter Four. This interface enabled the recreation of meso-periodic rhythms, as demonstrated in the composition portfolio. In addition, the grid-system in the TUBS model allowed for the creation of complex rhythms. It also allowed for the user to choose multiple simultaneous compositional parameters. These parameters (loudness, spectral flatness, and spectral centroid) produced interesting aesthetic results, which are demonstrated in the composition portfolio. It should be noted, however, that the use of parameters, other than loudness, caused sequences of timbral and dynamic variations, thereby reducing the efficacy of the performance model. This also occurred when combinations of compositional parameters were used simultaneously. However, it produced collateral compositional benefits insofar as the system produced uniquely varied timbral and dynamic effects, which enabled an exploration of contemporary percussive composition. Further work could include representing timbre multidimensionally, with a view to developing a timbre space for controlling both loudness and timbral parameters.

The broad bandwidth paradigm presented in this work was intended as a way of evaluating melody, rhythm, and spectral form (see Chapter Three). Indeed, this paradigm provided a useful framework upon which to visualise the melodic, rhythmic, and overall spectral form of Study No. 1 and No. 2, which are included in the composition portfolio.

Three compositional timbre parameters (loudness, spectral flatness, and spectral centroid) were used in this study. They were intended to simulate independency with meso-periodic rhythm, and to extend the compositional application of the model. These parameters produced some interesting compositional results. Firstly, the use of spectral flatness as a compositional parameter, with the tom-toms, produced unexpected discordancy and tension, which can be attributed to changes in amplitude. Study No. 3, in the composition portfolio, demonstrates this in particular. Secondly, the use of spectral centroid as a compositional parameter, and, at the same time, varying the levels of centroid hierarchically, produced a piece that was multi-layered (by perceived brightness). Study No. 4, in the composition portfolio, demonstrates this. In Study No. 9, spectral flatness conveyed the greatest sense of randomness in both temporal and timbral variation, compared to just timbral variation using spectral centroid. Further research might investigate the ways in which the current model can be extended into new compositional realms. Further research might, for example, consider replacing the percussive samples with non-musical sounds. Such research could use the same analysis paradigms and compositional parameters, or it might extend the compositional parameters to include other spectral features.

6.6 Summary Evaluation

In conclusion, this thesis presents a compositional tool that simulates human percussive performance variation. This compositional tool has been successfully applied to different genres of music to create rich and diverse rhythms, as demonstrated by the composition portfolio. This research makes several contributions to the fields of performance modelling, computer music, and percussive composition. It is of particular relevance to composers, the music-software community, percussive performers, and performance analysts. This research also highlights several areas for further work, which will be of interest to both the academic community and to industry.

Chapter 7

Appendices

"**W**hen words leave off, music begins."
– Heinrich Heine

A The PD-103 Software DVD

1 The PD-103 Software Installer Package

2 PD-103 Readme

3 Additional Software Components

B Data DVD

1 Text File Outputs of the MIRToolbox Feature Extraction

2 Questionnaires and Responses of the Selected Participants

3 Audio Reference Material: Moondog's *Bird's Lament*

4 The Max/MSP Video Analysis Patch

5 The Max/MSP Accelerometer Analysis Patch

6 Participant Video: Unskilled Overhead

7 Participant Video: Semi-skilled Overhead

8 Participant Video: Skilled Overhead

9 Java Code for the *MultiSwitch* External

10 Java Code for the *Fuzzy* External

11 Java Code for the *FuzzyAggregated* External

12 Java Code for the *PerformanceWeightings* External

13 Java Code for the *Velocity Curve* External

14 The Compositional Portfolio as .Wav Files

C Audio CD

1 Study No 1: African Meso-Periodic (I) (4:27)

2 Study No 2: African Meso-Periodic (II) (4:50)

3 Study No 3: Pitch Variation in the Tom-Toms (3:37)

4 Study No 4: Spectral Centroid and Isomorphic Rhythm (5:03)

5 Study No 5: Quark (1:00)

6 Study No 6: Demonstration of Improvisational Application (9:18)

7 Study No 7: Electroacoustic Piece (10:00)

8 Study No 8: Demonstration of Skill Levels (2:40)

9 Study No 9: Demonstration of Parametric Variation 3 x (1:45)

Total Running Time: (43:55)

References

Abend, W., Bizzi, E., and Morasso, P. (1982). Human arm trajectory formation. *Brain*, 105(2):331–348. 76, 77, 111, 173

Abernethy, B., Hanrahan, S. J., Kippers, V., Mackinnon, L. T., and Pandy, M. G. (2005). *The biophysical foundations of human movement* (2nd ed.). South Yarra: Palgrave Macmillan. 63

Adrien, J. M. (1991). The missing link: Modal synthesis. In De Poli et al., 1991 (Eds.). *Representations of musical signals* (pp. 269–298). Cambridge, MA: MIT Press. 4

Agawu, V. K. (2003). *Representing African Music: Postcolonial Notes, Queries, Positions.* New York, NY: Routledge. 93, 94

Aldridge, J. (1994). *Guide to Vintage Drums.* Anaheim Hills, CA: Centerstream Publications. 19, 49

Alén, O. (1995). Rhythm as duration of sounds in Tumba Francesa. *Ethnomusicology*, 39(1):55–71. doi:10.2307/852200. 75

Altenmüller, E., Wiesendanger, M., and Kesselring, J. (Eds.). (2006). *Music, Motor Control and the Brain.* Oxford, England: Oxford University Press. 261

Ames, C. (1987). Automated composition in retrospect: 1956-1986. *Leonardo*, 20(2):169–185. doi:10.2307/1578334. 82

Ames, C. (1990). Statistics and compositional balance. *Perspectives of New Music*, 28(1):80–111. doi:10.2307/833345. 31

Ames, C. (1992). Quantifying musical merit. *Journal of New Music Research*, 21(1):53–93. doi:10.1080/09298219208570600. 80

Antonelli, S. G. (2010). A Quantitative Study of the Quasistatic Tensioning Process of a Drumhead. Master's thesis, Cooper Union, Albert Nekern School of Engineering, New York, USA. Retrieved from ProQuest. (847871004). 47

Aramaki, M., Kronland-Martinet, R., Voinier, T., and Ystad, S. (2006). A percussive sound synthesizer based on physical and perceptual attributes. *Computer Music Journal*, 30(2):32–41. 4, 51, 125

Arom, S. (2004). *African Polyphony and Polyrhythm*. Cambridge, England: Cambridge University Press. 92, 93, 94

Aruin, A. and Shiratori, T. (2003). Anticipatory postural adjustments while sitting: the effects of different leg supports. *Experimental Brain Research*, 151(1):46–53. doi:10.1007/s00221–003–1456–y. 75

Aruin, A. S. and Latash, M. L. (1995). Directional specificity of postural muscles in feed-forward postural reactions during fast voluntary arm movements. *Experimental Brain Research*, 103(2):323–332. doi:10.1007/BF00231718. 74, 75

Avanzini, F. and Marogna, R. (2010). A Modular Physically Based Approach to the Sound Synthesis of Membrane Percussion Instruments. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(4):891–902. doi:10.1109/TASL.2009.2036903. 4

Barthet, M., Kronland-Martinet, R., and Ystad, S. (2008). Improving musical expressiveness by time-varying brightness shaping. In Martinet et al. (Eds.), *Computer Music Modeling and Retrieval, Sense of Sounds*, (pp. 313–336). New York: Springer Berlin Heidelberg. 125

Beauchamp, J. W. (2010). *Analysis, Synthesis, and Perception of Musical Sounds*. The Sound of Music. New York: Springer. 3, 29, 30, 50, 247

Bělohlávek, R. and Klir, G. J. (2011). *Concepts and Fuzzy Logic*. Cambridge, MA: MIT Press. 18

Benadon, F. (2006). Slicing the beat: Jazz eighth-notes as expressive microrhythm. *Ethnomusicology*, 50(1):73–98. 16

Benson, D. J. (2007). *Music: A Mathematical Offering*. Cambridge, England: Cambridge University Press. 41

Berg, R. E. and Stork, D. G. (1982). *The Physics of Sound*. Englewood Cliffs, NJ: Prentice Hall. 40, 41

Bergmann, M. (2008). *An Introduction to Many-Valued and Fuzzy Logic: Semantics, Algebras, and Derivation Systems*. Cambridge, England: Cambridge University Press. 18

Bernstein, N. (1967). The co-ordination and regulation of movements. Oxford, England: Pergamon Press Ltd. 58, 63

Bilbao, S. (2009). A modular percussion synthesis environment. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09), 1st-4th September, 2009*, Como, Italy. 4

Bilbao, S. (2010). Percussion synthesis based on models of nonlinear shell vibration. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(4):872–880. doi:10.1109/TASL.2009.2029710. 4, 53

Bilbao, S. (2012). Time domain simulation and sound synthesis for the snare drum. *The Journal of the Acoustical Society of America*, 131:914–925. doi:10.1121/1.3651240. 3, 4, 16

Billon, M., Semjen, A., and Stelmach, G. E. (1996). The timing effects of accent production in periodic finger-tapping sequences. *Journal of Motor Behavior*, 28(3):198–210. doi:10.1080/00222895.1996.9941745. 64

