Advancing the Art of Electronic Percussion

by

Adam Tindale
B.Mus., Queen's University, 2001
M.A., McGill University, 2004

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in Interdisciplinary Studies
(Music, Computer Science, and Electrical Engineering)

Advancing the Art of Electronic Percussion

by

Adam Tindale

B.Mus., Queen's University, 2001

M.A., McGill University, 2004

Supervisory Committee

Dr. George Tzanetakis, Co-Supervisor

(Department of Computer Science, Music, and Electrical Engineering)

Dr. Peter F. Driessen, Co-Supervisor

(Department of Electrical Engineering and Music)

Dr. W. Andrew Schloss, Co-Supervisor

(Department of Music and Computer Science)

Dr. Steve Gibson, Outside Member

(Department of Visual Arts)

Abstract

The goal of this project is to create a new instrument: the E-Drumset. This new interface addresses the lack of expressivity in current electronic percussion devices. The project combines Electrical Engineering for implementing hardware and digital signal processing, Computer Science for implementing musical and mapping software, and Music to devise new playing techniques and ways to combine them into a pedagogy and language of transmission.

Like an acoustic drumset, the E-Drumset consists of different components that can be arranged together as a whole. An acoustic drumset can be thought of as a collection of pedals, drums and cymbals. The E-Drumset consists of the E-Pedal, E-Drum and E-Cymbal. The technology utilized in the E-Drumset includes sensor technologies with newly developed technologies such as acoustically excited physical models and timbre-recognition based instruments. These new technologies are discussed and applied to situations beyond the E-Drumset.

Just building a new controller is not enough. It needs to be thoroughly tested in musical situations and to take into account feedback from musicians (both the player and other members of the ensemble) during the evaluation of the instrument.

Clear and attainable technical guidelines have not been devised for the E-Drumset. In the case of the radiodrum, a spatial controller, improvements can be summarized to be better resolution in space and time. In the case of the E-Drumset the goal is to offer a flexible interface to percussionists where electronic drums are often the bottleneck in bandwidth. There is no clear answer to questions such as how low the latency needs to be to satisfy a drummer; an issue that will be explored through the project.

The goals of the project are to provide the percussionist with an interface that they may sit down and use existing skills. Utilizing the great variety of gesture available to the expert, the E-Drumset allows the percussionist to explore all manners of controllers between acoustic instruments and electronic. To provide a smoother transition to the E-Drumset, notation and exercises for E-Drumset specific gestures

and techniques was devised.

The E-Drumset is a new instrument. Most new interfaces are derived to help lesser players achieve virtuosic ends, while other projects make a controller that is massively configurable where a more static instrument is appropriate. This project provides insight into the theory and practice of new musical interfaces while delivering novel forms of synthesis and gesture recognition appropriate for the E-Drumset.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgements

MISTIC at the University of Victoria has been my home for many years and all of the wonderful people there were a constant support. I couldn't have completed this work without any of you.

Robin Davies has had an incredible effect on me. His understanding during the more experimental phases of the design process were invaluable. His unassuming virtuosity is something to which I aspire.

I have to thank all of the people I have talked to about my ideas at conferences, talks and random encounters. Robert Ferguson was a constant sounding board, cheerleader, and invaluable help with my code. The Princeton crew and ChucK team provided endless help in the good and bad times and always treated me as one of their own, for that I am endlessly grateful to call you friends: Ge Wang, Ajay Kapur, Rebecca Fiebrink, Georg Essl, Matt Hoffman, Phil Davidson, and Perry Cook.

To all those who aren't on this list be sure I am kicking myself for your exclusion.

My colleagues at the Alberta College of Art and Design were endlessly supportive personally, professionally and collegially. My fellow faculty members in the Media Arts and Digital Technologies area were a constant support team: Mary Scott, Wayne Giles, Rita Mckeough, Alan Dunning, and Mitch Kern. Richard Brown, Gord Ferguson, Paul Jackson, and Tyler Los-Jones were great support during my adventures into cycling and bike commuting.

I have to thank coffee and all of the amazing people around the industry who helped me get here. Learning about coffee was a great escape from my real work, yet if provided great fuel. Colin Newell was invaluable reference on coffee, electronics, travel.

There is no way that I could have done this without my wife Katherine. She was relocated many times, left alone with the beast, and all without complaint.

Dedicated to my Grandfather: the first luthier in my family.

Forward

Throughout my dissertation I have strived to figure out how to not be alienated at parties when asked “what do you do?” During my first few years the answer was a long winded discussion of the state of computer music and the nature of control we have over our machines. Now I simply respond by saying that I make digital musical instruments. The answer satisfies most and allows us to discuss more appropriate higher concept topics rather than my summarizing of papers for my victims.

During my exploration of synthesis and gestural possibilities of an extended drumset controller I was exploring ambient and electroacoustic music. Many of the decisions about which types of synthesis to explore were based on these interests. There are other forms of synthesis available but to fully explore them all would be an endless project.

My undergraduate degree was focused on classical music performance. The notes I played were prescribed from composers. It became an important concern for me to democratize my music making and my instrument. I had to have the freedom to play the sounds in my head and to make mistakes that lead to new sounds. In performance there were many times that I was drawing upon the energy and needs of the audience and was inspired to make a sound. Sometimes this sound was available to me, other times it was not. I collected a log of my successes and failures. As time went on, my list of the former overshadowed the latter.

“A lot will depend on what kind of interfaces people use with their laptops, and if they become flexible enough. I’ll give you an example. There’s this percussionist called Glen Velez, and he just plays frame drums. But sometimes he doesn’t just strike the drum. He rubs the side of the drum as if he were polishing it. Or listen to a great tabla player like Zakir Hussain. Even though he’s just playing two drums, with two fundamental resonances, the range of sounds he can get is enormous. I’d ultimately like to see physical models that can react to the performer/programmer like that.” - Guy Sigsworth [1]

[1] http://www.applied-acoustics.com/guysigsworth.htm

CHAPTER

ONE

Introduction

1.1 Motivation

Drumming could be debated as the oldest form of music making; though more likely it was the first instrumental expression of music. Throughout the development of human culture there has been drums and drumming. The percussion family, the group of instruments on which drumming is performed, is the largest family of all instruments. Percussion instruments have been invented and improved throughout human history.

This project combines aspects of Computer Science, Electrical Engineering and Music to design, implement and test a new musical controller. The goal is the creation of an Electronic Drumset - or E-Drumset - that implements new technology and yields new aesthetic considerations.

The rate of advance in Music Technology has increased enormously, yet commercial electronic percussion has been stagnant for approximately 20 years. There is not even an entry for it in the standard book Percussion Instruments and their Histories [13]. Two of the world's most famous electronic percussionists, Bill Bruford and Danny Carey, are still using the Simmons SDX electronic drumset that was made in 1984. The hardware has not improved and the sounds have only evolved to imitate sounds that have already become popular. There is little or no push to advance the technology or the artistry of electronic percussion, for a number of reasons, some of them commercial. There are interfaces available that can transfer the gross gestures of percussion performance into data for making music (i.e., MIDI, OSC, etc.), but the majority of these devices are not able to convey the full range of expressive possibilities of an expert performer. Current percussion controllers only provide data on the velocity of the impact, forcing a single dimension of data to represent something as complex as musical gestures.

To provide more flexibility for the performer, we need to capture more of the complexity of gesture intention with our technology. Although it is possible to simply augment current models with more sensors, this thesis investigates another approach that I used in my M.A. [163]: timbre-recognition-based instruments used to infer the gesture by analyzing the timbre of a sound - an analog to the trained ear of a musician.

1.2 Design of New Controller

There are many important factors in the design of a new controller. This project intends to leverage the knowledge gained from the exploration and evaluation of the current models in order to create a novel and coherent interface. The following sections address the main topics in the process.

The current controllers available to drummers are mostly inhibitive: they impose limitations on the freedom of the player in terms of the types of sticks or surfaces that must be used, or similar limitations. A controller has been built that addresses this issue and attempts to free performers from all factors that inhibit their freedom.

1.3 Background

This thesis contains background and related work in the areas of gesture, gestural control, instrument design, new instruments for musical expression, electronic

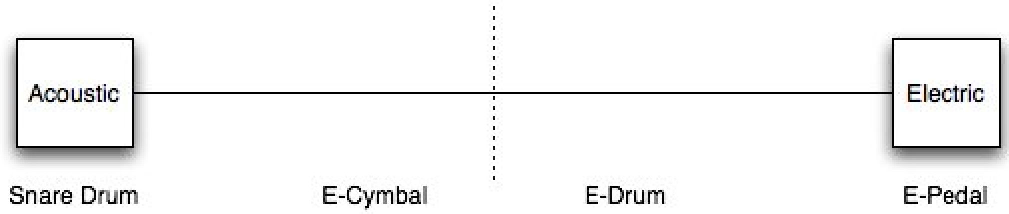


Figure 1.1: The range of instruments from acoustic to electronic.

percussion (commercial and academic) and related issues in percussion. The information garnered from this research informs the current work and provides advice from seminal electronic instrument builders such as Perry Cook and theorists such as Marcelo Wanderley.

Roberto Aimi [1, 2] has been working with percussion controllers and has invented many different controllers such as the beatbug and the echodrum. Recently, he has devised a new instrument called the convdrum that uses a convolution method to combine the sound of a drum or drumhead with an impulse response. Although this method allows for expressiveness with any excitation, it does lack the flexibility of the physical modeling approach.

The most flexible commercial percussion interface ever created is the Korg Wavedrum. This instrument used a combination of an acoustic drumhead, physical modeling and audio effects to create an original sound that was able to respond to any gesture that created a sound generated by the drumhead.

1.4 Breadth of Gesture

Percussion instruments are debatedly the most dynamic instruments in the world. The variety of gestures available to an expert performer is incredible. A survey of the standard techniques and extended techniques will be conducted in order to make sure that the new instrument will allow percussionists to utilize their techniques.

Gestural control of music through computers is a relatively recent occurrence, starting in the early 1980s. Even so, there are many gestural devices that utilize percussion or are inspired by percussion. Gestural language from these controllers is investigated to try to glean more techniques that may be useful for the percussionist.

Possibly the most important aspect of the E-Drumset is the ability for any type of drumstick or implement to affect the sound in a commensurate way.

Through the course of the development of percussion there has been a need to communicate gestures to other performers and composers. Once a gesture has become commonplace (or sometimes before) it must be represented in notation. The notation for these gestures will be studied so that I may be able to create a reasonable notation language for new gestures.

1.5 Project Goals

The first stage of the project was to implement software that can recognize the gestures of a performer by utilizing the concept of a timbre-recognition-based instrument. The software is implemented for a variety of instruments that will comprise the E-Drumset. Mapping strategies were devised to translate the extracted parameters to musical events. Physical models are utilized as a synthesis engine for the instrument, providing the instrument with many possibilities to exploit the expressiveness of an expert performer.

Just building a new controller is not enough. It needs to be thoroughly tested in musical situations and to take into account feedback from musicians (both the player and other members of the ensemble) during the evaluation of the instrument.

1.5.1 Controller

Instead of using sensors to capture the motion of the performer and translating this motion into musical meaning, timbre-recognition-based instruments analyze the sound produced by an acoustic instrument that is captured by a microphone. This sound is then processed and evaluated by machine learning algorithms to determine what specific timbre was produced by the instrument. Once the timbre has been labeled, it can be inferred what gesture of the musician caused the timbre.

Timbre-recognition based instruments offer many advantages over other techniques. Many modern performance situations are either amplified or are recorded; therefore, the microphones are already present. This means that there is no need for extra wires for sensors, or added weight to an instrument, which may interfere with the musician's performance. The repertoire of the gestures is easily expandable, any gesture that creates sound can be captured and subsequently labeled.

1.5.2 Mapping

A major issue is devising how data acquired from gesture capture devices can be mapped to meaningful parameters for a synthesis engine. A traditional percussive gesture is very cohesive and visceral, lending itself to many simple mappings (e.g. triggering events or samples). An investigation into the nature of mapping currently used in the field was conducted in order to discover how the instrument may utilize a combination of context-dependent and context-independent mapping strategies. A problem that I see with most drum controllers is that they are so configurable that they do not have any consistency in sound, whereas an acoustic drumset is static in its timbre palette and demands an experimental or virtuosic player to produce its more exotic sounds.

A comparison of the coherence of direct and indirect mappings will be explored from both the player and audience perspective. Mappings must be intuitive for the performer and communicable to the audience in order to create an authentic experience.

1.5.3 Synthesis Engine

Physical modelling offers the most potential for a synthesis engine capable of the flexibility needed for a true musical instrument. Currently, there are few models of percussion instruments and most are no more complex than modelling of ideal membranes. The resonating body of the drum is rarely explored.

Commercial electronic drum pads come equipped with a contact microphone and provide a physical membrane with an audio input to the computer. The sound of this membrane may be augmented with a virtual tube provided with digital waveguide physical modelling to simulate acoustic drum instruments.

1.6 Main Contributions

The research and development presented in this dissertation has yielded new ideas and implementations that will be outlined in this section. Supporting material in the form of audio and video materials is available online (http://www.adamtindale.com).

1.6.1 Software

- **Novel Position Tracking Approaches: Explicit and Implicit** :

  When a drum controller is struck it makes a different sound depending on where the strike position. Position tracking of percussion controllers can be done explicitly via analysis algorithms or implicitly by allowing the timbre change to be heard.

- **Acoustically Excited Synthesis Algorithms**:

  Utilizing the audio signal from commercial electronic drum controllers to directly drive synthesis methods. Control methods for classic synthesis algorithms have been added to aid in control and preservation of these methods.

- **Timbre-Recognition Based Instruments**:

  A new paradigm in gesture recognition systems that utilizes music information retrieval approaches. Timbre-recognition based instruments use real time classification techniques to determine the gesture used in order to provide high level information to the player.

- **Decoupled Analysis and Synthesis Software**:

  Separating the software tasks into independent pieces of software allows for additional applications for the software. Several projects using either the analysis or the synthesis systems were developed as a benefit of decoupled software components. Decoupling has not been previously explored in drum controllers.

- **Fuzzy Navigation Techniques**:

  The keyboard and mouse are very precise input devices for computers. Sometimes the user is looking for something similar to something else and require imprecision during navigation. The drum controller is utilized to provide control but variation in a method of fuzzy navigation that lies between random selection and precise selection.

- **Surrogate Sensor**:

  The surrogate sensor is the process of using a sensor in conjunction with machine learning techniques to train another sensor to replicate the results. This approach allows for less expensive sensors to be used in a performance system while retaining some of the benefits that a higher cost sensor.

1.6.2 Hardware

- **Hardware Generalization**:

  Current commercial electronic drums use drum pads with piezos embedded inside of them. The E-Drumset software allows for standard commercial hardware to be used with no modification. Drum pads are connected to a computer sound card and all features are available, allowing for owners of drum pads and computers to use the system with no other purchase or upgrade.

- **E-Pedal**:

  The E-Pedal is a modular hardware platform that can be installed on an acoustic bass drum pedal or a practice pedal. The E-Pedal not only tracks strikes but it also tracks the position of the pedal and the position of the foot. The controller allows for multidimensional control of continuous variables and discrete events.

1.6.3 Pedagogy

- **Pedagogy Feedback**:

  Developing a derivative instrument has the advantage of borrowing pedagogy from its parent. The new instrument however, may develop ideas that can be utilized on the parent instrument to provide similar or, more interestingly, different results.

- **Notation**:

  To communicate with traditionally trained drummers, notation for new gestures discovered was developed. The notation aims to be consistent with current models and only augment the set by notating the new gestural possibilities of the E-Drumset.

1.7 Supporting Activities

New ideas need proof of their viability. For digital musical instruments, viability is established on stage. The following section outlines ideas used to test the viability of the E-Drumset.

- **Visual Sound Source Identification Techniques**:

  Computer music performance in large ensembles has the problem of the audience being unable to determine the source of any particular sound. Relationships between gestures and sounds were explored in performance to allow the audience to decipher which player was creating the current sound.

- **Contextual Gestures**:

  Digital musical instruments provide an incredible amount of flexibility and allow the performer to change parameters of the instrument during performance. Contextual gestures is the notion that once a gestural language is established, it must remain consistent to establish authenticity with the audience. Therefore, the use of control gestures must be carefully considered to preserve the relationship with the audience.

- **Performance Locations**:

  Computer music is typically performed in a concert hall. Performances were held in many locations and situations to expose the general public to alternate controllers and alternate forms of music making.

- **Experimental Results**

  Experiments were carried out to verify the accuracy of machine learning algorithms employed. Upon examination of the results of the experiments decisions could be made about exactly which algorithms were to be used in performance.

CHAPTER

TWO

History of Electronic Percussion

I believe that the use of noise to make music will continue and increase until we reach a music produced through the aid of electrical instruments ...
– John Cage, 1937

2.1 Introduction

Electronic percussion has been around in different forms for more than fifty years. Throughout this chapter the concepts, players, and the evolving technology will be explored. This chapter is divided into four sections that outline major periods of development for electronic drums. For each period there be review of the major technological developments followed by be a survey of performance practice and theory enabled by the technological advances of the era.

The first period covers a pre-history of electronic drums and examines the theory and technology that has lead to current developments. The following three sections break up the next fifty years into reasonable divisions, ending with the present. The survey presented is an extensive look at the most influential representatives but is by no means exhaustive.

2.2 1850 - 1960

1850-1960	**1960-1980**	**1980-2000**	**2000-Present**
Early Electronic Music, First Drum Machines	Early Electronic Performance, First Electronic Drums	Solo Performers, MIDI, Simmons SDX	Integrated Electronics, Mandala Pad

2.2.1 Technology 1850 - 1960

Drum Machines

> Q - How many drummers does it take to change a light bulb?
>
> A - None. They have machines to do that now. [1]

Due to the popularity of drum machines, drummers began to worry about their future in bands. Many musicians would replace their drummer with a drum machine. Once the initial excitement over the new steadiness and interesting sound of the drum machines wore off, drummers began to learn to interact with musicians differently. The 1970s had been a time of extended drum solos and open forms (ie improvised music). Drummers had trained themselves to function in this musical world. The music of the 1980s was radically different in tone and required a style of drumming with few embellishments and steady beats; drum machines were perfect for this style. The shift can likely be attributed to the exploration of new technology combined with the search for a new sound to distinguish musicians from the music that came before.

Chamberlin Rhythmate

The Chamberlin Rhythmate was the very first dedicated drum machine. It was built in a garage in Uplands California in 1948 and continued production of various models until 1969. The Chamberlin Rhythmate is essentially a bank of tapes that can be played back at various speeds. The technology is exactly that of a Mellotron but preceded it by nearly twenty years. The tapes contain loops of acoustic drummers

[1] http://www.users.bigpond.com/prodigalson/drummer.htm

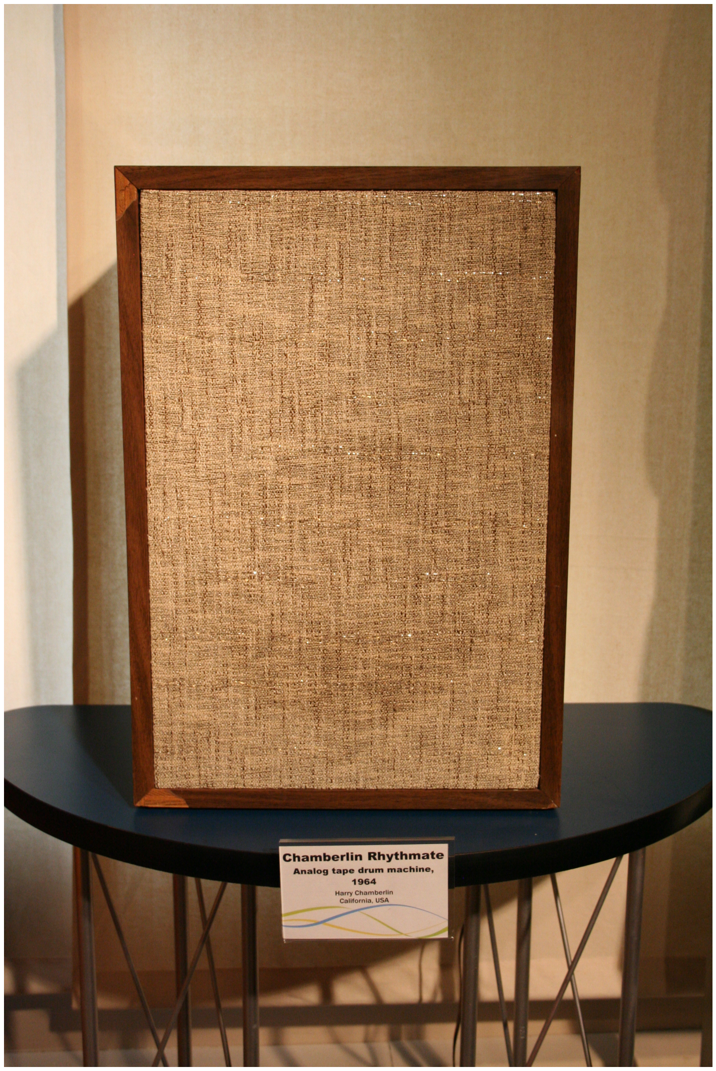


Figure 2.1: The Chamberlin Rhythmate.

playing patterns of various styles. Because the loops are recordings, it is not known how many recordings used the Rhythmate and how many are actual drummers. The Chamberlin company, run by Harry Chamberlin, also produced a keyboard that used the same technology.