Bilmes, J. A. (1993a). Techniques to foster drum machine expressivity. In *Proceedings of the International Computer Music Conference (ICMC)* (pp. 276–283). Retrieved from: http://ssli.ee.washington.edu/~bilmes/mypubs/bilmes1993-drum.pdf. 3, 4, 5

Bilmes, J. A. (1993b). Timing is of the essence: Perceptual and computational techniques for representing, learning, and reproducing expressive timing in percussive rhythm. Master's thesis, Massachussets Institute of Technology. Retrieved from: http://melodi.ee.washington.edu/~bilmes/mypubs/bilmes1993-mit-thesis.pdf. 20, 21

Black, D. (2003). *The Drummer's Toolkit*. Van Nuys, CA: Alfred Publishing Co. Inc. 44, 45, 47, 49

Blades, J. (1992). *Percussion Instruments and their History* (4th ed.). Westport, CT: Bold Strummer Ltd. 34

Blaine, T. and Perkis, T. (2000). The Jam-O-Drum Interactive Music System: A Study in Interaction Design. In *Proceedings of the conference on Designing interactive systems processes, practices, methods, and techniques - DIS '00* (pp. 165–173), New York, USA. doi:10.1145/347642.347705. 5, 23, 24

Boersma, P. and Weenink, D. (2013). Praat: doing phonetics by computer. [Computer program]. Available from: http://www.praat.org. 123

Bongers, B. (2000). Physical interfaces in the electronic arts. *Trends in Gestural Control of Music* (pp. 41–70). Available from: http://hapticity.net/pdf/nime2005_192-works_cited/10.1.1.90.4997.pdf. 101

Bork, I. (1983). Entwicklung von akustischen Optimierungsverfahren für Stabspiele und Membraninstrumente. *Development of acoustical optimization methods for musical bars and membrane instruments. PTB Report, Projekt*, 5267. 39

Bowman, F. (1958). *Introduction to Bessel Functions*. New York, NY: Dover Publications. 46

Brandmeyer, A., Timmers, R., Sadakata, M., and Desain, P. (2011). Learning expressive percussion performance under different visual feedback conditions. *Psychological Research*, 75(2):107–121. doi:10.1007/s00426–010–0291–6. 56, 63

Brent, W. (2009). Cepstral analysis tools for percussive timbre identification. In *Proceedings of the 3rd International Pure Data Convention*, Sao Paulo, Brazil. 126

Britto Jr, A. D. S., Gouyon, F., and Dixon, S. (Eds.). (2013). *The International Society for Music Information Retrieval. Conferences, Publications and Related Activities*. ISMIR, Curitiba, Brazil. 20

Budón, O. (2000). Composing with objects, networks, and time scales: an interview with Horacio Vaggione. *Computer Music Journal*, 24(3):9–22. 86

Butler, D. (1992). *The Musician's Guide to Perception and Cognition*. New York, NY: Schirmer Books. 50

Butterfield, M. W. (2010). Participatory Discrepancies and the Perception of Beats in Jazz. *Music Perception*, 27(3):157–176. doi:10.1525/MP.2010.27.3.157. 135

Butterfield, M. W. (2011). Why Do Jazz Musicians Swing Their Eighth Notes? *Music Theory Spectrum*, 33(1):3–26. doi:10.1525/mts.2011.33.1.3. 16

Calvin, S., Huys, R., and Jirsa, V. K. (2010). Interference effects in bimanual coordination are independent of movement type. *Journal of Experimental Psychology: Human Perception and Performance*, 36(6):1553–1564. doi:10.1037/a0020403. 63

Cannam, C., Landone, C., Sandler, M. B., and Bello, J. P. (2006). The Sonic Visualiser: A Visualisation Platform for Semantic Descriptors from Musical Signals. In Reiss, Joshua D. and Wiggins, Geraint A. (Eds.) *Proceedings of the 6th International Conference on Music Information Retrieval*, Victoria, Canada. 134

Carson, R. and Wanamaker, J. (1984). *International Drum Rudiments.* Van Nuys, CA: Alfred Publishing Co. Inc. 56, 57, 58, 60, 61

Chadwick, E. K. and van der Helm, F. (2003). Shoulder biomechanics. In *International Society of Biomechanics 2003 Congress XIX* (pp. 1–30). Dunedin, New Zealand. 79

Charbonneau, G. R. (1981). Timbre and the perceptual effects of three types of data reduction. *Computer Music Journal* (pp. 10–19). doi:10.2307/3679875. 50

Chen, G. and Pham, T. T. (2000). *Introduction to Fuzzy Sets, Fuzzy Logic, and Fuzzy Control Systems.* Boca Raton, FL: CRC Press. 18

Chion, M. (2009). *Guide des Objets Sonores.* (Dack, John and North, Christine, Trans.). Paris, France: Buchet/Chastel. (Original work published 1983). 51

Chowning, J. (2000). Digital sound synthesis, acoustics and perception: A rich intersection. In *Proceedings of the COST G-6 Conference on Digital Audio Effects*, Verona, Italy. 50

Christensen, M. G. and Jakobsson, A. (2009). *Multi-Pitch Estimation.* Synthesis Lectures on Speech and Audio Processing. Vol. 5, No. 1. (pp.1–160). Morgan & Claypool Publishers. doi:10.2200/S00178ED1V01Y200903SAP005. 46

Colton, S. (2008). Creativity Versus the Perception of Creativity in Computational Systems. In Ventura, D., Maher, M. L., and Colton, S. (Eds.). *Proceedings of the AAAI Spring Symposium on Creative Systems*, AAAI. Stanford, CA, USA. 80

Cook, P. R. (1997). Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal*, 21(3):38–49. 4, 104

Cook, P. R. (2002a). *Real sound synthesis for interactive applications.* Natick, MA: A. K. Peters Ltd. 50

Cook, P. R. (2002b). Toward physically-informed parametric synthesis of sound effects. In *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA-99),* (pp. 1–5), New Paltz, NY, USA. 104, 105

Cooper, G. and Meyer, L. B. (1963). *The Rhythmic Structure of Music*. University of Chicago Press. 87, 90, 184

Dahl, S. (1997a). Measurements of the motion of the hand and drumstick in a drumming sequence with interleaved accented strokes: A pilot study. In *Speech, Music and Hearing, Quarterly Progress and Status Report*. Vol. 4/1997, (pp. 1–6). Royal Institute of Technology, Stockholm. 66, 67, 68, 72, 73, 74, 110

Dahl, S. (1997b). Spectral changes in the tom-tom related to striking force. In *Speech, Music and Hearing, Quarterly Progress and Status Report*. Vol. 38, (pp. 59–66). Royal Institute of Technology, Stockholm. 36, 40, 64, 65, 125

Dahl, S. (2000). The Playing of an Accent - Preliminary Observations from Temporal and Kinematic Analysis of Percussionists. *Journal of New Music Research*, 29(3):225–234. 4, 15, 18, 68, 69, 71, 72, 77

Dahl, S. (2003). Striking movements: Movement strategies and expression in percussive playing. Master's thesis, Royal Institute of Technology, Stockholm, Department of Speech, Music, and Hearing, Available from: http://www.speech.kth.se/~sofiad/pdf/StrikingMovements.pdf. 66

Dahl, S. (2004). Playing the accent - comparing striking velocity and timing in an ostinato rhythm performed by four drummers. *ACTA Acustica united with Acustica*, 90(4):762–776. 65

Dahl, S. (2005). *On the beat: Human movement and timing in the production and perception of music*. Doctoral Thesis. 57, 65

Dahl, S. and Altenmüller, E. (2008). Motor Control in drumming: Influence of movement pattern on contact force and sound characteristics. In *Proceedings of Acoustics 2008* (pp. 1489–1494), Available from: http://vbn.aau.dk/files/18250599/DahlAltenmueller-Acoustics08. 3, 66

Dahl, S. and Altenmüller, E. (2013). Movements and grip influence the contact force and sound in drumming. *Acoustical Science of America ASA/EAA/SFA '08 Meeting Lay*. 3, 66

Dahl, S. and Bresin, R. (2001). Is the player more influenced by the auditory than the tactile feedback from the instrument? In Fernström, M., Brazil, E., and Marshall, M., (Eds.). *Proceedings of the COST-G6 Workshop on Digital Audio Effects* (pp. 194–197), Limerick, Ireland. 56, 63

Dahl, S. and Friberg, A. (2004). Expressiveness of musician's body movements in performances on marimba. In Camurri, A., and Volpe, G. (Eds.). *Gesture-based Communication in Human-Computer Interaction, Lecture Notes in Artificial Intelligence.* Vol. 2915. (pp. 479–486). New York: Springer. 55

Dahl, S., Granqvist, S., and Thomasson, M. (2000). Detection of drift in tempo. In *Detection of tempo in drift, report.* Vol. 41, No. 4. (pp. 19–28). Royal Institute of Technology, Stockholm. Available from: http://www.speech.kth.se/prod/publications/files/qpsr/2000/2000_41_4_019-028.pdf. 18