Wurlitzer Sideman

Wurlitzer Sideman, 1959, works with tape loops to create rhythms with rotary contacts. The Sideman sold rather well and made the rhythm machine concept popular. The taps create an impulse that is fed through an analog filter to create a pitched

Figure 2.2: The Wurlitzer Sideman.

Figure 2.3: Control panel for the Wurlitzer Sideman.

sound similar to a drum. The hihat and cymbals are simulated with the highest pitches, the toms and snare with mid range pitches and the bass drum with a low pitched filter. The synthesis method is simple yet effective, and is still in wide use today. The unit was designed to sit beside your organ with the controls on the top. It has its own power amp and speaker built in.

The Sideman was one of the most important developments in the history of electronic music. It was the product of a very progressive organ company that was trying to revolutionize the industry. They achieved their goals as the Sideman has almost universal respect.

The Sideman would gain even more fame when Wolfgang Flür of Kraftwerk [42] would hack the contacts to make the first full electronic drumset.

2.2.2 Performance 1850 - 1960

There are many histories of electronic music. This section will outline major innovations in technology, compositional ideas, and the movement towards live electronic music performance. The purpose of this section is to illustrate the highlights in history that lead to the present time. The first major event was the invention of electricity, without it there can be no electric music. The next event in my timeline may be a point of conjecture: the telegraph (invented in 1753). A classical definition of a musical instrument is that which is intended to produces music. A more modern definition is that which produces sound. With the classical definition a violin no longer is a musical instrument when used as a blunt weapon. While the telegraph was not intended to produce music, it did produce sound via electricity. In more modern times it has been used as a musical element in pop and classical music, as well as an element of sound design for television and movies.

The next major step forward in my timeline of electronic music is the work of Helmholtz and his treatise *On the Sensation of Tone* [68]. Helmholtz formalized the first notion of timbre as a sound that varies overtones while the fundamental frequency remains constant. Since the formalization of timbre in this method, timbre has become the focus of compositional techniques.

The first electronic instrument, by both the classical and modern definition, is the Electrisches Musikinstrument invented by Ernst Lorenz in 1885. This invention predates Thomas Cahill's Teleharmonium by almost twenty years. Electrical vibrations drove an electromagnet to create sound.

The next nearly fifty years saw numerous inventions of electronic instruments, most notably the Theremin, Electronic Sackbut, and the Ondes Martinot. Edgar Varese and the futurists theorized about a more democratic notion of music to include industrial sounds as musical elements. Modern interpretations of futurist literature allow the inclusion of all sound as elements to be manipulated by the composer. The notion of the composer is also recontextualized by a modern reading to include all those who manipulate sound. These modern viewpoints have lead to movements such as Sound Art, Noize Music, and Ambient Music to name a few.

Imaginary Landscape No. 1, John Cage, 1939

This landmark piece is composed for piano, chinese cymbal and two variable speed phonograph turntables. The pieces specifically calls for Victor Frequency Record 84522B and Victor Constant Note Record No. 24, Victor Frequency Record 84522A; records that contain fixed pitch sine waves. The record players are to be played by musicians who manipulate the pitch of the records by varying the speed of the record. An interesting development in this piece is the call for a fifth performer to control the mix of all the parts since it was originally intended for radio broadcast. This was possibly the first time a recording engineer is referred to as a musician.

Deserts, Edgar Varese, 1949-1954

This is first piece where live performers play with prerecorded tape. A major concern in this piece was synchronization between the musicians, a wind orchestra, and the tape piece - two monophonic tape players with discrete speakers placed at opposite sides of the stage. The tape and orchestra sections are alternated so to avoid the synchronization problem. The mental framework of an “organized sound” used in this piece would form the basis of Varese’s *Poeme Electronique* at the Brussels World Fair.

Musica su Due Dimensioni, Bruno Maderna, 1952

This piece is composed for solo flute and tape with a specific tape operator. The tape is stopped and started at different points in the piece, as indicated in the score to the tape operator. The alternation of duo and solo sections for the flute performer allows freedom during the sections where the tape is not present.

2.3 1960 - 1980

1850-1960	**1960-1980**	**1980-2000**	**2000-Present**
Early Electronic Music, First Drum Machines	Early Electronic Performance, First Electronic Drums	Solo Performers, MIDI, Simmons SDX	Integrated Electronics, Mandala Pad

2.3.1 Technology 1960 - 1980

Moog 1130 Drum Controller

The Moog 1130 drum controller (1973) featured a standard eight inch plastic head, a piezo transducer, and sensitivity knobs on the shell (see figure 2.4).

Drummers began attaching triggers to modular synthesizers to join synthesists and keyboard players in the electronic sound revolution. In response to the bourgeoning market, Robert Moog designed a controller to be easily integrated into the Moog series of synthesizers.

Electroharmonix Drum Units

In the 1970s, Electroharmonix made a series of drum trigger devices. The devices were housed in standard guitar effect pedals chassis'. The top of the devices had a piezo mounted inside a round rubber pad. Some of the devices also had an external input for standard drum trigger pads.

There were many different units in the series: The Space Drum, The Panic Button, The Sonic Boomer, The Rolling Thunder, The Super Space Drum, The Crash Pad, The Clap Track. The Sequencer Drum, The Clockworks Controller. Each unit had a slightly different synthesizer unit contained inside, except for the two sequencers: The Sequencer Drum and The Clockworks Controller.

Figure 2.4: Moog 1130 Drum controller

Boss Dr. Drum

The Boss Dr. Drum was a very similar design to the Electroharmonix drum trigger series but offered an external trigger option (see figure 2.5). Drummers could attach other drum pads or Piezos to the unit, preserving the life of the pad built in to the unit. The sound was produced with a swept oscillator and offered the drummer control over pitch, sweep depth, decay, and LFO via knobs on the unit.

Syndrum

The syndrum was the first autonomous electronic percussion interface. It included a piezo attached to shell that housed the electronics that produce the sounds. The sound generation unit was a voltage controlled oscillator with a voltage controlled amplifier that was connected to the Piezo to read onsets that controlled the amplitude envelope. With four switches on the side of the drum the user could control the volume, pitch, duration, and range of the pitch sweep. The sweep knob was configured so that a twelve o'clock position would be no sweep, completely clock wise

Figure 2.5: Boss Dr. Drum.

Figure 2.6: (a)Original Syndrum. (b) Syndrum 2 with outboard synthesizer and foot controller.

was maximum pitch sweep upwards, and completely counter-clockwise is maximum pitch sweep downwards.

Future versions of the syndrum would add an LFO to the main oscillator, switchable oscillator types (sine, square, sawtooth), and a low pass filter. The Syndrum Single, and later the Syndrum 2 (see figure 2.6), would see the separation of the drum controller from the synthesizer. The outboard synthesizer added multiple oscillators, filter envelopes, noise generators with a mixer, a configurable envelope generator, external voltage control, and an optional footswtich that could be used to control the unit during performance.

Synare 3

The Synare series (1-3) was made by Harvey Starr of Starr Instruments and was made popular due to the signature descending oscillator sound and the affordable price. The unit was built in the so-called "flying saucer" design where the top of the unit was a drum pad and the control knobs were arranged around the bevelled edges so that the the text could be read when looking down on the instrument.

2.3.2 Performance 1960 - 1980

Early Electronic Music Performance

The music of David Tudor represents the first usage of electronics that truly embraced the possibilities inherent in technology. David Tudor would set up tables of electronic devices and patch them together randomly. His performances involved manipulating the devices and discovering the sounds and relationships of the gestures. This method of performance is now common in the modern Noize Music scene.

The League of Automatic Music Composers

From 1978-1983, Jim Horton, Tim Perkis and John Bischoff formed the first live electronic music ensemble: The League of Automatic Music Composers. The group utilized a Comodore computer called the KIM-1 that was available at low with 1152 bytes of RAM, 2048 bytes of ROM and 30 I/O lines.

The League performed both composed music and improvised music. They utilized concepts of drones and tuning systems developed by LaMonte Young and Pauline Oliveros a decade earlier. They programmed these systems into the computer along with artificial intelligence software. The group often promoted themselves as computer mediators and their advertisements described the group as 50% machine and 50% human.

> "The scene at Mills seemed worlds away from the electronic music studios I had been exposed to. They still had the public access studio going at

that time, and they let me try out the electronic equipment myself and showed me how things worked. David (Behrman) was rehearsing with Rich Gold, John Bischoff, and Jim Horton, who were using tiny computers called KIMs. They were not exactly my image of what computers were like – a board about the size of a sheet of paper with a tiny keypad and a few chips." – George Lewis [2]

History of Electronic Percussion

Early percussion interfaces were made during the late 1960's, with simple acoustic pickups attached to surfaces that were struck; the signals from these transducers were routed through envelope followers (common items in the early modular synthesizers) that produced a voltage proportional to the strike intensity, together with a discriminated trigger pulse. – Joseph Paradiso [131]

Question- "One of the strangest pieces was 'Procession' (Every Good Boy Deserves Favour, 1971) which featured the pioneering work of Graeme Edge's electronic drum kit. How did that come about?"

Graeme- "...I'd got in touch with the professor of electronics at Sussex University, Brian Groves. We worked up an electronic drum kit, a marvelous idea. I had the control panel in front of me, it's old hat now but we were the first to do it. There were pieces of rubber with silver paper on the back with a silver coil that moved up and down inside a magnet that produced a signal, so it was touch sensitive. I had 5 snares across the top and then ten tom-toms and then a whole octave of bass drums underneath my feet and then four lots of 16 sequencers, two on each side. There was a gap–to play a space–a tambourine, ebony stick, snare and three tom-toms. This was pre-chip days, back then you did it all with transistors. So it had something like 500 transistors. The electronic

[2] **Composers and the Computer**. Curtis Roads. William Kaufman Publishers. pp. 79. 1985

> drums inside looked something like spaghetti. When it worked it was superb, but it was before its day, because it was so sensitive..." [3]

Wolfgang Flür

The first dedicated electronic drummer was Wolfgang Flür of the band Kraftwerk. Flür was an active member of Kraftwerk from 1973–1987. Before joining Kraftwerk Flür was an active drummer with jazz and dance bands. Flür started a band called The Spirits of Sound with guitarist Michael Rother. Rother was hired by Kraftwerk during their experimental instrumental music phase. Flür was approached by Florian Schneider-Esleben and Ralf Hütter in 1973 to join Kraftwerk. At the time, the three were all architects but soon would be the first full time electronic musicians in the world.

> If I pressed the key repeatedly, I thought that, with a little practice, I'd be able to play my own rhythms with my fingertips. Although it wasn't a very comfortable device to play, it had an electrifying sound, and it was quite advanced for the time.
>
> Florian and I were deeply enthusiastic, and that evening we played drums totally live, without the pre-programmed rhythms. How could we arrange thing so I could play it more comfortably, perhaps like a proper drum kit? After a lot of toing and froing, it occurred to us that we could detach the contact cables leading to the small keys and link them to other contacts.

The drum machine that the band discovered was the Wurlitzer Sideman (see section 2.2.1). Flür began to design a device that he could use to trigger the drum machine using percussive gestures from his training as a drummer. Using his training as a cabinet make and an architect he encased the machine into a chassis that was robust and covered in sparkling tile to give visual accent during live performance (see figure2.7).

[3] http://en.wikipedia.org/wiki/Electronic_drum

> It didn't take me long to find out that the device would have to be similar to a drum. You'd hold a stick in your hand and hold the flat surface.The stick would have to be from a conducting metal, such as copper and the flat surface would have to be made from a conductor, too. The plate would be rounded in shape, like a drum skin, and an individual disc would be provided for each sound.

Karl Bartos, a classically trained percussionist, joined the group in 1974 and Flür built a duplicate controller for him. This resulted in the well known stage orientation of two drummers in the middle with Florian and Ralf playing keyboards on the outside.

During a later phase of the band, smaller controllers were utilized to allow the band to move around the stage. Flür created a small version of drum controller with a single contact that could be held in one hand. The device was played in the same manner of a hand-held cowbell that was popularized during performances of the song *Pocket Calculator* (see figure 2.8).

The band continued to discuss integrating electronic percussion and gesture with their stage show. They began to develop a percussion cage that used arrays of light beams, originally with visible light and then later with infrared light, that were broken by the performer to trigger an action. This device was very similar to the drum cage in SensorBand [162]. During its premiere performance the device failed to trigger any sounds and resulted in an embarrassed Flür waving his hands madly in vain until he resorted to walking back to his main device to continue the performance. The device quickly fell out of favor with the band due to continued technical difficulties.

Flür eventually left Kraftwerk in 1987 to pursue solo projects. Flür's main recording project is a group called Yamo. Flür has embraced laptop performance and often performs solo in clubs. He also performs in an electronic percussion duo called X.

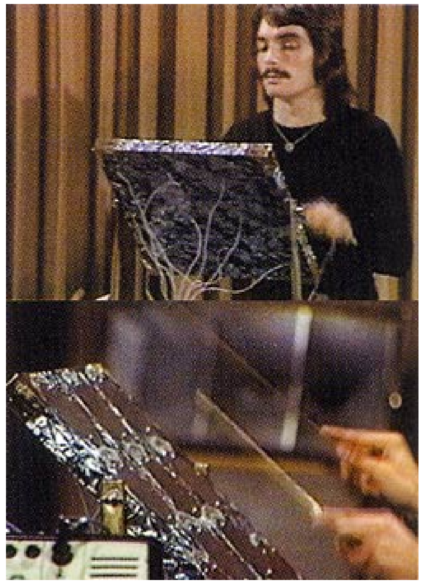


Figure 2.7: Wolfgang Flür using his electronic drumpads with Kraftwerk.

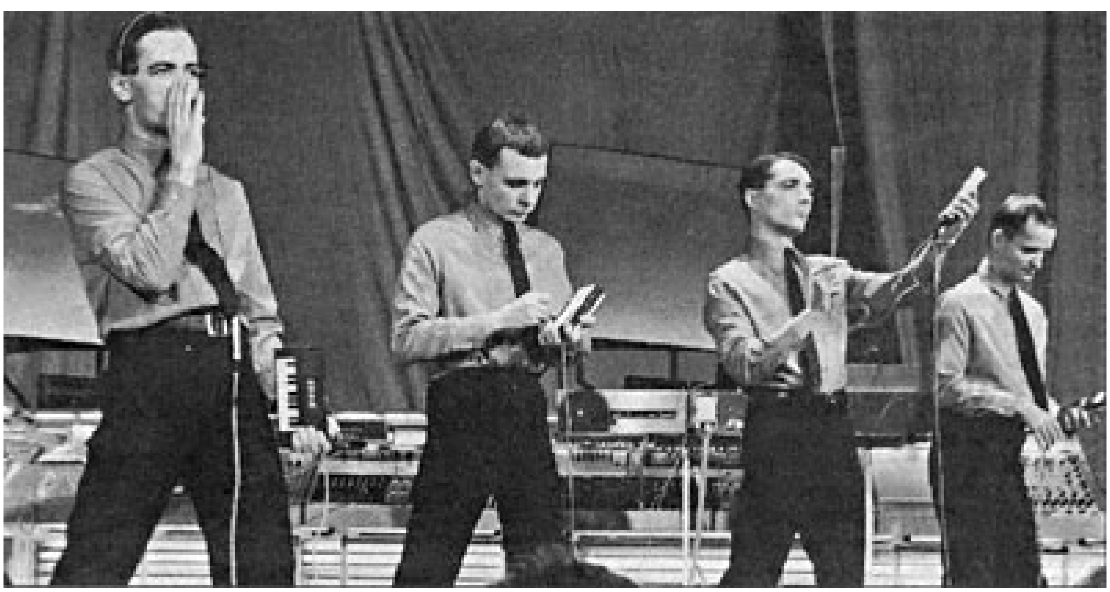


Figure 2.8: Kraftwerk playing smaller instruments. Flür can be seen playing the small devices like a cowbell (2nd from right).

2.4 1980-2000

1850-1960	1960-1980	1980-2000	2000-Present
Early Electronic Music, First Drum Machines	Early Electronic Performance, First Electronic Drums	Solo Performers, MIDI, Simmons SDX	Integrated Electronics, Mandala Pad

2.4.1 Technology 1980 - 2000

Roland TR Series

The Roland TR series (TR-606,TR-707, TR-808, TR-909) are still the most popular drum machines in use today, either directly or sampled. The TR series was produced between 1980 and 1990. With the exception of the Simmons drums, no other electronic percussion device, be it drum machine or electronic drums, had such a significant impact on music.

The TR series features a set of 16 buttons with LEDs on top that is the main interface to the built-in step sequencer (see figure 2.9). For each sound in the machine there is a virtual sequence controlled by the buttons. When switching control of the sounds the sequence the LEDs would switch to show which steps had a trigger or not. Only the TR-707 showed the complete sequence data at once by using a small screen with a matrix display.

The first three machines in the series used an internal synthesizer to generate sounds. The TR-909 (see figure 2.10) was the first dedicated drum machine to include samples for sound generation. The cymbal sounds were 6-Bit samples that could be altered via tuning and amplitude controls.

Simmons SDX

> So who is this instrument for? It's for drummers who want to create their own identities–drummers who want to user their own samples and have total control over their own performance. It's also for people who

Figure 2.9: The Roland TR-808.

Figure 2.10: The Roland TR-909.

> are interested in exploring alternatives in musical instruments. Some of these people might not be performers. But either way, the people who eventually use the SDX are going to be explorers, and it is going to be interesting to hear what they discover. [112]

Debatedly, the most advanced electronic drumset to ever be commercially available is the Simmons SDX drumkit, released in 1987. The kit featured multi-zone pads and a full sampling computer (see figure 2.11 and 2.12). Simmons dubbed the term "Zone Intelligence" to indicate that the pads were the first to include multiple zones with their own sensors. The snare drum pad included a separate zone for the rim, center, and edge of the drum. The cymbals were very similar pads with the zones representing the bow, bell, and edge of the cymbal. The tom pads could be ordered as a single, double, or triple zone configuration. The double zone pads had a rim zone and a remaining area zone.

One of the main problems with the Simmons pads was that they were very rigid, causing injury to drummers' wrists. The cymbals in the SDX included a flex feature so that when they were struck the cymbal would move in a similar fashion to an acoustic cymbal.

Due to the large price tag and development costs the SDX sold few units and was a factor in the company folding shortly after the release of the SDX. Simmons released a Drum Expander that was a rack unit that turned drum trigger signals into MIDI messages. Around the same time, AKAI began releasing samplers that had more memory and were much cheaper then a full SDX setup. For those drummers with a limited budget and who were more interested in bringing a few samples to the stage to augment an acoustic setup this solution offered a better fit.

Simmons has since reopened providing budget electronic drumsets. Dave Simmons is not involved in the business in any significant way.

Figure 2.11: A view of the Simmons SDX pads, controller, and acoustic instruments integrated.

Figure 2.12: Simmons SDX brain.

Radiodrum

The Radiodrum is a standard computer music controller and has been since its creation, even so, the history of the radio drum is a confusing topic. Where does the radio drum and the radio baton begin? When were they invented and by whom? This section will outline the history of the Radiodrum and its sibling, the Radio Baton. Current directions in development will be surveyed in terms of technology and performance.

There is one thing that can bring a new interface to the forefront of computer computer music: users. The Radiodrum has three people who have dedicated a significant amount of the their time and talents to mastering the instrument and either composing for the instrument themselves, or collaborating with other composers. Dr. Richard Boulanger and Dr. W. Andrew Schloss are two such individuals.

Beyond his technical contributions to the Radiodrum, Max V. Mathews uses it as a musical interface; Mathews uses the Radiodrum as a conducting interface. When used in conjunction with the Conductor program [109] (formerly Groove [110]) the Radiodrum controls tempo and other parameters of music playback.

Dr. Richard Boulanger, through extensive collaboration with Max Mathews, is one of the few virtuoso performers on the Radio Baton. He composed a multitude of pieces for the instrument in a solo, duo, chamber, and concerto setting.

Dr. W. Andrew Schloss has performed on the Radiodrum for nearly twenty years, almost as long as the Radiodrum has been available. Schloss performed solo on the Radiodrum in both pre-composed and improvised contexts. More recently he has joined with Hilario Duran to play Cuban jazz music and experiment with those forms. See figure 2.13.

Although the Radiodrum works very well as a spatial controller, it is not optimized for percussive gestures. The phenomenon of "duffs" is a byproduct of a low sample rate, whereby the velocity information of the strike is unreliable because no samples are captured during the high part of the stroke. The simple solution is to raise the sampling rate in order to capture the intermediate moments and provide

Figure 2.13: Andrew Schloss performing on the Radiodrum.

more accurate velocity information.

The newest method for processing these signals involves many steps [121]. The first involves calibrating the input signals in order to increase the spatial accuracy. This is done by sampling the signal when the sticks are closest and when they are furthest away from the drum. The means and variances of these signals are also collected and used to statistically eliminate the noise in the signals.

Roberto Bressin created an accurate physical model of the bodhran called the Vodhran [105] that utilizes the Radiodrum hardware. The bohdran is a traditional Irish drum played with a double ended stick. A custom stick was made where each end of the stick can be tracked. Traditional bohdran techniques can be accurately tracked and the gestures can be translated to the Vodhran for sonification.

Kjetil Falkenberg Hansen created a thimble controller that is easily placed on the finger tip of a performer [66]. This thimble acts as the stick of the Radiodrum. Hansen uses the thimble to track the motions of his hand on the surface of the Radiodrum. A variety of different scratching techniques are recognized by a PD patch and can be used for traditional DJ techniques or to control more abstract instruments.

Sofia Dahl has utilized the Radiodrum to control sounding objects in a system called Invisiball [17]. A thimble controller, similar to that on the DJ interface created by Falkenberg, tracks the position of the finger through the sensing space. The computer simulates a ball rolling on a surface that was being manipulated as the finger moved in space. Stretchable material is fit over the drum so to provide haptic feedback to the performer as well as a direct correlation between the effect of the material and that of the virtual surface.

Drumkat

The Drumkat is a force sensing resistor (FSR) based percussion controller distributed by Alternate Mode, previously Kat Percussion. The use of FSR technology allows for the pads to be placed together tightly without vibration insulation yet still preventing crosstalk or false triggering. FSRs allows for other features such as aftertouch to be included. The drumkat has ten discrete pads, four MIDI outputs, four MIDI inputs, multiple piezo trigger inputs, expression pedal inputs, and sustain pedal inputs. The internal software is able to generate a variety of midi messages.

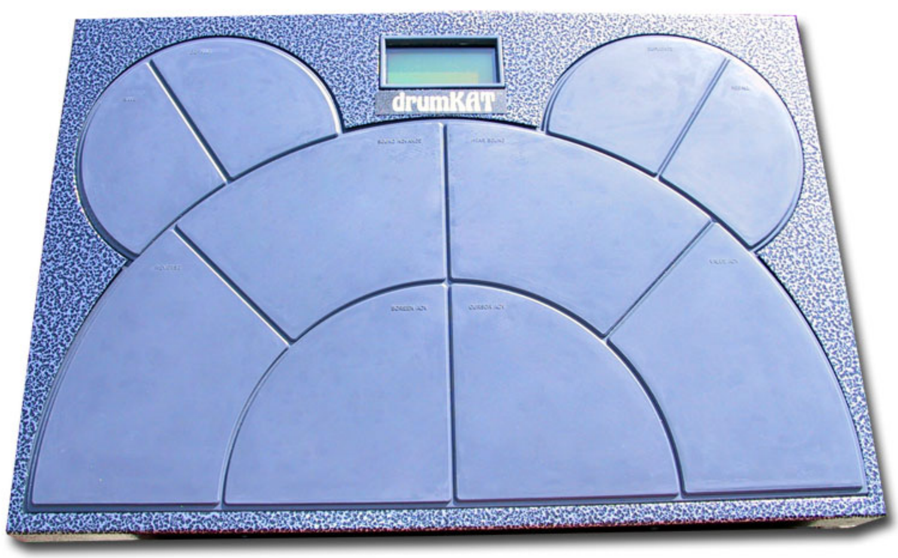


Figure 2.14: Alternate Mode Drumkat Turbo.

Figure 2.15: Alternate Mode Drumkat Turbo panel view.

The Drumkat has an interesting method of programming. The user depresses a sustain pedal to switch the controller from play mode to edit mode and then each pad has an edit function. For example, pad one advances to the next screen and pad two moves back one screen. The use of contextual gestures for programming makes programming the interface more friendly for drummers who may be new to technology and saves having to include other buttons on the machine specially for editing.

Alternate Mode shows incredible commitment to their customers by declaring a "War on Obsolescence."[4] The company has not changed the basic hardware design for nearly twenty years. Occasionally the company will offer a new machine with small variety to the layout of the pads or number of inputs. Instead of forcing customers to upgrade the hardware when new features are created, the company provides new firmware - some as paid upgrades for new features and some as free upgrades for bug fixes and minor feature additions.

Alternate Mode currently provides various editions of the Drumkat but also has a Malletkat, a controller with a pad layout like a xylophone or marimba without graduated bars, a Trapkat, another drum controller but with twenty pads, and a Pankat, a controller with a pad layout similar to steel pans.

Korg Wavedrum

The Korg Wavedrum had a short life in the commercial domain despite being a very advanced and versatile interface. There are three microphones underneath a

[4]http://www.alternatemode.com

standard drumhead that are submixed into a single signal. The microphones are positioned nearly an inch from the center of the drum sixty degrees apart from each other to provide even coverage across the whole drumhead.

An additional rim attachment can be affixed to the top of the Wavedrum. The attachment contains a contact microphone and a cable that allows the output of the contact microphone to be added to the signal chain of the Wavedrum. The rim attachment contains ridges so that a stick may be scraped across them to create scratching sounds.

Once the audio is collected it is transformed directly using digital waveguide physical modelling algorithms. An optional programmer and footpedal would allow the drummer to modify the algorithms on the fly. For each program, one or two parameters are assigned for control. In the case of programs to which two parameters are assigned, moving the pedal will simultaneously control both parameters. Precisely what is controlled over depends upon the make-up of the sound and how it has been programmed.

Figure 2.16: The Korg Wavedrum with rim attachment.

The RE1 remote editor can be used to control up to eight Wavedrum parameters. The RE1 is powered via the single multi-pin connector cable from the Wavedrum to the RE1. Once connected, there is instant interplay between the two units.

The Wavedrum provides various MIDI-related functions. It is not possible to change these MIDI functions from their factory settings when using the Wavedrum by itself, an RE1 Remote Editor is needed. With the RE1, it is possible to change the Global channel, Note number, Velocity and Aftertouch, Modulation, Program change. System Exclusive messages can be used to load and store the Wavedrum's program data.

The Wavedrum cannot be "played" into a sequencer in order that the sound you generate can be sent back to you, faithfully or otherwise. This is because the sound generating system of a Wavedrum is not purely under MIDI control. In fact, only on playback (audio playback, not MIDI) is it possible to get a true picture of what was done.

2.4.2 Performance 1980 - 2000

Roy Wooten a.k.a. Futureman

Roy Wooten, experimental drummer with Bela Fleck and the Flecktones, created the Drumitar in 1985 by modifying the SynthAxe (an act that caused some outrage since the SynthAxe was such an expensive item at the time). Futureman's experiments represent early experiments in alternate controllers and hardware hacking electronic musical instruments. The Drumitar was mounted on a strap and allowed Futureman to move about the stage in a manner similar the other electric instruments in the band, creating a new spatial relationship between drummer and audience, much the same way the SH-101 did for keyboard players. The spirit of the Drumitar would be commercialized into a device called the Zendrum. [5]

[5] http://www.zendrum.com

Figure 2.17: Futureman with his custom made Drumitar.

Tina Blaine

Tina Blaine, a.k.a. Bean, began using electronics as part of D'uckoo in the early 1980's. Like Mario DeCiutiis, Blaine was playing mallet instruments in increasingly loud environments and looking for ways to more accurately and reliably mic her instruments. Blaine began working with engineers to create a system where each marimba bar had a contact microphone for amplification. These instruments could then be attached to drum brains to trigger other devices via MIDI.

While in D'uckoo Blaine built many custom MIDI controllers and experimented with ideas long before they were approached by commercialization or academia. One such controller and concept was realized in the MIDI Ball the group featured at performances. The MIDI Ball was a large beach ball that the crowd would bounce from person to person, as is commonly seen at many concerts. The MIDI Ball had custom hardware to convert the audience interactions into MIDI signals and then route them back to the band wirelessly, more than a decade before any commercial product.

After D'uckoo, Blaine returned to her ethnomusicological roots performing in African drum ensembles, gamelan, and drum circles. More recently, Blaine has been performing with the Hang (pronounced "hung") and modifying its sound via a microphone and electronics.

Figure 2.18: Tina Blaine with D'uckoo.

Figure 2.19: D'uckoo with the MidiBall.

Mario DeCiutiis

Mario DeCiutiis is the Founder and CEO of Alternate Mode, the creators of the Drumkat MIDI controller. Deciutiis began his playing career as an acoustic drummer in the 1960s and after studying classical percussion transfered to playing vibraphone in popular music. In the 1970s, Deciutiis played in numerous funk bands where he had difficulty in high volume situations and began to experiment with micing and amplification techniques. He was not satisfied with the lack of reliability of the current technology for a heavily touring musician.

In a 1980s issue of Modern Drummer DeCiutiis read about Bill Katoski who was building a mallet synthesizer. DeCiutiis and Katoski began working together to develop the Malletkat. The prototype was completed in 1986 and was the beginning of the Kat company, which would later become Alternate Mode.

DeCiutiis used various synthesizers and samplers to experiment with alternative sounds and not just emulations of vibraphones. DeCiutiis often controls non-percussion sounds as a challenge so the he might improve his skills and conceptual ability[114]. More recently he has included the ArChaos software to control visual elements in his performances.

DeCiutiis is an active educator in schools and online, providing performance and tutorial videos for various Alternate Mode products on their website. He has been the Principle Percussionist with the Radio City Music Hall since 1979 and has used the Malletkat and synthesizers in musicals long before substituting musicians with keyboards become possible, let alone standard practice.

Bill Bruford

Bill Bruford has played drums in Yes, King Crimson, UK, Genesis, Earthworks (see Discography). Bruford was one of the earliest complete adopters of electronic drums and certainly one of the most famous. Bruford used the Simmons SDX (see section 2.4.1) as his main electronic instrument as part of King Crimson and Earthworks for many years.

Bruford became very frustrated with the state of electronic drums and that his preferred instrument was nearly a decade old and made by a company that folded shortly after release of the product. Bruford eventually stopped using electronic drums altogether and now uses acoustic drums exclusively, though his style of drumming has been influenced by his time with electronic drums.

During his time with Earthworks, Bruford revolutionized electronic drum performance by playing pitches, melodies, and chords via the electronics. The first incantation of Earthworks featured Bruford's electronics as the only chording instrument in the ensemble. Bruford functioned as both the keyboard player and the drummer, creating new patterns and textures.

The following is an excerpt and interview with Bill Bruford in Modern Drummer from 2001 [116] where Bruford explains his beginning and end of playing electronic drums.

> 6) Why did you play electronic drums?
>
> Rapidly advancing technology in the early '80s offered the drummer a kaleidoscopic range of sonic possibilities, at once seductive and intriguing. The ability to play melodies, chords, and repetitive pitched loops opened new horizons. Surely it would be easier to play "new music" if you have "new sounds" at your fingertips? Well, yes and no. It depends how much your new timbres become the essence and form of the piece without which the piece could not exist, rather than just the icing on the cake. But at the start the sky appeared to be the limit. I introduced them into King Crimson with considerable enthusiasm.
>
> 7) Why did you give up playing electronic drums?
>
> a) The excruciating amount of time needed to extract something interesting from a reluctant and unreliable group of disparate instruments played a part. The things only became interesting beyond their design capabilities, and when you intermarried them with other instruments and got them to talk to each other through MIDI. In the fourteen or fifteen

Figure 2.20: Bill Bruford with Simmons SDX integrated with an acoustic drumset.

years I was actively on board, I suppose I gave rise to no more than fourteen to fifteen compositions which were absolutely a function of electronic percussion, and whose charm arose uniquely from that instrument. At about one a year, that's not a great output, given the time it took.

b) The complexity of the instruments caused the bottom to fall out of the market and the manufacturers were forced to cater for the home-entertainment market, where everything remains on preset one, and comes ready cooked.

c) Shipping costs, maintenance, and unreliability, all took their toll.

2.5 2000 - Present

1850-1960	**1960-1980**	**1980-2000**	**2000-Present**
Early Electronic Music, First Drum Machines	Early Electronic Performance, First Electronic Drums	Solo Performers, MIDI, Simmons SDX	Integrated Electronics, Mandala Pad

2.5.1 Technology 2000 - Present

Since the year two thousand there have been a few technological advances of note but a number of new performers of electronic percussion. Danny Carey abandoned the Simmons SDX system as his main electronic component and began utilizing the Mandala Pads.

The technology available to the average person has increased to a point where many people are making their own devices. The New Interfaces for Musical Expression (NIME) conference has become the main venue for new electronic music interfaces in academia. Many electronic percussion devices have been demoed at this conference but most are abandoned after they are debuted. Many of the successful devices from the NIME conference are examined in Chapter 3 as part of the survey of gesture capture methodologies.

Roberto Aimi

Roberto Aimi received his Ph.D. from the MIT media lab where he focused on developing novel percussion interfaces [2]. Aimi created many interesting electronic devices, including the Beatbug for the Toy Orchestra [1]. The Beatbugs are a networked collection of triggers that are anthropomorphized to resemble common bugs by having a large body with two FSR sensors attached to resemble antennae. The Beatbugs are aimed at children to allow them to collaborate musically by passing musical elements to each other that they are able to manipulate using the FSRs.

The Convdrum is the work that most resembles the work presented in this thesis. A piezo microphone is attached to a drum and fed into a computer where a convolution algorithm is used. The sounds Aimi uses are acoustic percussion sounds so that the electronic drummer is able to have a full range of timbral possibilities.

Convolution offers many advantages as a synthesis method. When one uses a high quality sample of an acoustic instrument then the it is possible to get a high quality reproduction of the acoustic instrument. Since the impact of the strike of a drum creates a varied timbre depending on implement, strike position, and strike force, these variations are mapped onto the sample used in the convolution method. The result is the ability to vary timbre in an intuitive manner and to reproduce an accurate sound of an acoustic drum.

Convolution also has the potential to be implemented in a way that they latency is very low. The disadvantage of convolution is the ability to change sounds on the fly. Convolution is locked to the sample used as the impulse response and it is difficult to vary the sound of impulse response without creating artifacts in the sound. Convolution is also computationally expensive when compared with other methods. There are methods available that use Fast Fourier Transforms of various lengths to create a balanced tradeoff with latency and computational load. Aimi explores various methods in his dissertation, but this is not a major issue when running few channels of convolution on modern hardware.

Mandala Pad

Figure 2.21: The Mandala Pad USB version.

The Mandala drum pad by the Synesthesia Corporation is the first high resolution radial position drum stroke detection instrument. The pad uses a film surface to divide the pad into one hundred and twenty eight zones. There are two versions: one that has an included drum brain and one that connects to a computer via a USB port. The pad has an eleven inch diameter.

The original Mandala pad provided a drum brain that combined a synthesis engine and MIDI output. The module played various synthesis and sample based patches. The strike position could be mapped to effects or to change pitches. The one hundred and twenty eight zones could be divided into seven regions that could be assigned to pitches, samples, or presets. The module was programmable so that the player could create their own presets or use it to control and exterior synthesizer.

The USB version of the Mandala pad is completely bus powered allowing for a very portable system. The software included is a port of the software from the hardware version and has many of the same system features. However the software comes bundled with new software that has a snare drum sampler that contains over one thousand samples to illustrate the level of control possible with more than one hundred levels of strike position and velocity.

The Mandala pad is in use by many leading drummers, including Danny Carey of Tool, Pat Mastelotto of King Crimson and various side projects, Will Calhoun of Living Color, and Matt Chamberlain who plays with Tori Amos.

2.5.2 Performance 2000 - Present

Danny Carey

Danny Carey is the drummer for the band Tool. He integrates electronics amongst his acoustic drumset. For many years he has integrated the Simmons SDX and Korg Wavedrum but recently he has migrated his electronics to Mandala pads with Native Instruments Battery software.

Carey is a very loud and aggressive drummer and deeply committed to the craft. On the most recent Tool recording, *10,000 Days*, the band filled the studio with Helium in order to allow the transients of the drums to travel faster and accentuate the attack of the drums. Carey regularly plays and solos in odd time signatures with Tool and often utilizes melodic ostinatos on an electronic drum while using the rest of his limbs to function as an acoustic drummer.

> Most of the samples I have are in my Simmons SDX. I'm still using

Figure 2.22: Danny Carey with Simmons SDX and Korg Wavedrum integrated into an Acoustic Drumset.

> that antique — no one has come up with anything better. It's just ridiculous.
>
> The electronic drumming world is the most retarded thing. I'm so disgustapated[sic] about the whole thing. The fact that the Simmons SDX is the still the coolest thing out there and is twelve years old really frustrates me. Dennis Grzesik of Simmons still takes care of me, but I must say that the Simmons have been very reliable. That's why I'm able to keep playing them.
>
> I'll admit, ddrums and the Roland V-Drums feel great, and they don't have any MIDI lag like the old SDX does. But for me, the problem is that they don't have any surface intelligence, and that's one of the most important factors. To be able to bend pitches or change any kind of parameter as you move around the head in an intelligent way is key.
>
> MD: The state of electronic drumming today is an interesting point.
>
> Well, it *was* seen as that geeky '80s thing. But back then a lot of those guys were trying to imitate real drums, and that's why it was cheesy. They weren't using it as another instrument to add to the drums. That's

the way I've always looked at electronics.

Solo Electronic Drummers

A new trend in electronic percussion is to utilize technology to perform a modern version of the one-man-band. The classic image of a one-man-band can be seen in the movie Mary Poppins where the player has a variety of drums, cymbals, horns, and accessories that are controlled not only with the hands and feet but knees, elbows, mouth, and head in creative and entertaining ways.

As electronic drums moved from synthesis to sample-based engines new possibilities arose for drummers. Hitting a pad could trigger any instrument or sampled sound, allowing for drummers to play the drums but sound like an orchestra or passing train.