Dahl, S., Grossbach, M., and Altenmüller, E. (2010). Timing and motor control in drumming: Comparing players with and without focal dystonia. In *Abstract from Rhythm Production and Perception Workshop, Lille France.* 3

Dahl, S., Großbach, M., and Altenmüller, E. (2011). Effect of dynamic level in drumming: Measurement of striking velocity, force, and sound level. In Danish Acoustical Society (DAS), (Ed.). In *Proceedings of Forum Acusticum,* (pp. 621–624), Aalborg Congress & Culture Center, Aalborg, Denmark. (CD-ROM). 55, 77

D'Angelo (1999). Left & Right. [Recorded by M. E. Archer] On *Voodoo* [CD], New York, NY: Virgin Records America Inc. 88

Danielsen, A. (2010). *Musical rhythm in the age of digital reproduction.* Farnham, England: Ashgate Publishing. 87, 88, 91

Darke, G. (2005). Assessment of timbre using verbal attributes. In *Proceedings of the conference on Interdisciplinary Musicology. Montreal, Quebec.* Available from: http://www.oicrm.org/doc/2005/cim05/articles/darke_g_cim05.pdf. 125

de Paula, H. B., Loureiro, M. A., and Yehia, H. C. (2004). Timbre Representation of a Single Musical Instrument. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)* (pp. 546–549), Barcelona, Spain. 105

De Poli, G., Piccialli, A., and Roads, C. (1991). *Representations of musical signals.* Cambridge, MA: MIT Press. 240

Dean, R. T. (2009). *The Oxford Handbook of Computer Music*. New York, NY: Oxford University Press. 20, 86

Decker, S., Kendall, G., Schmidt, B., Ludwig, M., and Freed, D. (1986). A Modular Environment for Sound Synthesis and Composition. *Computer Music Journal*, 10(4):28–41. 81

Desain, P. and Honing, H. (1993). Tempo curves considered harmful. *Contemporary Music Review*, 7(2):123–138. doi:10.1080/07494469300640081. 56, 182

Deutsch, D. (1999). *The Psychology of Music* (2nd ed.). San Diego, CA: Academic Press. 50, 250, 262

Dodson, A. (2011). Expressive Asynchrony in a Recording of Chopin's Prelude No. 6 in B Minor by Vladimir de Pachmann. *Music Theory Spectrum*, 33(1):59–64. 134

Dolhansky, B., McPherson, A., and Kim, Y. (2011). Designing an Expressive Virtual Percussion Instrument. In *Proceedings of the 8th Sound and Music Computing Conference.* Retrieved from: http://server.smcnetwork.org/system/files/smc2011_submission_71.pdf. 136

Donnadieu, S. (2010). *Mental representation of the timbre of complex sounds.* In Beauchamp, J. (Ed.). *Analysis, Synthesis, and Perception of Musical Sounds.* (pp. 272–319). New York, NY: Springer. 50, 125

Drummond, J. (2009). Understanding interactive systems. *Organised Sound*, 14(2):124–133. doi:10.1017/S1355771809000235. 22, 23

Dubnov, S. (2006). Spectral anticipations. *Computer Music Journal*, 30(2):63–83. 126

Eigenfeldt, A. (2006). Kinetic Engine: Toward an intelligent improvising instrument. In *Proceedings of the Sound and Music Computing Conference,* (pp. 97–100). Marseille, France. 25

Eigenfeldt, A. (2010). Real-Time Timbral Organisation: Selecting Samples Based Upon Similarity. *Organised Sound*, 15(2):159–166. doi:10.1017/S1355771810000154. 100

Eigenfeldt, A., Burnett, A., and Pasquier, P. (2012). Evaluating musical metacreation in a live performance context. In *Proceedings of the 3rd International Conference on Computational Creativity,* pp. 140-144). Dublin, Ireland. Retrieved from: http://computationalcreativity.net/iccc2012/wp-content/uploads/2012/05/140-Eigenfeldt.pdf. 80

Ekwueme, L. E. (1974). Concepts of African musical theory. *Journal of Black Studies*, 5(1):35–64. 94, 95

Feiten, B. and Günzel, S. (1994). Automatic indexing of a sound database using self-organizing neural nets. *Computer Music Journal*, 18(3):53–65. 102

Fletcher, N. H. (1975). Analysis and synthesis of bass drum tones. *The Journal of the Acoustical Society of America*, 58(S1):S131–S131. 3, 40

Fletcher, N. H. (1978). Some experiments with the bass drum. *The Journal of the Acoustical Society of America*, 64(6):1570–1576. doi:10.1121/1.382140. 3, 37

Fletcher, N. H. (1999). The nonlinear physics of musical instruments. *Reports on Progress in Physics*, 62(5):723–764. doi:10.1088/0034–4885/62/5/202. 3

Fletcher, N. H. and Rossing, T. D. (1998). *The Physics of Musical Instruments* (2nd ed.). New York, NY: Springer. 3, 10, 19, 36, 37, 39, 40, 41, 42, 45, 58, 148, 175

Fontana, F. and Rocchesso, D. (1996). Simulations of Membrane-based Percussion Instruments. In *Proc. Workshop: Sound Synthesis by Physical Modeling,* (Firenze, Italy), Centro Tempo Reale. 4

Fontana, F. and Rocchesso, D. (1998). Physical modeling of membranes for percussion instruments. *ACTA Acustica united with Acustica*, 84(3):529–542. 4

Fourcade, P. and Cadoz, C. (2001). Physical modeling for sound synthesis with particle-interaction networks: study of the percussion phenomena and conception of a generic module for percussion. In *Proceedings of the International Computer Music Conference (ICMC)*, La Habana, Cuba. Retrieved from: http://quod.lib.umich.edu/cgi/p/pod/dod-idx/physical-modeling-for-sound-synthesis-with-particle.pdf?c=icmc;idno=bbp2372.2001.005. 4

Friberg, A. (1991). Generative rules for music performance: A formal description of a rule system. *Computer Music Journal*, 15(2):56–71. 28

Friberg, A. (1995). Matching the rule parameters of PHRASE ARCH to performances of “Träumerei”: A preliminary study. In Friberg, A. and Sundberg, J. (Eds.), *Proceedings of the KTH symposium on Grammars for music performance,* (pp. 37–44). Stockholm, Sweden. 29

Friberg, A. (2006). pDM: An Expressive Sequencer with Real-Time Control of the KTH Music-Performance Rules. *Computer Music Journal*, 30(1):37–48. 29

Friberg, A., Bresin, R., Frydén, L., and Sundberg, J. (1998). Musical Punctuation on the Microlevel: Automatic Identification and Performance of Small Melodic Units. *Journal of New Music Research*, 27(3):271–292. doi:10.1080/09298219808570749. 29

Friberg, A., Bresin, R., and Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology*, 2(2–3):145–161. doi:10.2478/v10053–008–0052–x. 3, 5

Friberg, A. and Sundberg, J. (1999). Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners. *The Journal of the Acoustical Society of America*, 105(2):1469–1484. doi:0001–4966/99/105(3)/1469/16. 136

Friberg, A. and Sundstrom, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, 19(3):333–349. 16, 30, 135, 182, 183

Fujii, S., Hirashima, M., Kudo, K., Ohtsuki, T., Nakamura, Y., and Oda, S. (2010). Synchronization Error of Drum Kit Playing with a Metronome at Different Tempi by Professional Drummers. *Music Perception*, 28(1):491–503. doi:10.1525/MP.2010.28.1.491. 62, 63

Fujii, S., Kudo, K., Ohtsuki, T., and Oda, S. (2009). Tapping performance and underlying wrist muscle activity of non-drummers, drummers, and the world's fastest drummer. *Neuroscience Letters*, 459(2):69–73. doi:10.1016/j.neulet.2009.04.055. 55, 62

Fujii, S. and Moritani, T. (2012a). Rise rate and timing variability of surface electromyographic activity during rhythmic drumming movements in the world's fastest drummer. *Journal of Electromyography and Kinesiology*, 22(1):60–66. 10.1016/j.jelekin.2011.10.004. 62

Fujii, S. and Moritani, T. (2012b). Spike shape analysis of surface electromyographic activity in wrist flexor and extensor muscles of the world's fastest drummer. *Neuroscience Letters*, 514(2):185–188. doi:10.1016/j.neulet.2012.02.089. 62

Fujii, S. and Oda, S. (2009). Effect of stick use on rapid unimanual tapping in drummers. *Perceptual and Motor Skills*, 108:962–970. doi:10.2466/pms.108.3.962–970. 65

fxpansion (2005). *BFD Premium Acoustic Drum Module, Reference Manual, v.1.5.* fxpansion, Burbank, CA, USA. Retrieved from: http://www.fxpansion1.com/resourceUploads/BFD_Manual_English.pdf. 125

Gabrielsson, A. (1999). *The Performance of Music.* In Deutsch (Ed.). *The Psychology of Music* (2nd Ed.). (pp. 501–602). San Diego: Academic Press. 56, 63

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music*, 31(3):221–272. doi:10.1177/03057356030313002. 56, 62, 63