The following sections highlight some drummers who perform in this fashion. There is a distinction made between drummers who use sequences and those who do not. In the 1970s it was common for drummers to play along with drum loops in disco music. When the idea of a complete musical experience came about there was a general attitude in the music industry that electronics had an element of magic since the performance could be sequenced and then lip synced, or it could be quantized, etc. The original solo electronic drum music was created without any loops or sequences, every note that was heard was explicitly triggered by the performer. This not only moved the art forward but displayed a high level of composition, forethought, and technical prowess. In more recent times these attitudes have relaxed and many drummers now uses loops and sequences in their solo performances. While still technically proficient it represents a shift in mindset of artists and audiences.

Electronic Drummers without Sequences

At the onset of solo electronic drummers there was an effort to provide authenticity to the audience. Since any gesture may trigger any sound, audiences became confused when watching electronic drummers. Were the sounds pre-recorded or were they actually being triggered by the drummer? Many videos at this time had a notice indicating to the audience that no pre-recorded loops were used in the performance.

Akira Jimbo is a Japanese virtuoso electronic drummer who specializes in independence, polyrhythms, and afro-cuban ostinatos. Jimbo integrates Yamaha electronic drums into his acoustic drumset and uses them to trigger electronic music that he accompanies. Jimbo's compositions sound like a full electronic Pop band but are produced entirely live, hit by hit.

Tony Verderosa a.k.a. the V-Man was one of the first dedicated electronic drummers. He developed a concept of the Drummer DJ where a drummer would play drum loops in the style of a Drum'n'Bass DJ [184]. Verderosa currently works for Yamaha producing samples for their line of electronic drums and as an artistic consultant.

Terry Bozzio is a virtuoso drummer who performed with the Brecker Brothers, Frank Zappa, and many other famous musicians. While he has experimented with electronic drums and percussion devices throughout his career, he primarily remained an acoustic drummer. Bozzio's drums represent an interesting shift from electronic concepts to acoustic concepts. Bozzio tunes his drums to pitches and has his cymbals designed to be very pitch centered, almost bell-like. Bozzio's setup includes nearly fifty drums and fifty cymbals. The result of his efforts (besides many hours of setup and teardown) is the audiences are able to sing his melodies.

Duracell, a.k.a. Andrew Dymond, plays an acoustic drumset fitted with triggers connected to a Nord Mico-Modular to perform eight bit style music.[6] Duracell's performances revolve around his athletic endurance as spectacle.

The CrossTalk electronic percussion ensemble at the University of Arizona is directed by Norman Weinberg and Robin Horn.[7] The ensemble utilizes various commercial electronic percussion devices. The group composes and arranges pieces specifically for the ensemble, and focuses primarily on various forms of popular and ethnic music.

[6] http://www.geocities.com/tribalboomboom/duracell.html
[7] http://www.crosstalk.arizona.edu

Jeff Quay is the drumset player for the Blue Man Group[8] Quay is a heavy user of the Drumkat. In his solo performances, he utilizes a Yamaha breath controller to provide continuous control in addition to foot controllers.

Electronic Drummers with Sequences

Electronic and acoustic drummers have begun playing with electronic sequences to create music. The drummers tend to play mostly acoustic drums as a method of authenticity to avoid confusing the audience as to the source of sound.

Jojo Mayer is a famous jazz drummer turned Drum'n'Bass innovator. Mayer began producing electronic music when he realized that because of his virtuosity he would be able to create the drum loops faster by playing them on a drumset. Mayer currently plays with Nerve: an ensemble dedicated to playing music generally produced exclusively with electronics.

the matic is an electronic music duo from Los Angeles that features McKay Garner, a.k.a. Drumbot, on electronic drums with guitarist and producer Adam Tenenbaum. Tenenbaum provides loops and electronic music while Garner provides percussion and pitched sounds. Garner uses a Drumkat with custom triggers arranged in traditional piano arrangement.

Tracey Kroll is an electronic drummer who performs solo[9] compositions with sequences and as part of the duo *Spinning Plates*[10]. KJ Sawka plays Drum'N'Bass music on a primarily acoustic drumset with sequences he composes himself[11].

2.6 Summary

In modern times the lines between a drum synthesizer, sampler, and sequencer have been blurred because there are many software packages that offer all of these features in one package. This section outlined the major developments over the last approximately fifty years. Two trends are now occurring: pure electronic drummers

[8] http://www.blueman.com
[9] http://www.myspace.com/tkroll
[10] www.myspace.com/splates
[11] http://www.myspace.com/kjsawka

are disappearing and electronic percussion elements are becoming more common in contemporary drum sets. The groundwork has been laid to integrate electronics into modern music on both an aesthetic and technological front.

Of the two trends outlined above, the second is disconcerting. Many of the original advocates feel that the tools have not evolved since their inception and that their exploration of the medium was complete. Acoustic drums offer potentially infinite variation while providing consistency with practice. Commercial electronic drums have only offered one potentially infinite device but today it sounds over-used and dated. The trend of these virtuosos abandoning the potential that electronics offers is the main motivation of this thesis and my solutions and innovations will be the focus of the remaining chapters.

The technology of electronic drums has remained relatively constant for nearly twenty years. As music and drumming has advanced, the tools of electronic drums have staled. The two potential outcomes are further innovation or the loss of an art. This thesis aims to inject new life into electronic percussion by examining the strengths and weakness of the devices surveyed in this chapter and creating a new design.

CHAPTER

THREE

Background

Civilization advances by extending the number of important operations which we can perform without thinking of them.

– Alfred North Whitehead

3.1 Introduction

Throughout this chapter previous related research will be discussed. This research falls into three main areas: Timbre Recognition, Signal Processing, and Physics. Aspects of these works will be drawn upon in subsequent chapters to outline the software designed to carry out the task of timbre recognition of the electronic drums and synthesis algorithms for live computer music performance.

The systems described in this chapter reflect the current state of research in areas related to electronic percussion. Currently, systems provide analysis of recordings and not live input. The goal of this thesis is to provide functionality from all of the research contained in this chapter into a real-time system.

3.2 Timbre Recognition

Automatic timbre recognition is a very new area of research. Most of the research presented in this section is categorized as automatic music transcription. These systems aim to emulate and automate the human behaviour of listening to music and identifying and transcribing the individual components. Since these systems must include an analysis phase that performs timbre recognition tasks, they are included in this section.

Andrew Schloss' Ph.D. thesis [151] is a pivotal study for many reasons; it was the first major investigation into automatic drum transcription, it is an early study of automatic music transcription, and it contributes some insight into automatic tempo tracking. The work was completed using computers and digital technology during a time when they were very limited and there were few centres in the world with the capacity for digital audio research. The study presents a system for capturing a performance and transcribing it using computers. Conga drums were used for testing and transcription, no other sound was present in the signal. Schloss cascaded three high-pass filters on the signal at specific frequencies so that only the attack of the strike is left. From other analysis, Schloss knew that there was a steady-state portion of the signal which occurs at approximately the same instance after an attack point. The steady-state portion of the waveform was analyzed using a spectral matching template that determined which drum was struck and if it was high or low. An envelope follower was then used in order to determine if the stroke was damped or not. When all of these analysis procedures were finished, Schloss was able to accurately detect which drum was hit and which playing technique was used to strike it.

Masataka Goto's beat tracking systems have application in drum transcription models. Goto's first published project [52] demonstrates a transcription for pop music where he works under the assumption that the bass drum is often on beats one and three while the snare drum is often on beats two and four. His system included methods for recognizing snare drum and bass drum signals and then deriving a tempo map by which he could perform his beat tracking.

Goto's method for recognizing the drums is relatively simple yet very robust. When a transient signal is recognized a Fast Fourier Transform (FFT) is done on that segment. Once a few of these segments are accumulated there are two prominent peaks at the bottom of the spectrum. The lowest peak corresponds to the bass drum and the second peak corresponds to the snare drum. This gives an estimation for the fundamental frequency of each drum. Subsequent segments are then analyzed and it can be determined whether they contain a snare drum or a bass drum. Once the drums are recognized, their positions can be inserted into the beat tracking system for analysis.

Goto then added another step to his system [53]. He was able to track the harmonic motion of a song using spectral analysis techniques. He made a generalization by stating that in pop music chord changes often occur on the first beat of the bar. This gave his system another clue which he could use to analyze and predict the beat of a piece of music.

A major step was taken when Goto was able to demonstrate these techniques in real time [50]. With the increase in computing power available and the optimization of his system for real-time performance Goto was able to demonstrate that his system could effectively track the tempo of a piece of music. The beat prediction algorithms were also improved upon, so that the system could make reasonable predictions about the tempo in the future and compare it with the input signals so as to still be performed in real time.

A major survey on the state of automatic classification of musical instruments was conducted by Herrera, Peters, and Dubnov [72]. Two different techniques are discussed throughout the paper: the taxonomic approach (demonstrated by Martin) and the perceptual approach (demonstrated by Grey in 1977) [60]. The definition proposed is outlined below:

> Perceptual description departs from taxonomic classification in that it tries to find features that explain human perception of sounds, while

> the latter is interested in assigning to sounds some label from a previously established taxonomy (family of musical instruments, instruments names, sound effects category ...). Therefore, the latter may be considered deterministic while the former is derived from experimental results using human subjects or artificial systems that simulate some of their perceptual processes.

Their proposed purposes for performing these types of classifications are either to provide labels for monophonic recordings or to provide indexes for locating the main instruments that are in a particular stream of music. Since the problem of dealing with polyphonic musical streams is yet to be solved, these applications are only aimed at monophonic streams of music.

The use of features is discussed from many different angles: micro-temporal and macro-temporal features describe a sound in small or large divisions of time. The relevant features for perceptual classification and taxonomic classification are discussed for each individual instrument. An extensive summary of the different types of classifiers used for automatic classification of musical instruments is given as part of the study.

Herrera also developed methods of recognizing and classifying percussion instruments for the purposes of integration into the MPEG-7 specification. His first major paper investigating techniques that apply to drum transcription is a survey of the major classification strategies for audio content analysis [70]. The survey shows results from major studies which utilize the different classifiers for recognition. There were no conclusive statements in the study since the different studies were classifying different signals. It was concluded that appropriate classification strategies need to be used to the task, and that they should be optimized for maximum efficiency and accuracy.

Fabien Gouyon joined Herrera's efforts and they released the first paper directed towards drum transcription [56]. Their study demonstrated exploratory techniques for labeling drums in the context of a Pop song. They used an extensive feature set as well as different classifiers to get initial results on the possibilities of recognizing drums. The study showed very high recognition rates with up to three drums being played simultaneously.

The project was furthered with the publication of their next paper [57]. The features that were introduced in the previous work were evaluated for their efficiency in classifying the input signals. The features were put into three categories that were useful to classifying signals into families of drums, then sub-families, and finally the specific drum. Many combinations were attempted and the most effective ones are presented in the paper. The classification strategy used was k-Nearest Neighbour (k-NN).

The next step in Hererra's project was a large-scale study of drum transcription [71]. Very large sets of training and test data were used. Additional features were added, such as zero-crossing calculations, from the previous studies. The resulting study demonstrated very high recognition rates utilizing the system developed over the previous years. The study also introduced the notion of labeling drums by their manufacturer. Although the acoustic instruments were difficult to classify, in terms of manufacturer, synthetic sounds produced by the famous Roland drum machines, TR-808 and TR-909 (see section 2.4.1), were successfully recognized.

Fujinaga has developed a system that is capable of recognizing orchestral instruments, and in later studies, achieving the same result in real-time [46, 47]. A total of thirty-nine orchestral instruments were used with samples from the McGill University Master Samples [127]. The system breaks the identification problem into a number of different hierarchies from three-timbre groupings to thirty-nine-timbre groupings.

An exemplar-based classifier, k-NN [30] was used to classify the instruments. The classifier works by creating a feature space based on a set of training data. When an unknown sample is input into the classifier, each value in the new feature vector is compared with the training data to determine the closest class. The new data is classified based on the class that has the majority of feature vectors nearest to the unknown input feature vector.

The evolution of the project from its inception to real-time performance is outlined in [47] as follows:

> In the first experiment (Fujinaga 1998), only the spectral shapes from manually selected steady-state portion of the sounds were used as data [148] to the recognition process. The features calculated from the spectral data included centroid and other higher order moments, such as skewness and kurtosis. The recognition rate for the 39-timbre group was 50% and for a 3-instrument group (clarinet, trumpet, and bowed violin) was 81%. In all cases, the standard leave-one-out cross-validation is used on the training set to calculate the recognition rate.
>
> In the second experiment [44], two improvements were made. First, spectral envelope data generated by the fiddle program was used as the basis of the analysis. Second, the features of dynamically changing spectrum envelopes, such as the velocity of the centroid and its variance were added. The recognition rate increased to 64% for the 39-timbre group and 98% for the 3-timbre group.

The real-time system reduced the analysis time from five hundred milliseconds to two hundred and fifty milliseconds and added more feature extraction algorithms. The recognition rate of the small groupings was between 95–100% and the thirty-nine-timbre group was increased to 68%.

Fujinaga's system is an example of a system that can accurately classify data by selecting features and a classifier that only needs a short window and minimal time to perform the classification. It is also important to note that throughout the publications by this project, the word timbre is used interchangeably with instrument.

Martin's doctoral dissertation [106] is an extensive and interesting work in instrument recognition. Martin designed software that was able to discriminate between non-percussive orchestral instruments with a high degree of accuracy.

Martin's idea was to provide tools that were appropriate for the problem at hand.

As such, a multiple-classifier hierarchical approach was taken. Different taxonomies were used during the classification process. An example of one the taxonomies used is illustrated below:

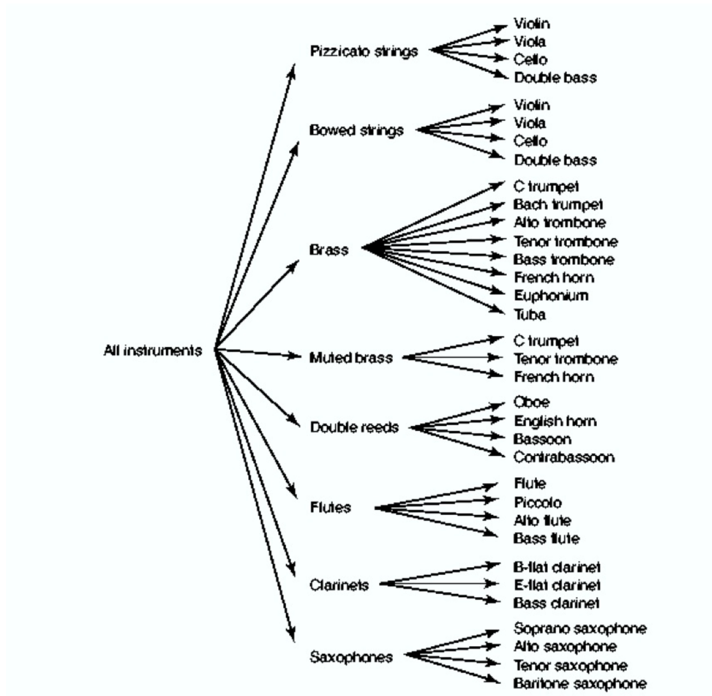


Figure 3.1: An example of a taxonomy by Martin.

Many different features were made available to the system in order to classify a signal. Thirty-one perceptually salient acoustic features based upon a human auditory model were used. At each node, different combinations and weightings of the features were fed to the classifier. These combinations were pre-determined by a set of training data.

One thousand twenty-three isolated tones from fifteen orchestral instruments were used as testing and training data for the system. The samples used were taken from the McGill Master Samples. A statistical pattern-recognition technique was used as a classifier.

The system was able to distinguish between pizzicato and sustained tones with 99% accuracy. Instrument families were recognized with 90% accuracy and individual instruments were recognized with over 70% accuracy. Perceptual tests were run by

Martin in order to compare the results of the system with the performance of humans. Only fourteen expert subjects were used for the tests and the system was found to perform better than three of them.

Traube created a system that is able to estimate the plucking point on the guitar [176, 175]. The work draws heavily on physics and signal processing but also includes perceptual tests of guitarists. The study demonstrates techniques to estimate the pluck position of the guitar based on signal processing algorithms.

Guitarists use words in order to describe the tone of the guitar in a fairly consistent manner. In fact, these words often show up on scores of guitar music. Also, relative terms, such as the brightness of the tone, are used to describe the timbre. Through perceptual tests Traube was able to make rough correspondences between the plucking point on the guitar and the types of descriptions used by guitarists.

Complex signal processing was combined with knowledge of the ideal physics in order to develop methods of estimating the plucking point on the guitar. By calculating the ideal spectra of plucking points on the guitar it was observed that the spectral shape was very similar to the impulse response of an FIR comb filter structure. The spectra could then be correlated with the ideal spectrum and a spectral centroid measurement in order to produce an estimation of the plucking point on the string of the guitar.

In his doctoral thesis, Jensen [82] presents theories for complex models of musical instruments. The initial idea for this research was to describe the transition between musical sounds. Jensen adopted the definition for the transition of musical sounds as the variance over time of pitch, loudness and timbre. Timbre then became the focus of the study.

Jensen works towards a comprehensive model of timbre by evaluating recordings of real instruments and correlating the data with research in perception. Jensen proposes three models of timbre: High Level Attribute (HLA), Minimum Description Attribute (MDA), and Instrument Description Attribute (IDA). The different models act as the basis for each successive model, i.e., the HLA model is an element to the

MDA model.

The parameters of the high-level attribute model are amplitude envelope, spectral envelope, frequency, and noise. The parameters of the minimal description attribute model are the same as the HLA model except that each attribute has a base value and a partial evolution value instead of complete envelope data. The parameters of the instrument description attribute model are the same as the IDA model except that the measurements are take in half-octave bands.

Jensen evaluates his models by demonstrating their ability to stand up to the analysis/resynthesis process. Also, the correlation between a Principle Component Analysis of the original waveform and the resynthesized waveform was performed. Listening tests were also performed in order to evaluate these models. A classification test based on one hundred and fifty samples from the McGill University Master Samples using a log likelihood of normal distributed data was conducted and made no errors.

Although the major contribution of the work is given through the process of creating the models, the models that are generated in this work do an excellent job of describing the original sound. Jensen offers models that can accurately describe an instrument tone throughout its range and timbral variations.

Klapuri contributed a significant amount of work to both music information retrieval and signal processing. He designed a system that is capable of accurately detecting the onset of discrete musical events [93]. Onset detection allows automatic transcription to be computed easily since the musical events can be segmented for recognition. The detection is done by combining onset data in twenty-one non-overlap bands. The system can detect how much time the onset takes and the intensity of each band through time.

Sillanpää produced a technical report that used aspects of Klapuri's system in order to recognize different instruments of the drumset [155]. Sillanpää's system was able to recognize the correct instrument with 87% accuracy using only a shape matrix derived from a bounded-Q transform [87]. The method was tested with two and three simultaneous instruments but was not as successful. Aspects of the system were improved upon by using high-level models as well as low-level models [156]. A metric grid was imposed onto the audio for use with a predictor to aid in the recognition rate. The system was able to recognize instruments very accurately but had difficulty when one instrument masked another (i.e., when one instrument is louder than another in such a way that the softer is imperceptible).

Further studies into drumset transcription system were done [67, 135]. The onset detection research was combined with periodicity measurements and pattern recognition with mel-frequency cepstrum coefficients as features and a gaussian mixture model classifier. Different genres of music were used in order to test the system with over twenty-eight hours of data. The system was able to accurately detect the presence of drums in a signal with 88.0% accuracy.

A novel approach was taken in the subsequent study [135]. The labelling of the instruments was based upon the placement of the instrument within the metric structure of the music. Tendencies in music could be infused in the system to increase recognition rate (i.e., the snare drum is usually played on beats two and four in Pop/Rock music).

These studies demonstrate the importance of segmenting techniques on the recognition rate. Sub-band techniques and mel-frequency cepstrum coefficients are used to great effect as features for percussive instrument recognition.

3.3 Physics

Relatively little work on the physics of the snare drum has been done when compared to a general drum. Most of the work summarizes the properties of ideal membranes and their interaction (i.e., coupling). This section will outline the major contributions to the physics of the snare drum and their relevance to the current project.

Rossing has published many important works in acoustics that focus on percussion instruments [145, 144, 41, 143]. In all discussions of drums, Rossing begins with a summary of the ideal membrane, such as the one included below.