Garry, M. I. and Franks, I. M. (2000). Reaction time differences in spatially constrained bilateral and unilateral movements. *Experimental Brain Research*, 131(2):236–243. doi:10.1007/s002219900299. 78

Georgopoulos, A. P., Kalaska, J. F., and Massey, J. T. (1981). Spatial trajectories and reaction times of aimed movements: effects of practice, uncertainty, and change in target location. *Journal of Neurophysiology*, 46(4):725–743. 78

Giordano, B. L. and McAdams, S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America*, 119(2):1171–1181. doi:10.1121/1.2149839. 125

Godøy, R. I. and Leman, M. E. (2010). *Musical Gestures: Sound, Movement, and Meaning.* New York, NY: Routledge. 76, 251

Goebl, W. (2011). Temporarily out of sync: Momentary temporal independence of a solo voice as expressive device. *Forum Acusticum 2011, Aalborg, Denmark.* 56

Goebl, W., Dixon, S., de poli, G., Friberg, A., Bresin, R., and Widmer, G. (2005). *'Sense' in Expressive Music Performance: Data Acquisition, Computational Studies, and Models.* In Polotti, P and Rocchesso, D (Eds.). *Sound to Sense Sense - Sound: A state of the art in Sound and Music Computing* (pp. 195–242). Berlin, Germany: Logos Verlag. 135

Goldberg, R. and Riek, L. (2000). *A Practical Handbook of Speech Coders (Discrete Mathematics and Its Applications).* Boca Raton, FL: CRC Press. 101

Gouyon, F., Fabig, L., and Bonada, J. (2003). Rhythmic Expressiveness Transformations of Audio recordings: Swing Modifications. In *Proceedings of the 6th International Conference on Digital Audio Effects (DAFX-03),* (pp. 94–99). Queen Mary University, London, UK. 4, 5

Greenwald, J. (2002). Hip-Hop Drumming: The Rhyme May Define, but the Groove Makes You Move. *Black Music Research Journal*, 22(2):259–271. 102

Hajian, A. Z., Sanchez, D. S., and Howe, R. D. (1997). Drum roll: increasing bandwidth through passive impedance modulation. In *Proceedings of the 1997 IEEE International Conference on Robotics and Automation,* (pp. 2294–2299). Albuquerque, New Mexico, IEEE. 136

Hall, D. E. (1991). *Musical acoustics* (2nd ed.). Pacific Grove, CA: Brooks/Cole Publishing Company. 3, 35, 36, 37

Halmrast, T., Guettler, K., Bader, R., and Godøy, R. I. (2010). *Gesture and Timbre.* In Godøy and Leman (Eds.). *Musical Gestures: Sound, Movement, and Meaning.* (pp. 183–211). New York, NY: Routledge. 56

Hardin, L. T. (1969). Bird's Lament. [Recorded by L. T. Hardin] On *Moondog* [CD], New York, NY: Columbia Masterworks (1969). 136

Hartmann, W. M. (1997). *Signals, Sound, and Sensation (Modern Acoustics and Signal Processing)* (3rd ed.). New York, NY: Springer. 50

Hellmer, K. (2005). Rendering timbre variance and time deviance in MIDI-controlled contemporary drum performances. C-Uppsats. Luleå Tekniska Universitet. Retrieved from: http://epubl.ltu.se/1402-1773/2005/017/LTU-CUPP-05017-SE.pdf. 16, 17, 18, 181

Hellmer, K. (2006). The Development of a Drum Machine Using the Steinberg VST-Specification. Unpublished master's thesis, Luleå University of Technology, School of Music. Retrieved from: https://pure.ltu.se/portal/files/30988700/LTU-DUPP-06054-SE.pdf. 3, 4, 5, 18, 85, 121

Henzie, C. (1960). *Amplitude and duration characteristics of snare drum tones.* PhD thesis, School of Education, Indiana University. 38, 39, 56

Herrera, P., Dehamel, A., and Gouyon, F. (2003). Automatic labeling of unpitched percussion sounds. In *Proceedings of the Audio Engineering Society 114th Convention.* (pp. 1378–1391). Audio Engineering Society, Amsterdam, The Netherlands. Available from: http://www.aes.org/e-lib/browse.cfm?elib=12599. 102

Herrera, P., Yeterian, A., and Gouyon, F. (2002). Automatic classification of drum sounds: a comparison of feature selection methods and classification techniques. In

Anagnostopoulou, C. and Ferrand, M. (Eds.). *Proceedings of the 2nd International Conference on Music and Artificial Intelligence* LNAI 2445. (pp. 69–80), New York, NY: Springer-Verlag. 102, 103

Hiller, L. and Baker, R. (1964). Computer Cantata: A study in compositional method. *Perspectives of New Music*, 3(1):62–90. 19, 82, 83, 84, 85

Hiller, L., Dick, R., Post, N., MacDonald, R., Marder, N., Rosen, R. J., Ritscher, G., Myhill, J., Ames, C., Isaacson, L. M., and O'Connor, A. (1986). Computer music retrospective (1957-1985). [Sound Recording], Mainz, Germany: Wergo. 3, 84

Hiller, L. and Isaacson, L. M. (1957). Illiac suite, for string quartet. *New Music Edition,* Bryn Mawr, PA: Theodore Presser Company. 82

Hoffmann, P. (2000). The new GENDYN program. *Computer Music Journal*, 24(2):31–38. 3, 84

Hollerbach, J. M. and Flash, T. (1982). Dynamic Interactions Between Limb Segments During Planar Arm Movement. *Biological Cybernetics*, 44:67–77. 77

Holmes, G. (1939). The cerebellum of man. *Brain*, 62:1–30. 75

Honing, H. (2001). From Time to Time: The Representation of Timing and Tempo. *Computer Music Journal*, 25(3):50–61. 88

Horner, A. (2003). Auto-programmable FM and wavetable synthesizers. *Contemporary Music Review*, 22(3):21–29. doi:10.1080/0749446032000150852. 103

Hove, M., Keller, P., and Krumhansl, C. (2007). Sensorimotor synchronization with chords containing tone-onset asynchronies. *Attention, Perception, & Psychophysics*, 69(5):699–708. 62

Howard, D. and Angus, J. (2009). *Acoustics and Psychoacoustics* (4th ed.). Oxford, England: Focal Press. 50

Howe, H. S. (1972). Compositional Limitations of Electronic Music Synthesizers. *Perspectives of New Music*, 10(2):120–129. 14

Hsu, W. (2006). Managing gesture and timbre for analysis and instrument control in an interactive environment. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME06),* (pp. 376–379), Paris, France. 23

Hsu, W. (2008). Two approaches for interaction management in timbre-aware improvisation systems. In *Proceedings of the International Computer Music Conference (ICMC)*, Belfast, UK. 23

Huber, D. M. and Runstein, R. E. (2005). *Modern Recording Techniques* (6th ed.). Oxford, England: Focal Press. 37, 40, 41, 49, 101

Hunt, B. R., Lipsman, R. L., Rosenberg, J. M., Coombes, K. R., Osborn, J. E., and Stuck, G. J. (2001). *A Guide to MATLAB: For Beginners and Experienced Users* (1st ed.). Cambridge, England: Cambridge University Press. 123

Iannarilli, F., Vannozzi, G., Iosa, M., Pesce, C., and Capranica, L. (2013). Effects of task complexity on rhythmic reproduction performance in adults. *Human Movement Science*, 32:203–213. doi:10.1016/j.humov.2012.12.004. 63

Jaffe, D. A. (1995). Ten Criteria for Evaluating Synthesis Techniques. *Computer Music Journal*, 19(1):76–87. 50, 51, 104, 108

Jaffe, D. A. and Smith, III, J. O. (1983). Extensions of the Karplus-Strong plucked-string algorithm. *Computer Music Journal*, 7(2):56–69. 40

James, R. D. (2001). Vordhosbn. On *Drukqs* [CD], Warp Records, UK. 87

JazzCorner.com (2012). Jazz Corner News: Buddy Rich 25th Anniversary Memorial Concert. (17th February, 2012). Retrieved from: http://www.jazzcorner.com/news/display.php?news=2604. 79

Jones, A. M. (1959). *Studies in African Music, Volume 1*. London, England: Oxford University Press. 91, 94

Jordanous, A. (2011). Evaluating evaluation: Assessing progress in computational creativity research. In *Proceedings of the Second International Conference on Computational Creativity (ICCC 2011)*, (pp. 102–107). México City, México. 80

Juslin, P. N. (2000). Cue Utilization in Communication of Emotion in Music Performance: Relating Performance to Perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26(6):1797–1813. doi:10.1037//0096–1523.26.6.1797. 125

Juslin, P. N., Friberg, A., and Bresin, R. (2002). Toward a computational model of expression in music performance: The GERM model. *Musicae Scientiae,* Special Issue 2001–2002, 5:63–122. doi:10.1177/10298649020050S104. 3, 25, 29, 30, 31, 33

Kahrs, M. and Brandenburg, K. E. (2002). *Applications of digital signal processing to audio and acoustics.* New York, NY: Kluwer Academic Publishers. 101, 104, 105

Kartomi, M. J. (1990). *On Concepts and Classifications of Musical Instruments.* Chicago, IL: University Of Chicago Press. 41