> A membrane, like a string can be tuned by changing its tension. One major difference between vibrations in the membrane and in the string, however, is that the mode frequencies in the ideal string are harmonics of the fundamental, but in the membrane they are not. Another difference is the in the membrane, nodal lines replace the nodes that occur along the string.

The formula for calculating the frequency of the modes of an ideal membrane is:

$$f_{mn} = \frac{1}{2\alpha}\sqrt{\frac{T}{\sigma}}\beta_{mn} \tag{3.1}$$

α = the radius of the membrane (meters)

T = the tension applied to the membrane (newtons)

σ = the area density (kg/m^2)

β_{mn} = values where the Bessel function becomes zero

The rest of the information offered in his writings is observations from experiments conducted in his lab. Although this is by no means a complete model of the snare drum, Rossing mentions many observations that are worth noting.

The interaction of the modes between the drum heads is very complex. Analysis is more complex when the drum heads are considered together. Potentially there are $X_{mn} * Y_{mn}$ number of modes where X_{mn} is the number of modes of the batter head and Y_{mn} is the number of modes of the snare head. Due to coupling, certain combinations of modes will occur most often and these are the modes that Rossing presents. Asymmetrical modes occur when the top head and the bottom head are vibrating in the same mode but 180° out of phase with each other. Symmetrical modes occur when the top and bottom heads are vibrating in the same mode in phase.

The effect of the snares on the drum are briefly discussed, but Rossing makes some interesting observations. Since the snares are on the bottom head of the drum and the drum is usually struck on the top there is usually a time for the snares to activate. Rossing demonstrates this effect by showing a plot of snare drum data and points out where the snares activate. The snares add higher frequency components to the sound that are perceived as noise. The tension of the snares also effects the sound, as Rossing demonstrates by plotting spectra of snare drums hit with varying snare tension. The plots indicate that low tension produces a flatter spectrum with less defined modes and that higher tension produces a more peaky spectrum with more defined modes.

Another factor in the sound of the drum is the decay of the modes. Not all modes are created equally. When modes are excited on the snare drum they interact over time, some decay faster than others due to the interference cause by other modes. Rossing provides some data for an experiment where the modal decay rates are measured for a drum that is suspended by elastics and a drum that is on a regular snare drum stand. The drums were struck off-center and in the center. The decay rates for the drum that is on the stand is significantly higher than the suspended drum suggesting that the coupling between the rims and shell are important factors in the vibration of the drum. Whether this affects the timbre in any significant way other than the decay rate is not discussed.

The first major study to focus specifically on the snare drum was published by Charles Henzie for his PhD in 1960 [69]. The purpose of the study was to determine how amplitude and duration are coupled in the snare drum sound, and to propose how that information will be used. The study analyzed the amplitude and duration characteristics of snare drums by comparing the results of his experiments. Two snare drums and many different sticks were used in the experiment. The concept of "stroke length" [69, 18] was used to control the amplitudes of the individual strikes of the snare drum for consistency. According to Henzie, stroke length refers to the height of the stick at the top of its arc. An expert player was used to strike the snare drum and a string was used as a reference point for the stroke length.

Henzie found that there was a direct correlation between amplitude and duration. He also discovered that the type of stick had a significant effect on these factors.

Henzie's conclusions only point out these links but do not offer any scientific insight as to how or why these phenomenon occur.

Henzie's conclusions and suggestions for the use of this information are entirely for pedagogical uses. Though he was examining students the findings can be used to refine professionals and be part of formalizing technique to extend the gestural and control possibilities of drummers.

Donald Wheeler published a brief article in Percussive Notes in 1989 which outlined some experiments and the results on the snare drum [100]. In the experiment a stick machine was used instead of a human to keep the results more constant. The experiment recorded snare drums with the snares turned on and off and then analyzed the recordings by taking FFTs and displaying the results on a three dimensional plot.

The interesting aspect of this study is the discovery that whether the snare mechanism is engaged or not, the fundamental frequency of the drum remains the same. Previous theories suggested that since the snare mechanism is contacted along the bottom head of the drum across a nodal line, then when it is engaged the fundamental frequency of the snare drum will not be excited and the drum would have a fundamental frequency equal to that of its second mode. Wheeler's plots show that the fundamental mode is still excited when the snares are engaged, thus the fundamental frequency remains the same whether the snares are engaged or not.

Zhao published his Masters thesis on snare drum acoustics [196]. Presently, this is the first and only significant study on the acoustics of the snare drum. The goal of the study is to build a scientific model for calculating the modes of the drum when it is struck and then implement it in software.

Although it is known that there are coupling effects between the modes of the drum and the different parts of the drum, little work has been done to accurately describe these effects in any detail. Zhao's work aims to provide a mathematical model for these relationships.

A two-mass model was employed to describe the relationship between the batter head and snare head. The interaction between the heads and other factors (excluding the snares) was examined and thoroughly described mathematically for the first mode of vibration.

Lewis and Beckford published a brief article in *Percussive Notes* in 2000 that showed the results of some interesting experiments done with snare drum batter heads [100]. Different batter heads from different manufacturers were placed on a snare drum and fixed to a uniform tension. The drum was struck and the sound recorded onto a computer and analyzed using FFTs. The results of the study show that the main mode of the drum was prominent in all of the drums but the rest of the spectrum varied greatly. This provides some factual basis for the head making a difference in the timbre of the instrument. The increase in the content in the higher parts of the spectrum may account for the perceptual rise in pitch when the snares are engaged.

3.4 Percussion Sensor Technology Overview

There are different methods of capturing the gestural information of percussive strikes using a variety of sensors. In this section a survey of the most common techniques will be described from simple methods using a Piezo to more modern sensor technology.

What information besides strike velocity is important and how can it be used to accurately translate the gestures of an expert performer? Geometrically speaking, there is the angle of incidence, velocity of the strike, the polar position of the strike on the surface, and the number of points of contact.[1] In order to capture these gestures and translate them into data usable by a computer, we must examine the issue of latency versus information gain and the effect upon real-time performance.

A cost/benefit analysis of controllers can be very complex. The following section will give guidelines to be considered when dealing with percussion interfaces as well as providing a thorough survey of the different models currently available and their

[1] Some playing techniques, such as the rimshot, require the player to hit the drum in multiple places simultaneously.

design decisions and consequences.

3.4.1 Piezo

Piezo sensors take advantage of the piezoelectric effect whereby mechanical energy is converted to electrical. Electrical charge results from the deformation of polarized crystals when pressure is applied [140].

These sensors are often found in older and cheaper commercial percussion trigger pads. In order to be effective they must be isolated from the vibrations of nearby impacts on different pads. This isolation is never perfect and is called crosstalk. Piezos in electronic percussion devices are sampled as standard audio and therefore provide the potential for very low latencies.

An example of an interface that uses these sensor to great advantage is the Jam-O-Drum [80]. This interface uses an array of commercial Piezo-based drum pads mounted into a surface to provide a collaborative instrument/installation for use with novice players.

The largest electronic percussion instrument, the Rhythm Tree [130], uses Piezos to detect impacts. The instrument is comprised of three hundred sensors that can detect a direct or indirect hit. The sensors are embedded in a soft rubber pad that can be lit up with an LED to provide visual feedback for the performer.

3.4.2 Force Sensing Resistors

Force sensing resistors (FSR) use the electrical property of resistance to measure the pressure exerted upon the sensor. FSRs are used in experimental and commercial drum controllers. The main advantage of an FSR for a drum controller is that the sensors only pick up strikes that directly strike the surface. This results in no crosstalk between the sensors, which is a major problem with Piezo based devices. There are two types of FSRs: ones that measure the amount of pressure and ones that measure the amount of pressure along an axis, thereby yielding position. FSRs provide a standard sensor sampling, as opposed to audio rates, and are therefore less accurate in time than a Piezo.

The Drumkat[2] is an example of a popular and powerful commercial device that utilizes these sensors. The Handsonic (HPD-15)[3] is another commercial device that uses FSRs to emulate a hand drum interface. The Buchla Thunder[4] is an alternate controller with more than a dozen pads that sense both pressure and position.

The Javamug has FSRs placed ergonomically onto a coffee mug that control the randomness of techno-latin drum loops [21]. This device demonstrates a simple mapping that leverages percussive techniques to control macro-level mappings in an interesting and simple manner.

The Electronic Tabla (ETabla) [84] is an interface designed to model the traditional playing technique of the North Indian classical Tabla. It uses FSRs in strategically placed areas of a custom built head to gather gestural imprints of Tabla performance.

The Electronic Dholak (EDholak) [84] is a multi-player Indian drum controller that uses both Piezo and FSRs for gesture capturing events inspired by the collaborative nature of the Jam-O-Drum. The Piezos are used to trigger time critical events while FSRs are used to control parameters that are less time critical, but give more expressive information than with just the Piezos alone.

3.4.3 Fiberoptic Sensing Pads

In this sensor technique a network of fiberoptic sensors detect pressure as well as position. An original device using this technique is the Tactex[5] Multi-Touch-Controller (MTC) which contains a grid of seventy-two sensors that can distinguish multiple sources of pressure. The STC-1000 - a newer version of the MTC pad - has been designed by the Mercurial Innovations Group.[6] This device is a Singe Touch Controller that outputs MIDI. The Ski is an example of a Tactex pad being used as an interface for musical expression [78].

[2] http://www.alternatemode.com
[3] http://www.rolandus.com
[4] http://www.buchla.com
[5] http://www.tactex.com
[6] http://www.thinkmig.com/

3.4.4 Capacitance

The Radiodrum is one of the oldest electronic music controllers [15]. Built by Bob Boie and improved by Max Mathews, it has undergone a great deal of improvement in accuracy of tracking while the user interface has remained nearly constant. There is an antenna at each of the four corners of the drum. The drum is played with two sticks that operate at different frequencies. The radio tracking technique depends on the electrical capacitance between the radio transmission antenna in the end of each mallet and the array of receiving antennas in the drum. The drum generates six separate analog signals that represent the x, y, z position of each stick versus time.

3.4.5 Microphone

A special mention must be made of the Korg Wavedrum. Most owners of this controller claim that it is as flexible as an acoustic instrument [146]. The Wavedrum has three contact microphones underneath the drumhead. The signals from the drumhead are combined with the synthesis engine of the drum to make sound. This results in the sound changing as the excitation is moved around the drumhead. These different signals are either used to excite the synthesis engine, are passed through the DSP chain, or are passed directly to the output after completing one of the previous tasks.

Timbre-recognition based instruments use timbre as a control parameter. Previous work of the author has integrated this technology into a percussion controller [168]. The system utilizes digital signal processing and machine learning techniques to classify the timbre of the instrument. The labels correspond to different playing techniques and striking implements identified to the system from human labelled instances. Since different playing techniques produce different timbres, it is a matter of collecting training instances of the desired technique and retraining the classifier.

3.4.6 Accelerometer

The PhiSEM controllers are shaker-like interfaces that utilize accelerometers to trigger physical models of shakers [21]. These interfaces demonstrate a simple use of percussive gestures to control physical models of shakers.

A recent use of accelerometers in a percussive interface has been developed by Dianna Young.[194] This interface used two two-axis accelerometers augmented with two single-axis accelerometers. The sensors were placed inside of a pair of Bachi sticks by hollowing out the end of the sticks to relay three axes of acceleration and the angular velocity of the sticks. A Bluetooth® emitter was also placed inside of the sticks so that the performer did not have any wires impeding their gestures.

3.4.7 Infrared

The Buchla Lightning uses infrared light tracking in order to track the position of the wireless sticks in two dimensions. Each stick uses a different frequency of light so that they may be tracked independently. The lightning also includes a button on each stick that may be used to trigger events or send metadata.

The Nintendo Wii uses an infrared camera inside of the controller to detect the orientation of the controller relative to the screen to allow players to aim at objects in the games. An array of infrared LED lights is provided to be mounted atop the screen. Simple geometric analysis is performed to provide the orientation. The Wii controller has been transformed into a low-cost interactive whiteboard and a low-cost head tracking device by Johnny Lee [98], and made famous by demonstrating these adaptations at the TED conference.[7]

3.4.8 Camera Tracking

While there are very few controllers that have employed video tracking for capturing percussive gestures there has been some interesting research on these motions using video capture. Sofia Dahl has done significant amount of work analyzing the motions

[7]http://www.ted.com

of drummers and percussionists using video capture in addition to motion capture via a Selspot system.[8] [24]

Motion capturing systems such as VICON[9], allow for more precise collection of human movement. Experiments with these type systems for musical gesture analysis is shown in [137]. Most motion tracking systems work at approximately 30 frames per second, therefore providing high latencies for percussive applications. Some systems provide cameras with high frame rates but they are usually extremely expensive and not portable.

3.5 Design Evaluation

By observing something you change its state. This is a consideration when designing any computer or sensor based instrument. For percussion instruments there are three main ways of augmenting the performance to capture the gesture: placing sensors on the drum, placing sensors on or in the stick, or detecting the motion of the body.

Throughout this section I will reflect upon the interfaces and sensors mentioned above and offer methods and considerations for evaluating a complete capture system. The needs of any project have to be assessed before the information provided here is to be of any real value. The end goal of any system should be to provide maximum control with minimal inhibitions without compromising the potential for advanced techniques. Percussion instruments are excellent examples of simple interfaces that retain the potential for virtuosity.

The bandwidth of a percussionist is a major consideration. The world record for fastest single stroke roll is currently one thousand forty-nine strokes per minute[10] which translates into approximately twenty Hertz. This only represents a single stroke roll; double stroke rolls have been recorded with up-wards of fourteen-hundred strokes per minute. The percussive interface must be able to capture these events with enough resolution to extract the desired information, such as: position, velocity,

[8] http://www.innovision-systems.com/
[9] http://www.vicon.com
[10] http://www.extremesportdrumming.com/

and angle. The sampling rate must be high enough to resolve the complexities of the motion of stick between the strokes.

3.5.1 Drum Surface

An augmented drum offers many possibilities for freedom but also many design problems. When embedding sensors into or on the drum they must be protected from the impact force of the striking implement. Many commercial products achieve this by using rubber pads with the sensor embedded inside or using a foam bed lying underneath the drumhead. This yields the possibility to use many different implements, thus allowing the performer to use their preferred traditional stick.

Wiring is a consideration in this case. In commercial devices the analog signal from the sensor is transmitted to a drum brain for processing. Another approach is to utilize the space inside of the drum shell to hold the wiring and circuitry and providing the processed data as output (the eTABLA is an example of this technique).

By placing sensors on the drum the vibrational properties of the instrument are altered. When fabricating a synthetic drum, not intended to make sound itself, this is acceptable but when extending a traditional drum with sensors, this can be a problem.

The Piezo offers a simple way of capturing percussive gestures. They were the first sensors used for this application and have been widely used in many of the available commercial interfaces. Using one Piezo it is possible to capture the timing of strike and a plausible indication of velocity. The disadvantage of these sensors is that you cannot capture position with a single sensor. It is possible to find position using a network of four Piezos, as has been shown [134].

FSRs give more precise pressure data than Piezos and some types also give a position parameter as well. Since you must touch an FSR in order to trigger it, it is necessary to have either a large sensor with a custom shape or a small surface that accommodates the shape of the sensor. Again, since it is triggered by pressure they greatly reduce crosstalk.

The main makers of the fiberoptic sensing pads are Tactex and the Mercurial Group. These interfaces come only as a prepackaged device. In order to incorporate them into new instruments they must be disassembled, as in the The Ski [79]. These devices provide position in two dimensions as well as pressure data. The MTC pad connects via serial to a computer running Max/MSP. There is a custom object that comes with the interface to connect directly to Max/MSP thus allowing easy mapping from physical gesture to messages in the program. The main advantage of the MTC is that it is capable of discerning different touch points, allowing for multiple simultaneous striking events. Besides being a prepackaged product, the main disadvantage for percussion is the inherent latency of the sensors.

Capacitance offers continuous data in three dimensions. Having the continuous z parameter offers a great deal of creative possibilities. There is a nonlinear response when using this technique that makes it difficult to accurately track absolute position. A major disadvantage is that both the drum *and* the sticks have to be augmented in order to capture their position relative to each other.

The Korg Wavedrum provides an interesting combination of controller and effects processor. The acoustic sound received by the embedded microphones is used to excite physical models. The result is that the performer is not limited by the data that the interface can capture. If the gesture makes sound, which it should in order to produce music, then it will be captured by the microphones and transformed into an output signal.

A new approach to capturing gestural data is the timbre-recognition based instrument. The main advantage is that the overhead for setting up this system is minimal; it only requires a microphone. Many current musical situations are recorded and most musicians have become accustomed to having their instruments captured by microphones. The precision of this system is constantly being improved. The main disadvantage is that it is yet be seriously tested in real musical situations. While it performs well in a closed studio environment and reasonably well with the addition of noise, the proper compensation for the interference of a musical environment seriously complicates the situation.

3.5.2 Stick Augmentation

Embedding sensor systems within a stick allows performers to play any drum, traditional, modern, augmented, or even thin air. However, the modified stick will never feel the same as a traditional one, with added weight from the sensor packaging. Transmitting data using wires obviously proves to be cumbersome. Wireless transmission may be an impressive laboratory stunt, but can result in loss of data and non-reliable systems not usable for performance on stage.

Accelerometers on sticks can aid in capturing kinematic performance gestures, though data will be noisy and cannot be used to derive accurate position of strikes. Adding a gyroscope into the mix, as accomplished in AoBochi, provides a reference in which to derive position, making for a more informative and expressive system.

The use of magnetometers as seen in the Vodhran captures orientation with respect to the north pole. These sensors are usually surface mount sensors and require custom PCB boards to built for experimentation, making them less easy to use than sensors like FSRs and Piezos.

Infrared light tracking allows for an elegant solution to wireless sticks. As the performer gets further away from the receptor the field of capture gets larger, the largest possible field being twelve feet high by twenty feet wide. One disadvantage of the Lightning is that when multiple instruments are put in the same space, a great deal of interference occurs and it is very difficult to control the system. The Lightning comes as a prepackaged product that is easily adapted to other tasks but is limited by its MIDI output.

3.5.3 Detecting Body Movement

Analyzing human body movements during drum performance offers another avenue of detecting a percussive gesture. A traditional technique is to analyze video camera footage which is completely unobtrusive requiring no sensors attached to sticks or drums. However the frame rate of the video camera (typically 30–120 frames per second) along with time needed for post processing does not allow for fast percussive interaction.

Using motion capture systems such as VICON, more precise gestures can be captured at high sample rates. A marker is placed on each of the joints of the hand, and in key positions along the length of the arm. These points are captured by six cameras at a high sample rate providing accurate temporal and spatial data that is superior to most sensor interfaces. However the system is quite expensive.

The major problem with video cameras or the VICON system is that camera systems are difficult to use on stage. Variable lighting, camera orientation, and calibration are all issues which can end up being a nightmare during sound check, and even worse, during a live performance.

Building devices that embed sensor systems directly to the skin or within clothing, as mentioned in Paradiso's work [131], are promising. For accurate percussive gestures, however, they might prove useful only in conjunction with other methods.

3.6 Summary

This chapter has presented a summary of the current state of research and technology surrounding electronic percussion. The computer music research community has done a significant amount of research into musical signal processing, music information retrieval, and acoustical analysis of musical instruments. These bodies of research can be leveraged when developing a new percussive interface, as will be demonstrated in the following chapters.

A survey of the sensor technologies currently available allows an informed decision on the type of sensors to pick for the E-Drumset. The issue of detecting percussive gestures becomes one of three solutions: augmenting the stick, augmenting the surface, or augmenting the player. Solutions most likely to adopted by the larger community are solutions that augment the surface because drummers are able to play uninhibited by additional technology or be restricted to a stick that has been pre-augmented. The Korg Wavedrum and the Simmons SDX are excellent examples of augmented surfaces that have become very popular devices despite limited availability.

CHAPTER

FOUR

The E-Drumset

Imitation is the sincerest of flattery.

– Charles Caleb Colton

4.1 Introduction

The E-Drumset is collection of software and hardware that allows alternatives to the acoustic components of the drumset. By mixing the components into different combinations it can be personalized for each player. The hardware can be any sound source that sends a signal into a computer so that it may be processed by the software.

This chapter will outline the basic concepts of the E-Drumset components in terms of hardware design, software design, and typical function. Synthesis methods utilized will be described and demonstrations of implementations with be presented.

4.2 Design Criteria

A survey of standard, extended and contemporary drum and percussion techniques can be found in Appendix D. Put simply, the E-Drumset was designed to allow for most, if not all, available acoustic gestures to be available and allowing for the potential for new gestures to evolve. The previous two chapters illustrate the research conducted that was used to determine the design criteria for the E-Drumset. The following is a general list of design requirements that was generated before developing the E-Drumset.

- Position Tracking
- Flexibility
- Expandability
- Programmability
- Openness

The main goal of the project was to allow for position tracking. The solution that was discovered was implicit position tracking, which will be discussed further through the chapter. Flexibility and expandability are related items. The final product had to be able to function in many situations and it also had to be expandable for large setups. Programmability became a requirement as many commercial interfaces have little to no ability for the user to change their sounds or mappings. The final criteria was the design had to be such that anyone would be able to replicate the results. A more thorough examination of design requirements and evaluation is discussed in the following chapter.

4.3 E-Drum

The E-Drum utilizes the concept of an acoustically driven physical model as the primary form of sound generation. Typically physical models are driven by synthetic excitations. By modifying the algorithms to accept audio input acoustically excited physical models are created Experiments using timbre-recognition based instruments will also be utilized to discover the advantages of the contrasting techniques and their different combinations. See figure 4.1 for the general software overview of the E-Drum.

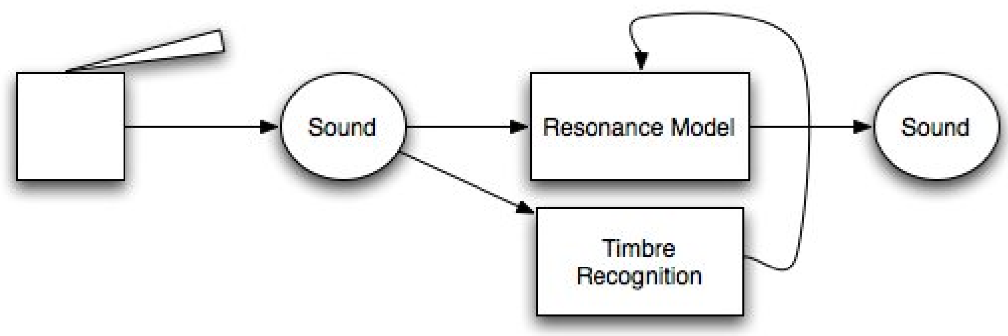


Figure 4.1: A diagram of the signal flow in the E-Drum software.

4.3.1 Hardware

The E-Drum was tested with different versions of hardware. Because of the universal practice of using piezo sensors in electronic drums the E-Drumset software works with any commercial electronic percussion pad. There are many types of pads on the market, but they consist of two main categories: rubber pads and pads with a mesh head. While both types of pads will give a response that can be modified through the synthesis algorithms, the mesh heads allows for greater response to brushes or scraping type sounds.

4.3.2 Mapping

Mapping drum gestures with standard drum interfaces is generally done with a one-to-one correspondence between gesture and result, usually only utilizing amplitude as a parameter for modifying the sound. By using the gesture recognition techniques proposed, the issues of mapping disappear as the instrument responds to the physical excitation and is modified by the synthesis algorithm. Timbre recognition techniques provide the strike position data that can be used modify the parameters of the synthesis without using a metacontroller, such as knobs or sliders.

This model allows for the use of brushes and scraping gestures that standard drum controllers are not able to capture. The dynamic possibilities of an instrument that morphs itself when it is struck will allow for different types of sound palettes as well as the flexibility to move quickly between them.

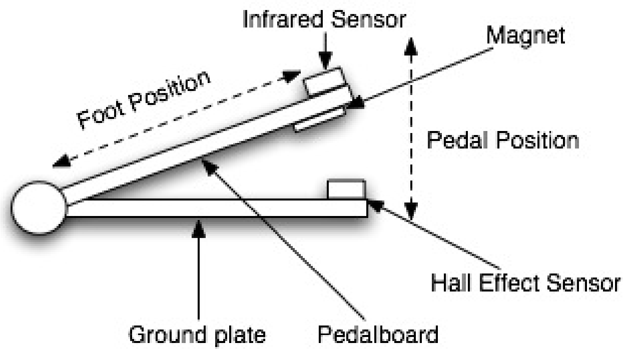


Figure 4.2: A Diagram of the E-Pedal and components.

4.4 E-Pedal

The E-Pedal is actually a sensor configuration for drum pedals. It consists of two range sensors that detect the pedal position, in the standard dimension, and the foot position along the pedalboard (see Figure 4.2). The idea with this simple package is that it may be attached to either a practice pedal[1] or to an acoustic pedal to create a meta-instrument or a hyperinstrument.

4.4.1 Hardware

The hardware package consists of two infrared distance sensors placed on the board. New sensors from Sharp - the GP2D120 - provide an effective range of 4 – 30cm, a more appropriate range than the other models in their line. A small chassis is fitted to the sensors so that the foot never impinges upon the minimum effective range and so that they may easily attached to the pedals. Small clips were attached to the chassis to attach them to acoustic or practice pedals.

4.4.2 Mapping

The algorithm used to detect discrete events is based on the algorithms proposed by Neville et al. [121]. The software uses statistical methods to determine the variance in the signal in order to adapt the pedal to different ambient light conditions. The

[1] http://www.hansenfutz.com

velocity and acceleration are measured when the pedal crosses a minimum point on the z axis. Thresholds are set for the velocity and acceleration to cross in order to trigger an event. The event can be mapped to a drum hit or to a toggle, etc.

An interesting possibility with the E-Pedal is to attach it to an acoustic drumset and thereby provide more control possibilities to the acoustic drummer. Another idea is to have the footboard position act as either a bass drum or hihat pedal (or both simultaneously) and have the foot position act as a standard volume pedal. To achieve this with current commercial technology requires the use of a pedal for each of the tasks mentioned.

4.5 E-Cymbal

The E-Cymbal combines an acoustic cymbal with a microphone and software to create a new hyperinstrument. The E-Cymbal provides the E-Drumset the facility for long notes without utilizing a drum roll. The software layout for the E-Cymbal can be seen in Figure 4.3. The E-Cymbal allows the drummer to utilize all of the standard playing techniques of the acoustic cymbal but provides new sound feedback and allows the player to explore the cymbal in a new way.

In an acoustic drumset mindset the cymbals provide long notes to the musician. In electronic music long notes can often last the course of a piece, in the case of drones or pads. The use of feed-backing delays and spectral delays have been investigated to create very long, yet varying textures. The E-Cymbal rounds out the sound palette of the E-Drumset, providing the ability to play very long and very short notes intrinsically.

4.5.1 Hardware

The hardware for the E-Cymbal is simply a microphone; a contact microphone or a standard microphone will suffice to capture the sound of an acoustic cymbal. The main advantage of the E-Cymbal is that it is an easy conversion from an acoustic cymbal to an E-Cymbal. As a result a number of cymbals were converted by the

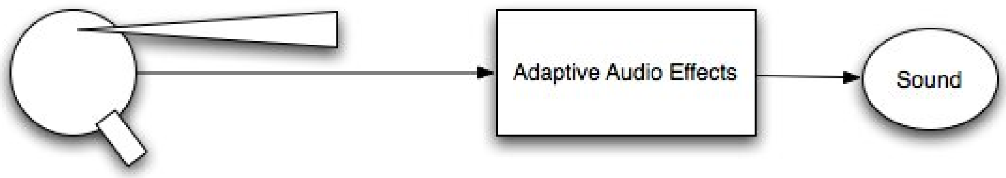


Figure 4.3: A Diagram of the E-Cymbal.

author. The mounting of a microphone to a cymbal can be done in a non-permanent way to avoid potentially damaging the original instrument.

4.6 A Note about Latency

In recent times latencies have lowered as computer power has increased. When one gives a talk with a new interface often the question of latency is raised, usually aiming for the smallest amount of time possible. Teemu Mäki-Patola [120] delivered a paper explaining that latency is relative and gave an example of a piano player: generally, the piano requires latencies that are very low but during slow chorales latencies of up to 100ms were tolerable to subjects.

In electronic music performance there are two latencies: the latency of the digital system and the latency of the acoustic system. By disconnecting the sound from the generating source, electronic music is able to regain some latency that is lost with purely acoustic systems. A musician may monitor the sound via headphones, reducing the acoustic latency from the instrument to the ear, and place the speaker system at the front of the stage. This allows the player to save a millisecond or two. The audience saves a few milliseconds allowing the visual elements on stage to be closer to the sound stimulus.

The latency available in modern digital signal processing systems is either very small or adjustable to be small. Low latencies are crucial for a number of percussion gestures but not all. While performing on the E-Drumset a number of drone or steady state sounds were explored that were inspired by the idea of the chorale pianist. Latency was not an issue for these types of sounds.

During the course of the study various software packages were utilized. Packages that utilized RTAudio[2] were configurable to give imperceptible latencies. The main latencies in these cases came from the synthesis methods used. When utilizing filters to create pitches, especially recursive filters, latency would accrue. A perceptual solution is to have a multiple synthesis chain where one chain is noisier for attacks and the other is pitchy for a steady state portion of the sound, much like an acoustic drum.

4.7 Dynamic Range

Dynamic range is another crucial factor in a percussion instrument. A drum under the control of a master is able to be imperceptibly soft and damagingly loud. The available dynamic range of the E-Drumset is quite large but is dependent on two factors: noise floor and the headroom. Tuning these parameters appropriately will allow for maximum dynamic range.

In a digital system each bit may represent 6DB of dynamic range. Current standard audio interfaces utilize 16-Bit quantization allowing for 96DB of possible dynamic range. For most sound cards used in the experiment the noise floor was 2 Bits. The difference between the noise floor and the quietest possible gesture was another 2 Bits.This leaves 12 Bits of resolution left or 72DB of dynamic range. See figure 4.4.

The range of an acoustic drum is somewhere in the range of 60 decibels if we assume that a quiet sound is 40 decibels and a loud sound is 100 decibels. Therefore, the dynamic range of the E-Drumset is greater, except that in practice it often is not. The input stage offers suitable dynamic range but it is relative to the line level signal. When the signal is amplified often the dynamic range is shrunk due to the need make the sound audible. Under ideal conditions, where the sound system is loud enough to mask the acoustic sound of the drum pads, the full dynamic range is available.

[2] http://www.music.mcgill.ca/ gary/rtaudio/ - see Glossary for more information.

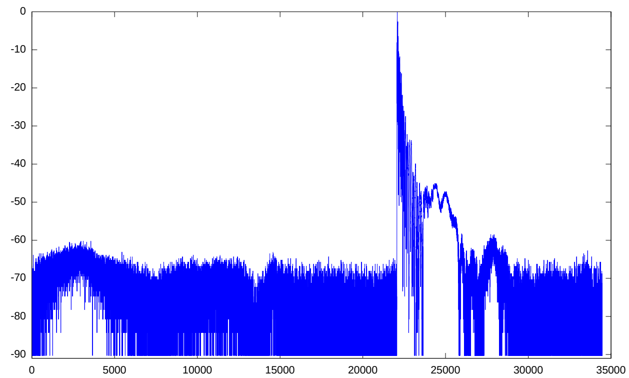


Figure 4.4: Plot of minimal gesture and maximum gesture input.

4.8 Personal Setup

The first conception of the E-Drumset was to make an instrument that utilized various components but remained in static configuration for myself, to allow me to get to know the components in a specific context. This setup was quite large and consisted of the following components:

- 4 E-Drums
- 6 E-Pedals
- 2 E-Cymbals
- 1 Snare Drum
- 2 Cymbals
- 2 Expression pedals
- 4 Sustain Pedals

This setup changed over the course of preparing the dissertation. The configuration listed above (see Figure 4.5) remained static for nearly two years but eventually new performance opportunities arose that didn't call for all components. See Figure 4.6 to see the E-Drumset in concert.

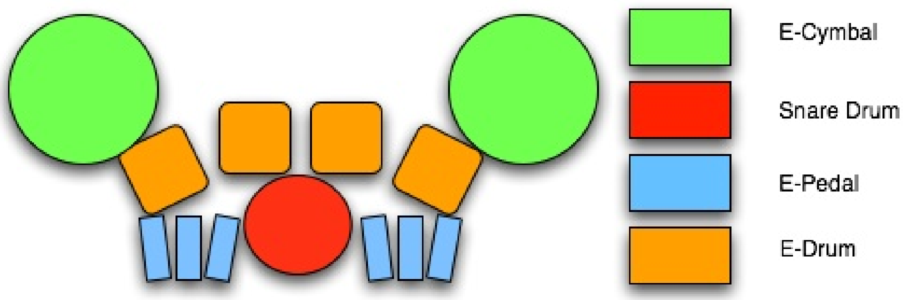


Figure 4.5: Initial E-Drumset layout.

Figure 4.6: The author performing on the E-Drumset.

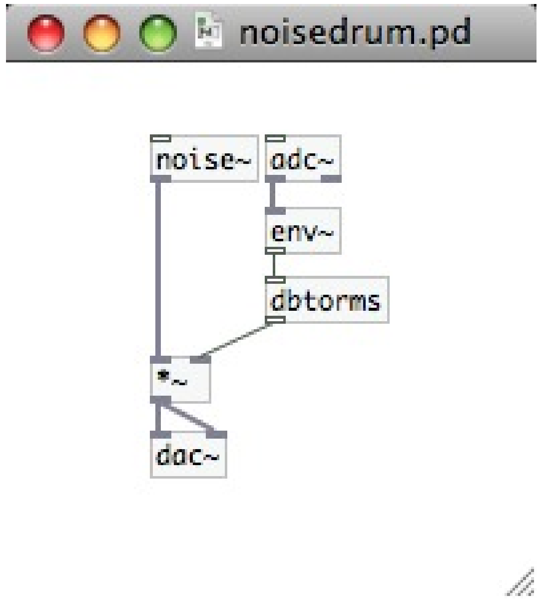


Figure 4.7: Simple Pure Data patch using the amplitude envelope of the input.

4.9 Modifying Algorithms

The concept behind the synthesis algorithms utilized in the E-Drum is to modify existing synthesis algorithms by modulating components with audio from the microphone. In some cases this is raw audio and in other it is the envelope of the strike. Figure 4.7 shows a simple envelope tracking program written in the Pure Data[3] multimedia programming language. The envelope of the input is used to control the volume of a white noise generator.

4.10 Synthesis

4.10.1 Onset Detection

An audio stream can be segmented into discrete events, such as in the MIDI paradigm. Having a segregated stream has many advantages: classification is easier because the beginning and end of an event are known (or assumed), the computer saves processing power by not classifying needlessly. Onset detection is important for percussion

[3]http://www.puredata.info - see Glossary for more information

because although there are many different algorithms that may detect the onset equally well, they all detect the onsets at different points in the envelope. The algorithm that maintains accuracy but detects the onset the soonest is to be favoured.

The simplest method of event detection is to use threshold crossing on the amplitude envelope of the sound. For the E-Drumset this is somewhat acceptable because all the drums are segregated into discrete signals. With other hardware implementations where there is more crosstalk or background noise a more robust solution is needed. The two main methods of onset detection utilized in the E-Drumset are discussed below. In the future other onset algorithms will be investigated for potential improvements.

Spectral flatness is a very effective measure of onsets as it is a measure of noisiness to pitchiness within a signal. Spectral flatness is computed by taking the ratio of the geometric mean to the arithmetic mean. Miller Puckette's bonk object for Max/MSP and Pure Data [139] uses a spectral flatness measure using 11 filters in a Bounded-Q configuration in order to detect onsets.

$$\frac{\sqrt{\prod_{i=0}^{n-1} x_n}}{\frac{1}{n}\sum_{i=0}^{n-1} x_i}$$

Marsyas[4] offers many peak picking algorithms appropriate for event detection. The PeakerAdaptive marsystem unit in Marsyas tracks local maxima and adapts to current signal levels. PeakerAdaptive has the flexibility to add a relax period or hysteresis for determining when the algorithm is begin looking for new onsets.

4.10.2 Additive Synthesis

Additive synthesis is the superposition of oscillators or unit generators. Amplitude of the input was used to modulate the synthesis algorithm, allowing the player to play the tones intuitively. The amplitude could modulate the entire collection of oscillators or have scaled effect on each oscillator. Additive synthesis offers the potential for complete resynthesis if enough overtones with appropriately variable

[4] http://www.marsyas.sourceforge.net

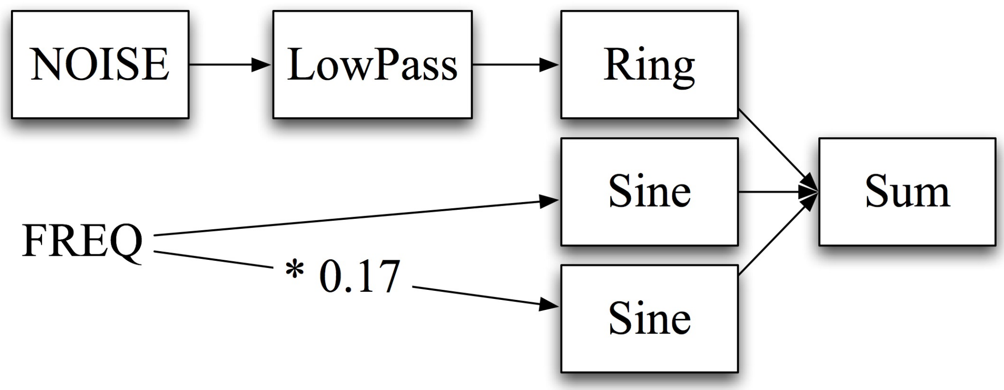


Figure 4.8: Signal diagram of Risset drum algorithm.

amplitudes are used. For a complete resynthesis an infinite amount of partials is necessary, though approximation is possible and often convincing.

Similar to Figure 4.7, the global envelope of the additive synthesis may controlled by the amplitude envelope of the drum.

Risset Algorithms

In 1963 Jean Claude Risset compiled a library of synthesis algorithms that he had developed in the MUSIC V programming language. A number of percussion algorithms were developed. The bell algorithm and the drum algorithm were ported to the E-Drum software. Originally, the algorithms were triggered with impulses. The E-Drum software modifies these algorithms to use an audio input so that a drum pad may be processed through the Risset models. Figures 4.8 and 4.9 depict the original Risset algorithms as outlined in his catalog of synthesis[141].

4.10.3 Physical Modelling

Physical Modelling synthesis can be done in two main ways: digital waveguides [157] or mass-spring simulations. For this work the digital waveguide method was employed due to the ease of implementation and the efficiency of the algorithm. The low computational requirements allowed for patches with many variations, thus

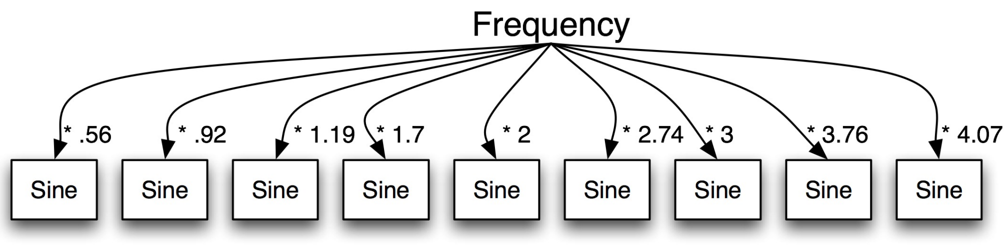


Figure 4.9: Signal diagram of Risset bell algorithm.