Katayose, H., Hashida, M., de poli, G., and Hirata, K. (2012). On Evaluating Systems for Generating Expressive Music Performance: the Rencon Experience. *Journal of New Music Research*, 41(4):299–310. doi:10.1080/09298215.2012.745579. 19, 80

Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, 2(3):275–283. 8

Keil, C. (1995). The Theory of Participatory Discrepancies: A Progress Report. *Ethnomusicology,* Special Issue: Participatory Discrepancies, 39(1):1–19. 133

Kelso, J. A., Buchanan, J. J., and Wallace, S. A. (1991). Order parameters for the neural organization of single, multijoint limb movement patterns. *Experimental Brain Research*, 85:432–444. 73, 74, 77, 110

Kelso, J. A., Southard, D., and Goodman, D. (1979). On the nature of human interlimb coordination. *Science*, 203(4384):1029–1031. 63

Kilchenmann, L. and Senn, O. (2011). "Play in time, but don't play time": Analyzing timing profiles in drum performances. *International Symposium on Performance Science,* (pp. 593–598). Retrieved from: http://www.legacyweb.rcm.ac.uk/cache/fl0026853.pdf. 135

Kippen, J. and Bel, B. (1994). Computers, Composition and the Challenge of "New Music" in Modern India. *Leonardo Music Journal*, 4:79–84. 15

Kirke, A. and Miranda, E. R. (2009). A survey of computer systems for expressive music performance. *ACM Computing Surveys*, 42(1):1–41. doi:10.1145/1592451.1592454. 3, 4, 19, 33, 82

Kiusalaas, J. (2005). *Numerical Methods in Engineering with MATLAB.* Cambridge, England: Cambridge University Press. 123

Klatzky, R. L., van den Doel, K., and Krotkov, E. P. (2000). Perception of material from contact sounds. *Presence: Teleoperators and Virtual Environments*, 9(4):399–410. 51

Klingbeil, M. (2009). *Spectral analysis, editing, and resynthesis: Methods and applications*. PhD thesis, Columbia University, School of Arts and Sciences. Retreived from: http://www.klingbeil.com/data/Klingbeil_Dissertation_web.pdf. 101, 102

Koblick, J. (2007). *The Home Recording Studio Guide: Get The Pro Sound*. North Charleston, SC: CreateSpace. 39, 41

Koetting, J. (1970). Analysis and notation of west African drum ensemble music. *Selected Reports in Ethnomusicology*, 1(3):115–146. 91, 186, 187

Kubik, G. (2010). *Theory of African Music*. Vol. 1. Chicago, IL: University of Chicago Press. 94

Kunkler-Peck, A. J. and Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance*, 26(1):279–294. 51

Kurzweil, R. (1992). *The Age of Intelligent Machines*. Cambridge, Mass. : MIT Press. 267

Laird, J., Masri, P., and Canagarajah, C. (1998). Efficient and Accurate Synthesis of Circular Membranes using Digital Waveguides. In *Proceedings of the IEE Colloquium on Audio and Music Technology: The Challenge of Creative of DSP,* (pp. 12/1–12/6). IEE Digest 98/470. 4

Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception and Psychophysics*, 62(7):1426–1439. doi:10.3758/BF03212144. 4, 50

Lartillot, O. (2010). *MIRtoolbox* User's Manual *v*1.3. *Finnish Centre of Excellence in Interdisciplinary Music Research University of Jyväskylä, Finland.* Version 1.5 available from: https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/MIRtoolbox1.5Guide . 123, 124, 125

Lartillot, O. and Toiviainen, P. (2007). A Matlab Toolbox for Musical Feature Extraction From Audio. In *Proceedings of the 10th International Conference on Digital Audio Effects*, Bordeaux, France. 7, 121, 123

Lartillot, O., Toiviainen, P., and Eerola, T. (2008). *A matlab toolbox for music information retrieval*. In Preisach, C., Burkhardt, H., Schmidt-Thieme, L., and Decker, R., (Eds.). *Data analysis, machine learning and applications*. Freiburg, Germany: Springer Verlag. 7

Laske, O. (1989). Composition theory: An enrichment of music theory. *Interface*, 18(1-2):45–59. doi:10.1080/09298218908570537. 81

Laukka, P. and Gabrielsson, A. (2000). Emotional Expression in Drumming Performance. *Psychology of Music*, 28(2):181–189. doi:10.1177/0305735600282007. 55

Lee, S. (2010). Hand Biomechanics in Skilled Pianists Playing a Scale in Thirds. *Medical Problems of Performing Artists*, 25(4):167–174. 55

Legge, K. A. and Fletcher, N. H. (1989). Nonlinearity, chaos, and the sound of shallow gongs. *The Journal of the Acoustical Society of America*, 86(6):2439–2443. doi:10.1121/1.398451. 4, 28

Leonard, III, N. (1996). "Legacy: San Lazaro" The integration of composition, performance, and computer programming. *Computers and Mathematics with Applications*, 32(1):89–92. 23, 24

Lerdahl, F. and Jackendoff, R. S. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press. 90, 184, 207

Lewis, R. and Beckford, J. (2000). Measuring tonal characteristics of snare drum batter heads. 38, 56

Lippe, C. (1994). Real-time granular sampling using the IRCAM signal processing workstation. *Contemporary Music Review*, 10(2):149–155. doi:10.1080/07494469400640381. 103

London, J. (2012). *Hearing in Time: Psychological Aspects of Musical Meter* (2nd ed.). New York, NY: Oxford University Press. 90

Loureiro, M., De Paula, H., and Yehia, H. C. (2004). Timbre classification of a single musical instrument. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR),* (pp. 546–549). Barcelona, Spain. doi:10.1.1.106.416. 105

Lutfi, R. A. and Liu, C.-J. (2007). Individual differences in source identification from synthesized impact sounds. *The Journal of the Acoustical Society of America*, 122(2):1017–1028. doi:10.1121/1.2751269. 51

Macon, M. W., McCree, A., Lai, W., and Viswanathan, V. (1998). Efficient analysis/synthesis of percussion musical instrument sounds using an all-pole model. In

Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing. Vol. 6. (pp. 3589–3592), Seattle, USA. 3, 4

Maestre, E., Ramirez, R., Kersten, S., and Serra, X. (2009). Expressive concatenative synthesis by reusing samples from real performance recordings. *Computer Music Journal*, 33(4):23–42. 115

Maki-Patola, T., Hämäläinen, P., and Kanerva, A. (2006). The augmented djembe drum: sculpting rhythms. In *Proceedings of the 2006 conference on New interfaces for Musical Expression (NIME06),* (pp. 364–369). Paris, France. 81

Manning, P. (2004). *Electronic and Computer Music,* Revised and Expanded Edition. New York, NY: Oxford University Press. 16

Marozeau, J. and de Cheveigne, A. (2007). The effect of fundamental frequency on the brightness dimension of timbre. *The Journal of the Acoustical Society of America*, 121(1):383–387. doi:10.1121/1.2384910. 125

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* Pacific Grove, CA: Brooks/Cole Publishing Company. 10, 59, 60, 113

Martinet, K., Sølvi, Y., and Jensen, K. (2008). Computer Music Modeling and Retrieval, Sense of Sounds. *Proceedings of the 4th International Symposium, CMMR 2007*, Copenhagen, Denmark. New York, NY: Springer Berlin Heidelberg. 241, 261

Masri, P. (1996). *Computer Modelling of Sound for Transformation and Synthesis of Musical Signals.* PhD thesis, University of Bristol, England. Retrieved from: http://www.mp3-tech.org/programmer/docs/Masri_thesis.pdf. 3

Mathews, M. V. (1963). The Digital Computer as a Musical Instrument. *Science, New Series*, 142(3592):553–557. 50, 81

Mazzola, G. B. (1995). Inverse Performance Theory. In *Proceedings of the 1995 International Computer Music Conference (ICMC)*, San Francisco, USA. 20

Mazzola, G. B. (2011). *Musical Performance.* A Comprehensive Approach: Theory, Analytical Tools, and Case Studies. New York, NY: Springer-Verlag. 20, 26, 27, 28

Mazzola, G. B. and Thalmann, F. (2011). *Musical Composition and Gestural Diagrams.* In Agon, C., Andreatta, M., Assayag, G., Amiot, E., Bresson, J. and Mandereau, J., (Eds.). *Mathematics and Computation in Music*. (pp. 151–164). New York: Springer. 27

Mazzola, G. B. and Zahorka, O. (1994). Tempo Curves Revisited: Hierarchies of Performance Fields. *Computer Music Journal*, 18(1):40–52. 183

McKenna, T. (2000). *Where Digital Music Technology and Law Collide – Contemporary Issues of Digital Sampling, Appropriation and Copyright Law,* Student Prize, 2000 (1). Journal of Information, Law and Technology (JILT), Retrieved from: http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2000_1/mckenna/. 101

Mei, C. (1969). Free Vibrations of Circular Membranes under Arbitrary Tension by the Finite-Element Method. *The Journal of the Acoustical Society of America*, 46:693–700. doi:10.1121/1.1911750. 47

Miranda, E. R. (2001). *Composing Music with Computers.* Boston, MA: Focal Press. 82

Mirza, S. M. (2010). *Introduction to Matlab.* Pakistan Institute of Engineering and Applied Sciences, Department of Physics and Applied Sciences, Islamabad, Pakistan. Retrieved from: http://worldtracker.org/media/library/Mathematics/All%20MATLAB%20Books%20Collection/Introduction%20to%20MATLAB%20-%20Sikander%20M.%20Mirza.pdf. 123

Moore, F. R. (1990). *Elements of Computer Music.* Upper Saddle River, NJ: Prentice Hall. 6

Moravcsik, M. J. (2001). *Musical Sound: An Introduction to the Physics of Music*. New York, NY: Kluwer Academic/Plenum Publishers. 35, 36, 41

Murakami, Y. and Miura, M. (2008). Automatic classification of drum-rhythm patterns employed in popular music. In *Proceedings of the International Conference on Music Perception and Congnition (ICMPC10),* (pp. 450–454), Sapporo, Japan. 49

Murphy, D., Kelloniemi, A., Mullen, J., and Shelley, S. (2007). Acoustic modeling using the digital waveguide mesh. *IEEE Signal Processing Magazine*, 24(2):55–66. doi:10.1109/MSP.2007.323264. 4, 52, 53

Musician's Friend Inc. (2013). Musician's Friend: Musical Instruments Store. Retrieved from: http://www.musiciansfriend.com/concert-percussion/yamaha-css1455-concert-steel-snare-drum-with-ss745a-stand, http://www.musiciansfriend.com/drums-percussion/zildjian-a-custom-crash-cymbal, and http://www.musiciansfriend.com/drums-percussion/roc-n-soc-original-saddle-drum-throne. 79

Myhill, J. (1979a). Controlled Indeterminacy A First Step Towards a Semi-Stochastic Music Language. *Computer Music Journal*, 3(3):12–14. 84

Myhill, J. (1979b). Some Simplifications and Improvements of the Stochastic Music Program. In *Proceedings of the 1978 International Computer Music Conference*, (pp. 272–317). Illinois, USA. Retrieved from: http://hdl.handle.net/2027/spo.bbp2372.1978.023. 84

Naveda, L., Gouyon, F., Guedes, C., and Leman, M. (2009). Multidimensional microtiming in Samba music. In *12th Brazilian Symposium on Computer Music, SBCM*, (pp. 1–12). Recife, Pernambuco, Brazil. 134

Naveda, L., Gouyon, F., Guedes, C., and Leman, M. (2011). Microtiming patterns and interactions with musical properties in samba music. *Journal of New Music Research*, 40(3):225–238. doi:10.1080/09298215.2011.603833. 134

O'Sullivan, L. and Boland, F. (2010). Towards a Fuzzy Logic Approach to Drum Pattern Humanisation. In *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10).* Graz, Austria. 19

Paine, G. (2002). Interactivity, where to from here? *Organised Sound*, 7(3):295–304. doi:10.1017/S1355771802003096. 22

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, 48:115–138. 56

Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach.* Berlin, Germany: Springer. 97, 99

PAS (2014). Percussive Arts Society International Drum Rudiments. Available from: http://www.pas.org/Learn/Rudiments/RudimentsOnline.aspx. 56

Passino, K. M. and Yurkovich, S. (1998). *Fuzzy control.* Menlo Park, CA: Addison Wesley Longman, Inc. 18

Peaden, B. R. and Worland, R. (2011). Acoustic effects of holes on commercial cymbals. *The Journal of the Acoustical Society of America*, 129(4):2616. doi:10.1121/1.3588691. 3, 45

Pearce, M. and Wiggins, G. (2001). Towards a framework for the evaluation of machine compositions. In *Proceedings of the AISB'01 Symposium on Artificial Intelligence and Creativity in the Arts and Sciences.* (pp. 22–32). Brighton, UK: SSAISB. 80

Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C. H., Avanzini, F., Puce, A., and Pollick, F. E. (2009). Multisensory integration of drumming actions: musical expertise affects perceived audiovisual asynchrony. *Experimental Brain Researchr*, 198(2-3):339–352. doi:10.1007/s00221–009–1817–2. 63, 135

Pfordresher, P. and Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychological Research*, 66(1):71–79. doi:10.1007/s004260100075. 56, 63

Pfordresher, P. Q. and Benitez, B. (2007). Temporal coordination between actions and sound during sequence production. *Human Movement Science*, 26(5):742–756. doi:10.1016/j.humov.2007.07.006. 63

Pierce, J. R. (1992). *The Science of Musical Sound* (Rev. Ed.). New York, NY: W. H. Freeman and Co. 29, 50

Pinksterboer, H. (1992). *The Cymbal Book.* Milwaukee, WI: Hal Leonard Corporation. 19, 44, 45

Pressing, J. (1990). Cybernetic issues in interactive performance systems. *Computer Music Journal*, 14(1):12–25. 22

Pressnitzer, D. and McAdams, S. (2000). Acoustics, psychoacoustics and spectral music. *Contemporary Music Review*, 19(2):33–59. doi:10.1080/07494460000640251. 125

Prögler, J. (1995). Searching for Swing: Participatory Discrepancies in the Jazz Rhythm Section. *Ethnomusicology,* Special Issue: Participatory Discrepancies, 39(1):21–54. 16, 135

Puckette, M. (2007). *The theory and technique of electronic music.* Singapore: World Scientific Press. 101

Quarteroni, A., Saleri, F., and Gervasio, P. (2010). *Scientific Computing with MATLAB and Octave* (2nd ed.). New York, NY: Springer Berlin Heidelberg. 123

Raichel, D. R. (2006). *The Science and Applications of Acoustics* (2nd ed.). New York, NY: Springer. 35

Ramesh, V. M. and Sahasrabuddhe, H. V. (2008). Exploring Data Analysis in Music Using Tool Praat. In *1st International Conference on Emerging Trends in Engineering and Technology (ICETET '08),* (pp. 508–509). Nagpur, India, doi:10.1109/ICETET.2008.83. 123

Rath, M. and Waltermann, M. (2008). Exploring the Perceptual Relevance of Inherent Variability of Drum Sounds. In Martinet et al. (Eds.), *Computer Music Modeling and Retrieval, Sense of Sounds.* (pp. 303–312). doi:10.1007/978-3-540-85035-9_21, New York: Springer Berlin Heidelberg. 16

Register, A. H. (2007). *A Guide to MATLAB Object-Oriented Programming.* Boca Raton, FL: Chapman and Hall/CRC. 123

Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's "Träumerei". *Journal of the Acoustical Society of America*, 92(5):2546–2568. 30

Repp, B. H. (1994). Relational invariance of expressive microstructure across global tempo changes in music performance: An exploratory study. *Psychological Research*, 56:269–284. 30, 135

Repp, B. H. (1996). Pedal Timing and Tempo in Expressive Piano Performance: A Preliminary Investigation. *Psychology of Music*, 24:199–221. doi:10.1177/0305735696242011. 30

Repp, B. H. (1997). The effect of tempo on pedal timing in piano performance. *Psychological Research*, 60:164–172. 55

Repp, B. H. (1999). Detecting deviations from metronomic timing in music: Effects of perceptual structure on the mental timekeeper. *Perception, & Psychophysics*, 61:529–548. 30, 55

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, 12(6):969–992. 62

Repp, B. H. (2006). *Musical synchronization.* In Altenmüller, E., Wiesendanger, M., and Kesselring, J. (Eds.). *Music, Motor Control and the Brain.* Oxford, England: Oxford University Press. 30, 62

Rhaouti, L., Chaigne, A., and Joly, P. (1999). Time-domain modeling and numerical simulation of a kettledrum. *The Journal of the Acoustical Society of America*, 105(6):3545–3562. doi:10.1121/1.424679. 47, 48

Richardson, P. (2010). *Acoustic analysis and tuning of cylindrical membranophones.* PhD thesis, Anglia Ruskin University, United Kingdom. 3, 47

Riskedal, E. (2002). *Drum analysis.* PhD thesis, University of Bergen, Norway. 20

Risset, J.-C. and Wessel, D. (1999). *Exploration of Timbre by Analysis and Synthesis.* In Deutsch (Ed.). *The Psychology of Music* (2nd ed.). San Diego, CA: Academic Press. 28, 29, 50, 125

Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17(1):67–99. doi:10.1007/s11023–007–9066–2. 80

Roads, C. (1978). Automated granular synthesis of sound. *Computer Music Journal*, 2(2):61–62. 101

Roads, C. (1996). *The Computer Music Tutorial.* Cambridge, MA: MIT Press. 14, 27, 30, 31, 36, 54, 101

Roads, C. (2004). *Microsound.* Cambridge, MA: MIT Press. 100, 101

Roads, C. (2005). The art of articulation: The electroacoustic music of Horacio Vaggione. *Contemporary Music Review*, 24(4):295–309. doi:10.1080/07494460500172121. 101

Rocchesso, D. (2004). Introduction to sound processing. University of Verona, Italy. 101