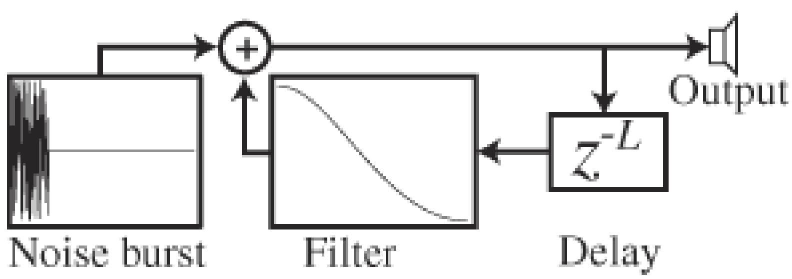


Figure 4.10: Signal Diagram of the Karplus-Strong algorithm.

enabling more experimentation. Most simulations utilized one or two dimensional waveguide models.

The classic Karplus-Strong algorithm is a recirculating delay line with a moving average filter (See Figure 4.10). The noise generator shown in the diagram was replaced with the audio input of a drumpad connected to a computer. This algorithm was one of the first to be implemented in the E-Drumset software. The original expectation was that the timbral variance of the output would be minimal due to the filter and that it would sound like any other Karplus-Strong. Timbral variation was noticeable and quite significant considering the limited nature of the algorithm.

The filter can be replaced by a number of other filters available with modern software. Band-pass filters tuned to the fundamental frequency of the delay lines create a more focused attack to the algorithm. Interesting effects are achieved by tuning the filters lower or higher than the fundamental frequency. The attack can be made more noisy by lowering the Q factor of the filters.

Sustain can be modified by controlling the gain of the reflection filter. Alterna-

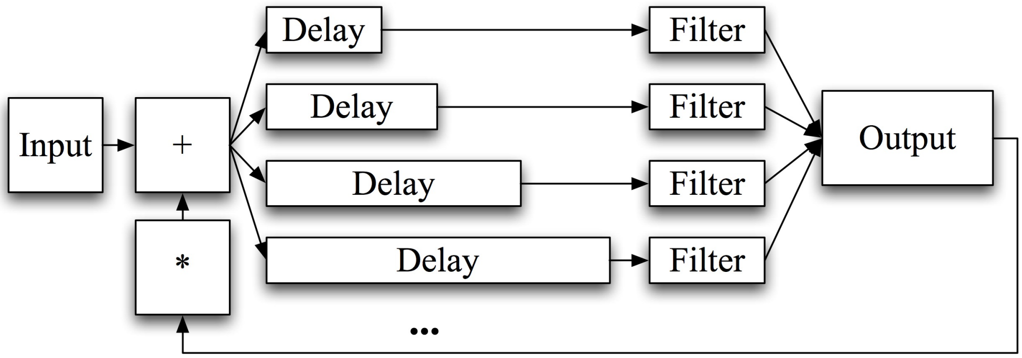


Figure 4.11: Signal Diagram of a Banded Waveguides Network.

tively, the filter used may be set to be a ringing filter that will extend the duration of an excitation. Modifying the duration of the event allows for greater variability in the control of sound produced by this algorithm.

The banded waveguides concept is an array of single dimensional waveguides with a summed feedback section. They were developed by Georg Essl and Perry Cook [36] and have been used for percussion and friction synthesis. See Figure 4.11 for a system diagram of a banded waveguide network.

The length of the delay lines corresponds to the fundamental frequency of the overtone. The filters in each tap of the waveguide control the amplitude and timbral decay for each overtone. The algorithm is a more specialized mutli-dimensional Karplus-Strong algorithm. To allow for audio input into the system the noise generator is substituted for a generic audio input allowing for drum excitation.

Waveshaping

Waveshaping is a classic synthesis operation where a table is stored and the sample value of an audio signal is used to reference a lookup table. Waveshaping may also be performed by processing an audio signal through a function. Simple waveshaping may be performed by using a sine function. Figure 4.12 shows an implementation of waveshaping in Pure Data.

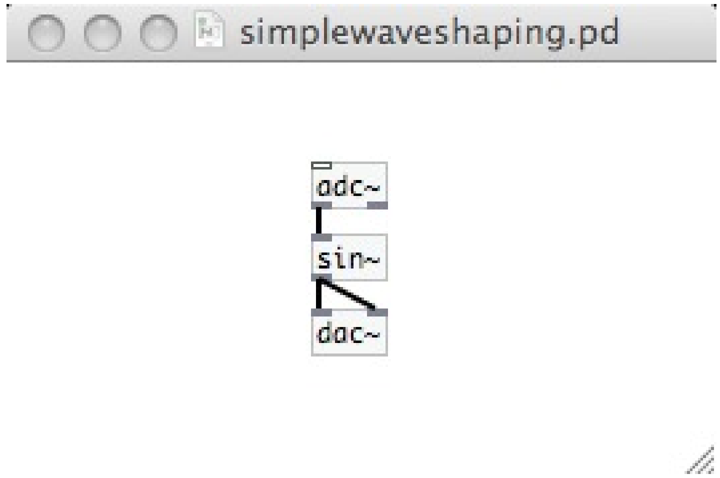


Figure 4.12: A simple Pure Data patch for sine based waveshaping.

$$y[n] = sin(x[n])$$

Interesting effects can be achieved by putting arbitrary sound files into the waveshaping lookup table. Since the operation is only a table lookup it is computationally cheap.

Supercollider[5] utilizes a distortion function that is direct manipulation utilizing a hypertangent operation. In practice the hypertangent is stored in a lookup table and the incoming signal is referenced against it.

$$y[n] = tanh(x[n])$$

Many other waveshaping functions are available freely in code or pseudo-code at musicdsp.org. During the course of the project many of these algorithms were implemented as part of the E-Drumset codebase.

[5]http://www.supercollider.sourceforge.net - see Glossary for more information.

Spectral Rotation

Spectral rotation is the offsetting of FFT transforms. The latency of the transform depends on the length of the FFT window used. A ChucK[6] program follows that rotates the input by 300 positions, nearly half of the frequency resolution in this example.

```
adc =>  FFT fft => blackhole;
IFFT ifft => dac;
1024 => fft.size;
300  => int rotation;
fft.size()/2 => int halfFFTlength;

// use this to hold contents
complex input[halfFFTlength];
complex output[halfFFTlength];
while( true )
{
    // perform fft
    fft.upchuck();
    // get contents of fft
    fft.spectrum( input );

    for ( 0 => int i; i< halfFFTlength ; i++)
    input[(i + rotation) % halfFFTlength] => output[i];

    // perform ifft
    ifft.transform( output );

    // advance time
```

[6]http://www.chuck.cs.princeton.edu - see Glossary for more information.

```
    fft.size()::samp => now;
}
```

Spatialization

Modern recording practice generally utilizes one of two models for spatializing drums across a stereo field. The instruments of the drumset are moved from left to right in the configuration of the physical drumset. The divergence in approaches occurs when the spatialization is oriented from the audience point of view or the drummer's point of view. The center instruments, the snare drum and bass drum, remain in the center but the hihat may be on the left side for the drummer viewpoint or the right side for the audience view point.

In the E-Drumset the toms are arranged in a non-standard setup. The spatialization of the toms may be arranged in a number of configurations. There are two sub-headings for spatialization: orientation and taxonomy. Orientation is merely a flipping of the stereo image. Taxonomy changes the spread from being the physical locations of the drums or the descending pitch model. In the descending pitch model the toms are arranged from one side to the other in order of their pitches.

Figure 4.13 illustrates descending pitch model and the physical model for arranging the pitches on the four toms of the E-Drumset. Typically the toms are arranged from one speaker to another from high pitch to a low pitch. This can be confusing for an audience listening to the E-Drumset because the lowest two pitches are on opposite sides of the kit. Therefore two methods of panning can be employed, descending pitch or physical orientation.

Morphing

The morphing of sounds into each other or from one to the other has been an interesting research area in recent times. Modal synthesis algorithms were implemented during the project but the interesting part of the project was choosing which overtones to emphasize. Collections of overtone ratios were collected from various articles

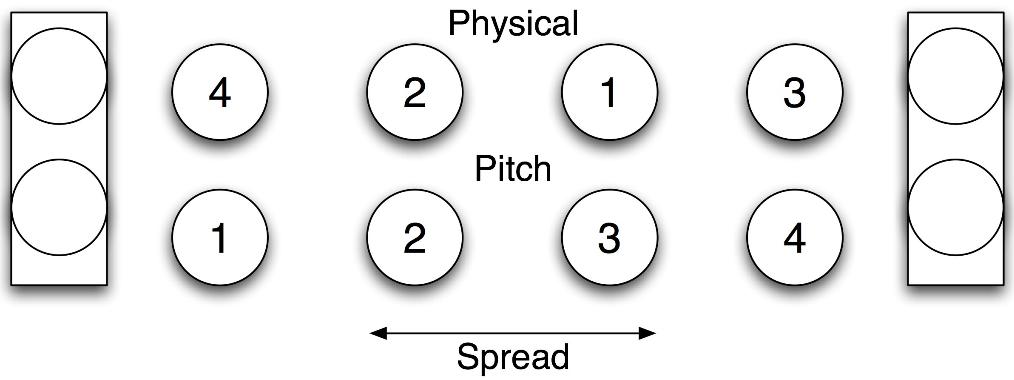


Figure 4.13: An illustration for the tom spatialization models.

on synthesis and physics. Morphing was achieved by doing linear interpolation between the overtone ratios. Various "hybrid" sounds were created by choosing a fundamental frequency for the sound and then determining the ratio of overtones. See Figure 4.14 for a screenshot of a Pure Data patch that implements modal synthesis morphing.

4.11 Conclusion

A number of synthesis methods were explored with the E-Drumset. The only method that was deliberately excluded from study was convolution due to the thorough coverage in the application to percussion interfaces by Roberto Aimi. Applying these algorithms to percussion has yielded extensions to the original algorithms. Acoustically excited physical models are simply established algorithms with new inputs but the result is a new field of sound.

The general design goals outlined at the beginning of the chapter are addressed throughout this chapter. Flexibility is achieved by allowing for the number of synthesis algorithms explored in this chapter. Expandability is inherent in the E-Drumset by its modular nature. Programmability and openness are demonstrated through the software practice of the author. Position tracking is outlined in the following chapter.

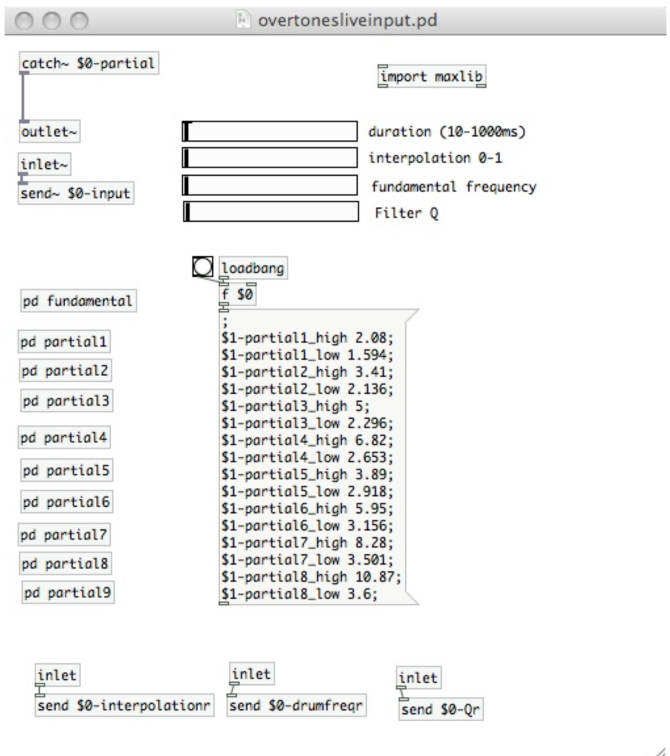


Figure 4.14: A Pure Data patch to interpolation between an ideal membrane and ideal free plate.

The design principles outlined are intended to provide the E-Drumset with the highest degree of success as a real instrument. The following two chapters address the success of the design through experiments to determine accuracy and descriptions of the E-Drumset in practice.

CHAPTER

FIVE

Experimental Results

In theory, there is no difference between theory and practice. In practice, there is.

– Yogi Bera

5.1 Introduction

During the course of study concepts of electronic drumming were investigated. Analysis and synthesis software was devised and tested in many different environments. The following section outlines the results of the software in various experiments. This chapter will explore a qualitative evaluation of design criteria and a lengthy exploration of the results of the indirect acquisition of percussive gestures.

5.2 Qualitative Evaluation

The design criteria for a new interface have been discussed (see Section 4.2). There are a number of other electronic drum interfaces that improve on standard designs. As representative examples the Radiodrum, Wave Drum, Snare Drum and the E-Drumset were evaluated with respect the design criteria (see table 5.1).

Affordances refer to the ability of the interface to allow for flexibility beyond its original purpose or primary function [49]. The Radiodrum is the only interface that

Device	Affordances	Programmability	Implicit	Explicit	Free Space
Radiodrum	No	Yes	No	Yes	Yes
Wave Drum	Yes	Yes (*)	Yes	No	No
Snare Drum	Yes	No	Yes	No	No
E-Drumset	Yes	Yes	Yes	Yes	No

Table 5.1: A comparison of the E-Drumset with current models.

does not allow for affordances. The Radiodrum has only one type of output, being the positional information of the sticks above the surface. This information can be converted into strike data that can be used to trigger events.

Programmability is the ability to reprogram the interface. Of course, the Snare Drum cannot be programmed (though it can be tuned and adjusted). The E-Drumset and Radiodrum are both arbitrary controllers that can map their data to various outputs. The Wave Drum is an all-in-one unit that has the synthesis unit embedded into the controller. An optional programmer was available for the Wave Drum that would allow for the presets to be modified. Although reprogrammability is available for the Wave Drum the external programmer was necessary to access this feature.

Implicit is short for implicit position tracking that occurs when the timbral variations of the interface correlate with the variation in position of the strike. Being a controller, the Radiodrum does not offer this functionality.

Explicit refers to explicit position tracking where the interface outputs a data stream referencing the position of the sticks. Having no analysis component, the Snare Drum does not offer this functionality.

Free Space refers to the ability to perform gestures in free space and have them recognized by the controller. This is a feature of an active controller and the only active controller surveyed is the Radiodrum.

The E-Drumset lacks free space gestures but fulfills other design requirements. Most importantly, it allows for drummers with an existing electronic drumset to connect it to a computer and gain all the new functionality described in this thesis. No other controller or electronic drum interface has this ability.

5.3 Surrogate Sensors for Detecting Drum Strike Position

There are two main approaches to sensing instrumental gestures. In indirect acquisition, traditional acoustical instruments are extended/modified with a variety of sensors such as force sensing resistors (FSR), and accelerometers. The purpose of these sensors is to measure various aspects of the gestures of the performers interacting with their instruments. A variety of such “hyper- instruments” have been proposed [103, 84, 194]. However, there are many pitfalls in creating such sensor-based controller systems. Purchasing microcontrollers and certain sensors can be expensive. The massive tangle of wires interconnecting one unit to the next can get failure-prone. Things that can go wrong include: simple analog circuitry break down, or sensors wearing out right before a performance forcing musicians to carry a soldering iron along with their tuning fork. However, the biggest problem with hyper instruments is that there usually is only one version. Therefore only one performer, typically the designer/builder, is the only one that can benefit from the data acquired and utilize the instrument in performances. Finally, musical instruments especially the ones played by professionals can be very expensive and therefore any invasive modification to attach sensors is bound to be met with resistance if not absolute horror.

These problems have motivated researchers to work on indirect acquisition in which the musical instrument is not modified in any way. The only input is provided by non-invasive sensors typically one or more microphones. The recorded audio then needs to be analyzed in order to measure the various desired gestures. Probably the most common and familiar example of indirect acquisition is the use of automatic pitch detectors to turn monophonic acoustic instruments into MIDI instruments. In most cases indirect acquisition doesn’t directly capture the intended measurement and the signal needs to be analyzed further to extract the desired information. Frequently this analysis is achieved by using real-time signal processing techniques. More recently an additional stage of supervised machine learning has been utilized in order to "train" the information extraction algorithm. The disadvantage of indirect acquisition is the significant effort required to develop the signal processing algorithms. In addition, if machine learning is utilized the training of the system can be time consuming and labor intensive.

This approach of using direct sensors to "learn" indirect acquisition models has some nice characteristics. Large amounts of training data can be collected with minimum effort just by playing the enhanced instrument with the sensors. Once the system is trained and provided the accuracy and performance of the learned surrogate sensor is satisfactory there is no need for direct sensors or invasive modifications to the instrument [167].

The traditional use of machine learning in audio analysis has been in classification where the output of the system is an ordinal value (for example the instrument name). As a first case study of the proposed method a system is described for classifying percussive gestures using indirect acquisition. More specifically the strike position of a stick on a snare drum is automatically inferred from the audio recording. A Radiodrum controller is used as the direct sensor in order to train the indirect acquisition. Regression is also explored as a method of clasification, which refers to machine learning systems where the output is a continuous variable. One of the challenges in regression is obtaining large amounts of data for training, which is much easier using the proposed approach.

5.3.1 Radiodrum

The Radiodrum was used as the direct sensor to be removed once training was complete. A drum pad was placed on top of the Radiodrum surface and the sticks of the Radiodrum were used to capture position data. MIDI data was captured from the Radiodrum.

The diameter of the drumpad was 8 inches. The drum was struck at the center and the edge, a distance of 4 inches. The maximum MIDI value was 75 and the minimum MIDI value was 59. When hitting the center of the drum the mean MIDI

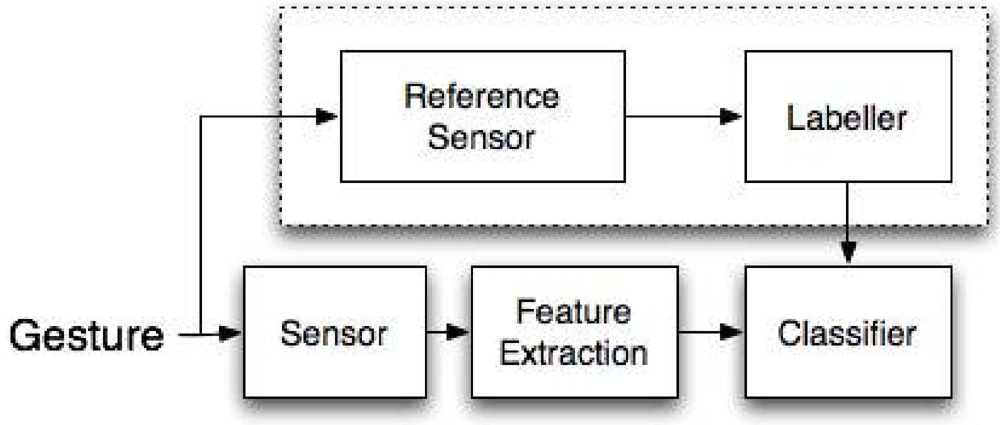
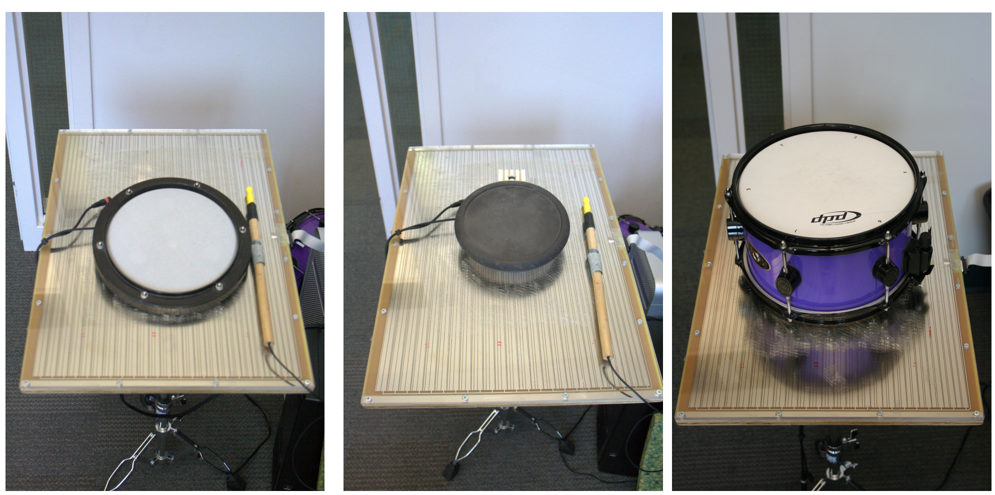


Figure 5.1: System Diagram of a Surrogate Sensor Network. Once training is complete the blocks in dotted lines are eliminated.

value was 66 and when hitting the edge the mean MIDI value was 40. The difference between two means is 8 values to quantize 4 inches or one value being roughly equivalent to half an inch. The problem is that there is a standard deviation of approximately MIDI 3 values when using the Radiodrum. With only 8 MIDI values to quantize the examined distance and such a large variance the data acquired by the Radiodrum is not appropriate when a high degree of precision is desired.