Roederer, J. G. (2008). *Generation of Musical Sounds, Complex Tones, and the Perception of Timbre.* In *The Physics and Psychophysics of Music: An Introduction.* (pp. 113–166). doi:10.1007/978-0-387-09474-8_4. New York, NY: Springer. 50, 51

Rogers, J. (2011). *Your Band's First Gig: Getting The Sound Right.* Raleigh, NC: Lulu Press. 39

Rosenbaum, D. A. (2010). *Human Motor Control* (2nd ed.). Burlington, MA: Elsevier. 59, 60, 61

Rossing, T. D. (1992). Acoustics of Drums. *Physics Today*, 45(3):40–47. Available from: http://link.aip.org/link/phtoad/v45/i3/p40/s1. 3, 35, 37

Rossing, T. D. (2000). *Science of Percussion Instruments.* Singapore: World Scientific Publishing Company. 3, 34, 35, 43, 44, 45, 52

Rossing, T. D., Bork, I., Zhao, H., and Fystrom, D. (1992). Acoustics of snare drums. *The Journal of the Acoustical Society of America*, 92(1):84–94. doi:10.1121/1.404080. 3, 37, 38, 40, 47

Rossing, T. D. and Fletcher, N. H. (1983). Nonlinear vibrations in plates and gongs. *The Journal of the Acoustical Society of America*, 73(1):345–351. doi:10.1121/1.388816. 3

Rossing, T. D. and Fletcher, N. H. (2004). *Principles of Vibration and Sound* (2nd ed.). New York, NY: Springer-Verlag. 3, 35

Rowe, R. (1993). *Interactive Music Systems.* Machine Listening and Composing. Cambridge, MA: MIT Press. 22

Rowe, R. (2001). *Machine Musicianship.* Cambridge, MA: MIT Press. 87

Runeson, S. and Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112(4):585–615. 31, 133

Russ, M. (2004). *Sound Synthesis and Sampling,* (2nd ed.). Burlington, MA: Focal Press. 101

Sandler, M. B. (1990). Analysis and synthesis of atonal percussion using high order linear predictive coding. *Applied Acoustics*, 30(1):247–264. doi:10.1016/0003–682X(90)90046–W. 54

Santos, M. J. and Aruin, A. S. (2008). Role of lateral muscles and body orientation in feedforward postural control. *Experimental Brain Research*, 184(4):547–559. doi:10.1007/s00221–007–1123–9. 75

Schedel, M. (2002). Anticipating interactivity: Henry Cowell and the Rhythmicon. *Organised Sound*, 7(3):247–254. doi:10.1017/S1355771802003047. 22

Schedin, S., Gren, P., and Rossing, T. D. (1998). Transient wave response of a cymbal using double-pulsed TV holography. *The Journal of the Acoustical Society of America*, 103(2):1217–1220. doi:10.1121/1.421254. 44

Schloss, W. A. (1985). *On the automatic transcription of percussive music: from acoustic signal to high-level analysis.* PhD thesis, CCRMA Department of Music, Stanford, USA. 41, 88, 89, 90, 91, 92, 97, 98, 99, 195, 200, 201

Schögler, B. (2000). Studying temporal co-ordination in jazz duets. *Musicae Scientiaea, Special Issue*, 3:75–91. doi:10.1177/10298649000030S106. 30

Schroedl, S. (2002). *Drum Tuning: The Ultimate Guide*. Milwaukee, WI: Hal Leonard Corporation. 47

Schroedl, S. (2003). *One Hundred and One Drum Tips*. Milwaukee, WI: Hal Leonard Corporation. 40, 46

Schubert, E. and Wolfe, J. (2006). Does timbral brightness scale with frequency and spectral centroid? *ACTA Acustica United with Acustica*, 92:820–825. 125

Schwarz, D. (2004). *Data-driven concatenative sound synthesis*. PhD thesis, IRCAM, Paris, France. 3, 101, 103, 108

Sedes, A. (2005). Aspects of time in the music of Horacio Vaggione. *Contemporary Music Review*, 24(4):327–334. doi:10.1080/07494460500172188. 101

Serra, M. H. (1993). Stochastic composition and stochastic timbre: GENDY3 by Iannis Xenakis. *Perspectives of New Music*, 31(1):236–257. 84

Serra, M. H., Rubine, D., and Dannenberg, R. B. (1990). Analysis and synthesis of tones by spectral interpolation. *Journal of the Audio Engineering Society*, 38(3):111–128. 30

Serra, X. (1989). *A system for sound analysis/transformation/synthesis based on a deterministic plus stochastic decomposition*. PhD thesis, Department of Music, Stanford University. 30

Sethares, W. A. (2005). *Tuning, Timbre, Spectrum, Scale* (2nd ed.). London, UK: Springer-Verlag. 26, 50

Shaffer, L. H. (1989). Cognition and affect in musical performance. *Contemporary Music Review*, 4(1):381–389. doi:10.1080/07494468900640431. 63, 74

Shove, P. and Repp, B. H. (1995). *Musical motion and performance: Theoretical and empirical perspectives*. In Rink, J., (Ed.), *The Practice of Performance*. Cambridge, England: Cambridge University Press. 55

Şimşekli, U., Jylhä, A., Erkut, C., and Cemgil, A. T. (2011). Real-Time Recognition of Percussive Sounds by a Model-Based Method. *EURASIP Journal on Advances in Signal Processing*, 2011(Article ID 291860):1–14. doi:10.1155/2011/291860. 21

Smith, III, J. O. (1991). Viewpoints on the History of Digital Synthesis. In *Proceedings of the 1991 International Computer Music Conference,* (pp. 1–10), San Francisco: International Computer Music Association. 4, 104

Smith, III, J. O. (2004). Virtual acoustic musical instruments: Review and update. *Journal of New Music Research*, 33(3):283–304. doi:10.1080/0929821042000317859. 101, 108

Smith, L. (1973). Henry Cowell's Rhythmicana. *Anuario Interamericano de Investigacion Musical*, 9:134–147. 22

Sockut, G. H. and Goldberg, R. P. (1979). Database Reorganization - Principles and Practice. *Computing Surveys*, 11(4):371–395. 102

Souvignier, T. (2003). *Loops and Grooves: The Musician's Guide to Groove Machines and Loop Sequencers*. Milwaukee, WI: Hal Leonard Corporation. 102

Starr, E. (2009). *The Everything Music Composition Book: A step-by-step guide to writing music*. Avon, MA: Adams Media. 49

Strong, J. (2006). *Drums For Dummies* (2nd ed.). Hoboken, NJ: Wiley Publishing Inc. 34, 49

Sturm, B. L. (2006). Adaptive concatenative sound synthesis and its application to micromontage composition. *Computer Music Journal*, 30(4):46–66. 100

Sullivan, C. (1990). Extending the Karplus-Strong algorithm to synthesize electric guitar timbres with distortion and feedback. *Computer Music Journal*, 14(3):26–37. 19

Sweeney, P. (2004a). *Beginning Drumset*. Van Nuys, CA: Alfred Publishing Co. Inc. 176

Sweeney, P. (2004b). *Drums for the Absolute Beginner*. Van Nuys, CA: Alfred Publishing Co. Inc. 34, 39, 47

Tidemann, A. and Demiris, Y. (2008). A Drum Machine That Learns to Groove. In *Proceedings of the 31st annual German conference on Advances in Artificial Intelligence (KI '08)*, (pp. 144–151). Heidelberg, Germany. 102

Tindale, A. R. (2004). Classification of snare drum sounds using neural networks. Master's thesis, Mcgill University, Montreal, Canada. 37, 39, 121, 122

Tindale, A. R., Kapur, A., Schloss, W. A., and Tzanetakis, G. (2005a). Indirect Acquisition of Percussion Gestures using Timbre Recognition. In *Proceedings of the Conference on Interdisciplinary Musicology (CIM05)*, Montreal, Canada. 122

Tindale, A. R., Kapur, A., Tzanetakis, G., Driessen, P., and Schloss, W. A. (2005b). A Comparison of Sensor Strategies for Capturing Percussive Gestures. In *Proceedings of the 2005 conference on New Interfaces for Musical Expression (NIME05)*, (pp. 200–203). Vancouver, Canada. 136

Tindale, A. R., Kapur, A., Tzanetakis, G., and Fujinaga, I. (2004). Retrieval of Percussion Gestures using Timbre Classification Techniques. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, (pp. 541–544). Barcelona, Spain. 121, 122

Tipei, S. (1975). MP1: A Computer Program for Music Composition. In *Second Annual Music Computation Conference*, University of Illinois at Urbana-Champaign. 85, 86

Tipei, S. (1984). Maiden Voyages for trumpet, piano, three slide projectors and two-channel tape, on the 25th Anniversary of the Experimental Music Studios record album. *School of Music, University of Illinois at Urbana-Champaign*. 86

Tipei, S. (1987). "Maiden Voyages": A Score Produced with MP1. *Computer Music Journal*, 11(2):49–58. 85, 86

Tipei, S. (1989). The Computer: A Composer's Collaborator. *Leonardo*, 22(2):189–195. 85, 86

Todorov, E. and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11):1226–1235. 62

Toulson, E. (2009). Percussion Acoustics and Quantitative Drum Tuning. In *Audio Engineering Society (AES) Tutorial, T3*, New York, USA. Retrieved from: http://www.robtoulson.rt60.co.uk/rt_docs/RT_AES09_T3_DrumTuning.pdf. 3

Toulson, E., Cuny-Cringy, C., Robinson, P., and Richardson, P. (2008). The perception and importance of drum tuning in live performance and music production. In *Proceedings of The Art of Record Production Conference*, Lowell, MA. 47

Touze, C. and Chaigne, A. (2000). Lyapunov exponents from experimental time series: application to cymbal vibrations. *ACTA Acustica United with Acustica*, 86(3):557–567. 45

Touze, C., Chaigne, A., Rossing, T. D., and Schedin, S. (1998). Analysis of cymbal vibrations using nonlinear signal processing methods. In *Proceedings of the International Symposium on Music Acoustics*, (pp. 377–382). Leavenworth, WA. 4

Trautmann, L., Petrausch, S., and Rabenstein, P. (2001). Physical modeling of drums by transfer function methods. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (pp. 3385–3388). Salt Lake City, UT. 4

Trefethen, L. N. (2000). *Spectral methods in MATLAB.* Philadelphia, PA: Society for Industrial Mathematics. 123

Truax, B. (1987). Real-Time Granulation of Sampled Sound with the DMX-1000. In Berg, P., (Ed.). *Proceedings of the 1987 International Computer Music Conference*, (pp. 138–145). San Francisco: International Computer Music Association. 103

Tucker, S. and Brown, G. J. (2002). Investigating the perception of the size, shape and material of damped and free vibrating plates. *Department of Computer Science, University of Sheffield, Sheffield, UK, Technical Report CS-02-10.* 51

Turcan, P. and Wasson, M. (2003). *Fundamentals of Audio and Video Programming for Games (Pro-Developer).* Redmond, WA: Microsoft Press. 125

Turkle, S. (1992). *Growing up in the Age of Intelligent Machines: Reconstructions of the Psychological and Reconsiderations of the Human.* In Kurzweil (Ed.). *The Age of Intelligent Machines.* Cambridge, MA: MIT Press. 14

Turvey, M. T. (1990). Coordination. *American psychologist*, 45(8):938–953. 58

Uhle, C. and Dittmar, C. (2004). Generation of musical scores of percussive un-pitched instruments from automatically detected events. In *Proceedings of the AES 116th Convention,* Paper 6032, Berlin, Germany. 20

Vaggione, H. (1994). Timbre as syntax: A spectral modeling approach. *Contemporary Music Review*, 10(2):73–83. doi:10.1080/07494469400640311. 103, 104

van den Doel, K. and Pai, D. K. (2003). *Modal synthesis for vibrating objects.* In Greenebaum, K (Ed.). *Audio Anecdotes III.* Natick, MA: A K Peters. 4

Van Steelant, D., Tanghe, K., Degroeve, S., De Baets, B., Leman, M., Martens, J., and De Mulder, T. (2004). Classification of percussive sounds using support vector machines. *Proceedings of the annual machine learning conference of Belgium and The Netherlands, Brussels, Belgium.* 103

von Hornbostel, E. M. and Sachs, C. (1914). Systematik der musikinstrumente Ein versuch. *Zeitschrift für Ethnologie*, 46(4-5):553–590. 19, 41

von Hornbostel, E. M. and Sachs, C. (1961). Classification of musical instruments/ (Baines, Anthony and Wachsmann, Klaus P., Trans.). *The Galpin Society Journal*, 14:3–29. (Original work published 1914). 41

Waadeland, C. H. (2000). *Rhythmic movements and moveable rhythms: syntheses of expressive timing by means of rhythmic frequency modulation.* PhD thesis, Norwegian University of Science and Technology, Norway. 4, 5, 29, 33

Waadeland, C. H. (2001). "It Don't Mean a Thing If It Ain't Got That Swing" - Simulating Expressive Timing by Modulated Movements. *Journal of New Music Research*, 30(1):23–37. doi:10.1076/jnmr.30.1.23.7123. 3, 4, 18, 29

Wagner, A. (2006). Analysis of drumbeats - interaction between drummer, drumstick and instrument. Master's thesis, KTH Royal Institute of Technology, Stockholm, Sweden. Retrieved from: http://www.speech.kth.se/prod/publications/files/1698.pdf. 65, 70, 71, 136

Warren, R. M. (1982). *Auditory Perception: A New Synthesis.* Elmsford, NY: Pergamon Press. 19, 50

Washburne, C. (1998). Play It "Con Filin!": The Swing and Expression of Salsa. *Latin American Music Review/Revista de Música Latinoamericana*, 19(2):160–185. 14, 95

Whalley, I. (2009). Software agents in music and sound art research/creative work: Current state and a possible direction. *Organised Sound*, 14(2):156–167. doi:10.1017/S1355771809000260. 24, 25

Wilbur, C. (1996). Vibrational modes of cymbals and Chladni's law. *The Journal of the Acoustical Society of America*, 100(4):2843. 3

Wilbur, C. and Rossing, T. D. (1997). Visualizing Modes of Vibration in Cymbals with Electronic Holography. In *Proceedings of International Symposium on Simulation, Visualization and Auralization for Acoustic Research and Education (ASVA 97)*, (pp. 675-680). Tokyo, Japan. 43, 45

Wing, A. M., Church, R., and Gentner, D. (1989). Variability in the timing of responses during repetitive tapping with alternate hands. *Psychological Research*, 51(1):28–37. 62, 66

Wing, A. M. and Kristofferson, A. (1973a). Response delays and the timing of discrete motor responses. *Attention, Perception, & Psychophysics*, 14(1):5–12. 62

Wing, A. M. and Kristofferson, A. (1973b). The timing of interresponse intervals. *Perception and Psychophysics*, 13(3):455–460. doi:10.3758/BF03205802. 62

Winkler, T. (1998). *Composing Interactive Music*. Techniques and Ideas Using Max. Cambridge, MA: MIT Press. 22

Winter, D. A. (2009). *Biomechanics and Motor Control of Human Movement* (4th ed.). Hoboken, NJ: Wiley Publishing Inc. 136

Worland, R. (2008). Drum tuning: an experimental analysis of membrane modes under non-uniform tension. In *Proceedings of the 156th Meeting of the Acoustical Society of America (ASA)*, Miami, FL. doi:10.1121/1.3138888. 48

Worland, R. (2010). Normal modes of a musical drumhead under non-uniform tension. *The Journal of the Acoustical Society of America*, 127(1):525–533. 48, 49

Wright, M. and Berdahl, E. (2006). Towards machine learning of expressive microtiming in Brazilian drumming. In Zannos, I., (Ed.). In *Proceedings of the 2006 International Computer Music Conference (ICMC)*, (pp. 572–575). New Orleans, USA. 4, 19

Wright, M., Schloss, W. A., and Tzanetakis, G. (2008). Analyzing Afro-Cuban rhythm using rotation-aware Clave template matching with dynamic programming. In *Proceedings of the International Conference on Music Information Retrieval*, (pp. 647–652). Philadelphia, PA. 4, 5

Xenakis, I. (1991). More Thorough Stochastic Music. In *Proceedings of the International Computer Music Conference (ICMC)*, (pp. 517–520). Montreal, Canada. 19, 84, 85

Xenakis, I. (1992). *Formalized Music: Thought and Mathematics in Composition* Revised Ed.). Hillsdale, NY: Pendragon Press. 81, 84, 100

Xenakis, I. and Brown, R. (1989). Concerning time. *Perspectives of New Music*, 27(1):84–92. 86

Young, D. and Fujinaga, I. (2004). Aobachi: a new interface for Japanese drumming. In *Proceedings of the 2004 Conference on New Interfaces for Musical Expression (NIME04)*, (pp. 23–26). Hamamatsu, Japan. 136

Yu, L. (2004). Fundamental frequency of a circular membrane with a strip of small length. *Zeitschrift für Angewandte Mathematik und Physik (ZAMP)*, 55(3):539–544. doi:10.1007/s00033–004–3004–4. 39

Yu, L. and Wang, C. (2001). Fundamental frequencies of a circular membrane with a centered strip. *Journal of Sound and Vibration*, 239(2):363–368. doi:10.1006/jsvi.2000.3061. 39

Yuen, J. and Horner, A. (1997). Hybrid sampling-wavetable synthesis with genetic algorithms. *Journal of the Audio Engineering Society*, 45(5):316–330. 103

Zagaykevych, A. and Zavada, I. (2007). Development of electronic music in Ukraine: emergence of a research methodology. *Organised Sound*, 12(2):153–165. doi:10.1017/S1355771807001811. 101

Zavada, I. (2008). *3D Composer: A Software for Micro-composition.* PhD thesis, Sydney Conservatorium of Music, University of Sydney, Australia. 81

Zils, A. and Pachet, F. (2001). Musical mosaicing. In *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01)*, Limerick, Ireland. 20, 101

Zölzer, U. (2008). *Digital Audio Signal Processing.* New York, NY: John Wiley & Sons. 101