(a) Hart mesh pad. (b) Roland Pad (c) PDP snare drum.

Figure 5.2: Radiodrum with various drums for experimentation.

Another test of the Radiodrum measured the jitter of a stationary stick. A stick was placed at the edge and the center and left for approximately 10 seconds. The data was recorded as a signal in Max/MSP to gain accurate resolution in time. Confirming the standard deviation observed with strikes, the data varied by 3 values when the stick was left in the same position. The jitter measurements illustrate the effect of ambient noise on the Radiodrum and help to explain the lack of precision.

	Minimum	Maximum	Mean	Standard Deviation	Variance
Center	66	78	73.366	2.1602	4.6666
Edge	59	67	63.043	1.9136	3.662

Table 5.2: Statistics on Radiodrum Accuracy with 100 hits at each location.

My original conclusion from this experiment was that the low correlation was not necessarily bad because the input of label data had so much variance. I decided to collect data and label the instances with discrete classes and continuous data to determine if the noise of the Radiodrum was a significant factor in the low correlation results. Further investigation into the use of discrete labelling is discussed in the following section.

5.3.2 Gesture Recognition

My previous research [168] has focused providing discrete symbolic labels for snare drum hits. The snare drum would be struck at different positions with different implements and the resulting sound would be processed with feature extractors and then classified using Weka[1].

Classification refers to the prediction of discrete categorical outputs from real-valued inputs. A variety of classifiers have been proposed in the machine learning literature [117] with different characteristics in respect to training speed, generalization, accuracy and complexity. The main goal of our experiments was to provide evidence to support the idea of using direct sensors to train surrogate sensors in the context of musical gesture detection. Therefore experimental results are provided using a few representative classification methods. Regression refers to the prediction of real-valued outputs from real-valued inputs. Multivariate regression refers

[1]http://www.cs.waikato.ac.nz/ml/weka/ - see Glossary for more information

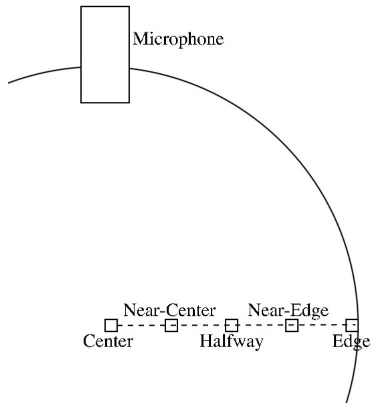


Figure 5.3: Diagram of radial position labels on the snare drum.

to predicting a single real-valued output from multiple real-valued inputs. A classic example is predicting the height of a person using their measured weight and age. There are a variety of methods proposed in the machine learning [117] literature for regression. Ordinal regression is a specialized form of regression where the predicted output consists of discrete labels that are ordered. For example when predicting the strike position in relation to the center of a drum it can be either a continuous value (regression) or an ordinal value with values such as center, middle, and edge (ordinal regression). Ordinal regression problems can be treated as classification problems that do not assume order among the labels but there are also specialized techniques that attempt to utilize the order information.

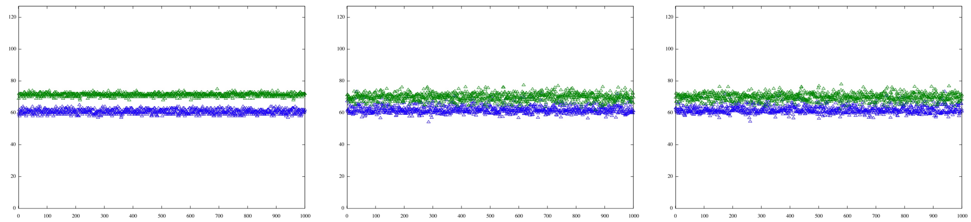
For some of the experiments described below, linear regression is used where the output is formed as a linear combination of the inputs with an additional constant factor. Linear regression is fast to compute and therefore useful for doing repetitive experiments for exploring diffent parameter settings. A more powerful back propagation neural network [117] is also utilized that can deal with non-linear combinations of the input data. The neural network is slower to train but provides better regression performance. Finally, the M5 prime decision tree based regression algorithm was also used [75]. The performance of regression is measured by a correlation coefficient that ranges from 0.0 to 1.0 where 1.0 indicates a perfect fit. In the case of gestural control, there is significant amount of noise and the direct sensor data doesn't necessarily reflect directly the gesture to be captured. Therefore, the correlation coefficient can mainly be used as a relative performance measure between different algorithms rather than an absolute indication of audio-based gestural capturing. The automatically annotated features and direct sensor labels are exported into the Weka [2] machine learning framework for training and evaluation [192]. For evaluation and to avoid over-fitting the surrogate sensors a 50% percentage split is employed where half the collected data is used for training and the remaining is used for testing. This ensure pairs of correlated feature vectors that are close together in time do not get split into training and testing.

In each experiment, unless explicitly mentioned the hits were regularly spaced in time. For each hit the radial position was measured and the hit was labeled as either "edge" or "center" using thresholding of the Radiodrum input. Audio features are also extracted in real-time using input from a microphone. The features and sensor measurements are then used for training classifiers. The setup can be viewed in Figure 5.2.

In the first experiment the electronic drum pad was hit in the center and at the edge. 1000 samples of each strike location were captured and used for classification. Figure 5.4a shows a graph of the MIDI data captured by the Radiodrum for each strike. Figure 5.4b shows a graph of the predicted output from a PACE regression classifier. The result was a correlation coefficient of 0.8369 with an absolute error of 2.3401 and a mean squared error of 2.9784. The graph clearly shows enough separation between the two classes. The data was then divided into two symbolic classes: Center and Edge. The data was run through the PACE regression classifier using the mean of the Radiodrum input for each class. The results were slightly improved - a correlation coefficient of 0.8628 with an absolute error of 2.0303 and a mean squared error of 2.6758.

The error achieved in the regression tests suggest that the algorithm has an ac-

[2] http://www.cs.waikato.ac.nz/ml/weka/



(a) Radiodrum Input (b) Surrogate Sensor (c) Surrogate Sensor with discrete classes

Figure 5.4: Regression results for predicting drum strike position using a surrogate sensor.

Table 5.3: Percentages of correctly classified drum pad hits (center, halfway, or edge).

	Classifier	**Ordinal**
ZeroR	36.5285	NA
SMO	80.3109	76.1658
Naive Bayes	76.6839	76.1658
J48	86.2694	89.1192
KNN	88.342	88.342
Neural Network	88.8601	90.4145

curacy of approximately 1 centimeter. Each MIDI value provided by the Radiodrum corresponds to approximately 0.5 centimeters and with an error of approximately 2, depending on the algorithm, this leads to a worst case error of 1 centimeter. Therefore even though the trained “surrogate” is not as accurate as the Radiodrum input it still provides enough resolution to discriminate between center and edge easily.

Table 5.3 shows classification results for predicting whether a mesh electronic drum pad was hit in the center, halfway, or the edge. As can be seen excellent classification results can be obtained using the surrogate sensor approach. A total of 348 drum hits were used for this experiment.

Table 5.4: Percentages of correctly classified snare drum hits

	ZeroR	**NB**	**MLP**	**MLR**	**SMO**
Snares	53	92	91	91	92
No Snares	57	93	94	95	95
Improvisation	59	79	77	78	78

Table 5.4 shows classification results for predicting whether an acoustic snare drum was hit in the center or the edge. The **Snares, No Snares** rows are calculated using approximately 1000 drum hits with the snares engaged/not engaged. All the results are based on 10-fold cross-validation. The trivial *ZeroR* classifier is used as a baseline. The following classifiers are used: *Naive Bayes* (NB), *Multi-Layer Perceptron* (MPL), *Multinomial Logistic Regression* (MLR), and *Support Vector Machine* trained using sequential minimal optimization (SMO). The results are consistent between different classifier types and show that indirect acquisition using audio-based features trained using direct sensors is feasible. The **Improvisation** row is calculated using 200 drum hits of an improvisation rather than the more controlled input used in the other cases where the percussionist was asked to alternate regularly between hitting the edge and the center of the drum. Even though the results are not as good as the cleaner previous rows they demonstrate that any performance can potentially be used as training data.

Ordinal regression [43] was computed for all tests to evaluate any difference. Tracking of strike position is a candidate for ordinal regression because the classes are ordered. Marginal improvements on some classifiers were obtained when ordinal regression was applied (see Figure 5.3).

An experiment was conducted to train a regression classifier using the Radiodrum as the direct sensor. Data was collected by playing on the drum moving gradually from edge to center and back to edge for a total of 1057 strikes. This experiment illustrates the "surrogate" sensor in the intended application of rapid data collection and training of a classifier.

Table 5.5: Radiodrum regression from with 1057 instances moving from center to edge.

Regression Type	Correlation Coefficient
Linear Regression	0.4594
SMO	0.6522
Pace	0.6579
Neural Network	0.6737
M5' Regression	0.7394

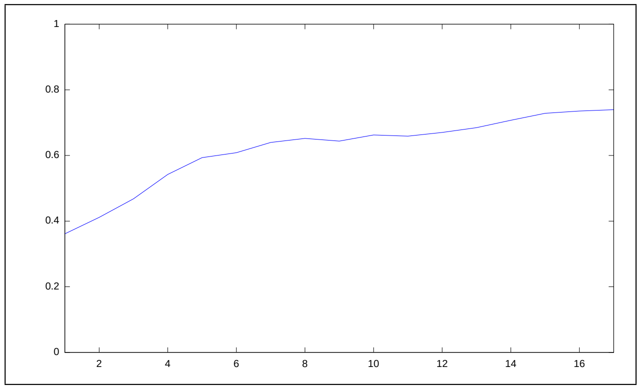


Figure 5.5: The effect of more features to the correlation coefficient in drum regression.

To verify the effectiveness of the features used for classification an experiment was conducted to progressively add features. The feature vector was reduced to 1 element and then increased until all 17 features were included (see Figure 5.5). The plot shows an increasing line as features are added back into the vector and the correlation coefficient increases.

5.3.3 Classifying Hands

A test was conducted to determine if a classifier could be trained that would be able to differentiate between the left and right stick of a player. Two hundred and eighty one samples were collected and with a ZeroR classifier 50.8897% instances were classified correctly. A subset of the samples was taken and the ZeroR classifier returned 57%. (see Table 5.3.3)

Regression Type	Classification (Full Set)	Classification (100)
ZeroR	50.8897	57
Bayesian	68.3274	85
SVM	66.548	80
Neural Network	78.2918	89

Table 5.6: Classification results to differentiate right and left sticks.

In both cases the classifiers were able to classify at least ten percent better than the ZeroR baseline. The Neural Network, though often subject to overfitting, was able to achieve 27.4021% and 22% excess of the baseline, respectively. Though these results are not high enough for performance, they do provide an interesting result. These results lead one to believe that it is possible that with further research it will be possible to detect a multitude of semantic data from the audio stream.

5.4 Summary

The E-Drumset software provides two methods of recognizing gestures: implicit and explicit position tracking. Explicit position tracking is achieved using machine learning to determine the timbre produced by a drum controller and then infer position based on labeled training data. The explicit position tracking can be tailored to a specific application by using different classification algorithms outlined in this chapter. Implicit position tracking is achieved by using the acoustic signal of a drum controller to allow the listener to hear to timbral variations that naturally occur when the controller is struck at different locations. The accuracy of these systems is sufficient for performance, as will be demonstrated in the following chapter.

The concept of a surrogate sensor is used to "train" machine learning model based on audio feature extraction for indirect acquisition of music gestures. Once the model is trained and its performance is satisfactory the direct sensors can be discarded. Large amounts of training data for machine learning may be collected with minimum effort just by playing the instrument. In addition, the learned indirect acquisition method allows capturing of non-trivial gestures without modifications to the instrument. The idea of using direct sensors to train indirect acquisition methods can be applied to other area of interactive media and data fusion.

CHAPTER

SIX

Supporting Activities

The most beautiful thing we can experience is the mysterious. It is the source of all true art and science.

– Albert Einstein

6.1 Introduction

One of the main goals of this thesis is to transition the E-Drumset from a proof-of-concept state into a performing instrument. My belief is that this transition requires a champion, or virtuoso, to perform regularly and push the boundaries of the interface technically and aesthetically. During the course of study the E-Drumset was presented in demos and performances. Additional software was created to explore additional contexts for electronic percussion. This chapter outlines these supporting activities.

6.2 Demos

On October 2, 2006 Minister of Advanced Education, Murray Coell, visited the University of Victoria's Computer Science department. During this visit he encountered the E-Drum demo and began playing (see figure 6.1). The Minister is a former pa-

Figure 6.1: Minister of Advanced Education Playing the E-Drum in a Demo Session with George Tzanetakis.

rade drummer, as can be seen by his use of traditional grip. The immediacy of the Minister's connection to the drum demonstrate clearly its success as a device that is intuitive to a drummer and allows them to transfer their skill set effortlessly.

In 2005 an early working prototype of the E-Drumset was presented at the West Coast Composers seminar at the University of British Columbia [164]. As part of the presentation an improvisation of the University of Victoria constituents at the conference was performed to demonstrate the flexibility of the interfaces presented during the demo period. The University of Victoria team included: Ajay Kapur, Randy Jones, and Jeff Morton.

6.3 Performance Ensembles

6.3.1 i send data live

i send data live is an experimental ambient and electroacoustic duo consisting of

Adam Tindale and Robin Davies. The name of the ensemble is a anagram of our last names. In the ensemble I play various forms of electronic percussion and Robin plays laptop exclusively. This project served as the main venue of experimentation for the various developmental stages of the E-Drumset.

As stated by many others [152, 195], laptop performance suffers from the problem of sound source identification: who is generating which sound. In performance with *i send data live*, the relationship between the gestures and the sounds of the electronic percussion setup used are very direct: a sound is generated only when the drum is hit. Through the course of a performance the listener is able to identify the relationship between my gestures and the sound then correctly infer that the remaining sounds must be generated by the laptop. This technique has proven to be very successful in illustrating sound sources to audiences and giving them an appreciation for the laptop's contribution to the sound. We measure success by noting that most of the audience members who inquire about the technical production after the show approach Robin to see how he is achieving his results. This indicates to me that they understand how I am making sound and need no further explanation.

Interactive Futures 2006

The first major performance of *i send data live* was at Interactive Futures 2006 [166]. During this performance the initial design of the E-Pedal was used in conjunction with an Alternate Mode Drumkat. The E-Pedal utilized a Sharp GPD120 infrared distance sensor. During this performance it was determined that the latency of the sensor was inappropriate for live performance.

Alberta College of Art and Design Invited Performance 2007

i send data live was invited to give a performance and lecture at the Alberta College of Art and Design. This performance utilized a combination of E-Drumset and the Alternate Mode Drumkat. This performance was designed to be a transition away from the Drumkat to a full E-Drumset setup for the group. The performance began with older material that relied upon the Drumkat and then moved to newer material developed around the E-Drumset.

Text, Media and Improvisation 2008

This performance utilized a travelling version of the E-Drumset. The setup included a Hart Multipad connected to an Edirol soundcard and an Asus EEEPC running ChucK on Debian Linux. An E-Pedal connected to an Arduino microcontroller was used to capture the hall effect sensor and infrared sensor data.

The Hart Multipad is an array of 6 pads arranged in two rows of three. Each pad contains a piezo microphone but unlike other devices the audio outputs are directly available. Each pad may be connected discretely to a computer or other sound modification device. The Hart Multipad is therefore ideal as a portable, smaller scale, E-Drumset device and will be a core component of future *i send data live* repertoire.

6.3.2 i send data live with Visual Jockeys

For nearly half the performances of *i send data live* visuals were performed by VJs, or visual jockeys. *Love and Olson* provided the visuals for many of the shows, integrating the concepts of narrative and ambience into the visual material. The visual aspect of the shows could emphasize the experimental nature of the sound design, further informing the audience as to the intention of the performance.

6.3.3 LUVLEE

LUVLEE is an acronym for Live University of Victoria Laptop and Electroacoustic Ensemble. The group is comprised of three members, originally David Berger, Kirk McNally and myself. After his graduation, David Berger was replaced by Jeff Morton. The group performed at many functions in the University. The members of the ensemble were equal members in creation and direction of the ensemble.

The group began as a fully improvisational ensemble. There were many successful performances with this format but the audience seemed alienated during longer performances. Silent films were utilized as visual material but there were a number of other unforeseen advantages. The audience was able to enjoy longer performances because the visual material was narrative, not just associative. The ensemble became more focused because we had to develop material that would emphasize certain moments. The group moved towards precision in composition and execution rather than improvisation. The visual material also gave the audience a focus, the gestures of laptop performance can be interesting but have many established problems of resembling office work.

6.3.4 Andrea Revel

Andrea Revel is a folk/pop singer who experiments with technology integrated into traditional form. I have used the E-Drumset in many different forms in performance with Andrea Revel (see figure 6.2). The two main configurations are my personal E-Drumset as outlined in Chapter 4 and a simplified setup where just the E-Pedals are used with an electronic bass drum pad, a cymbal and snare drum with a variety of drumsticks.

The small setup is designed to provide enough functionality to perform in the group appropriately. Simple music is often best served with simple drums and drumming. The E-Pedal allows me to tune the bass drum appropriately for each song and for each playing environment. The small setup has the added bonus that I am able to carry the whole setup on my back in one load.

6.3.5 The Electron Orchestra Withdrawal

The Electron Orchestra Withdrawal is an exploration in expanding the prototypical rock ensemble: guitar, bass, drums, and vocals. The group consists of Clinker (aka Gary James Joynes) on bass and electronics, Les Robot on Guitar, Jackson 2Bears on turntables and electronics, and me on E-Drumset. The group is able to explore electronic music, rock music, ambient and metered structures through improvisation.

Figure 6.2: Performing with Andrea Revel using a E-Drum pedal.

Sonic Boom was an event hosted by EMMEDIA artist run centre in Calgary, Alberta. The show featured Clinker and Jackson 2Bears and was the premiere event for the ensemble. During the performance large scale visuals were projected inside of a hemispherical dome, which brought a level of grandeur to the performance (see figure 6.3).

6.4 Software Applications

Three main software applications were developed during the course of the study. The applications will be presented in order of their development: Thief, Strike-A-Tune, and the UVATT Radiodrum music therapy software. Two of these applications, Thief and the UVATT Radiodrum software, utilize the Radiodrum hardware. The focus on the Radiodrum was due to local needs of advancing the hardware, as well as making myself more familiar with the advantages and disadvantages of this interface.

Figure 6.3: Performing with E-Drumset and the Electron Orchestra Withdrawal.

6.4.1 Thief

Thief is a software framework for improvisation using the Radiodrum. Many musical controllers are not capable of producing sound on their own, rather they translate gestures captured into information to be mapped into musical instructions. Thief is built to provide the Radiodrum with an interactive method of capturing musical information for playback and manipulation. Thief was written for Andrew Schloss to be part of his improvisational toolkit.

The software is intended for use with a Radiodrum performer and another performer that provides MIDI output. Thief "listens" to a MIDI stream and keeps a record of all previous events. These events may be used and manipulated in many different ways by the Radiodrum performer. This approach for generating meta-level data for the Radiodrum performer allows a more flexible and cohesive environment for improvisation. No longer is the Radiodrum performer restricted to preprogrammed pitch material; As the performer generating the MIDI stream creates

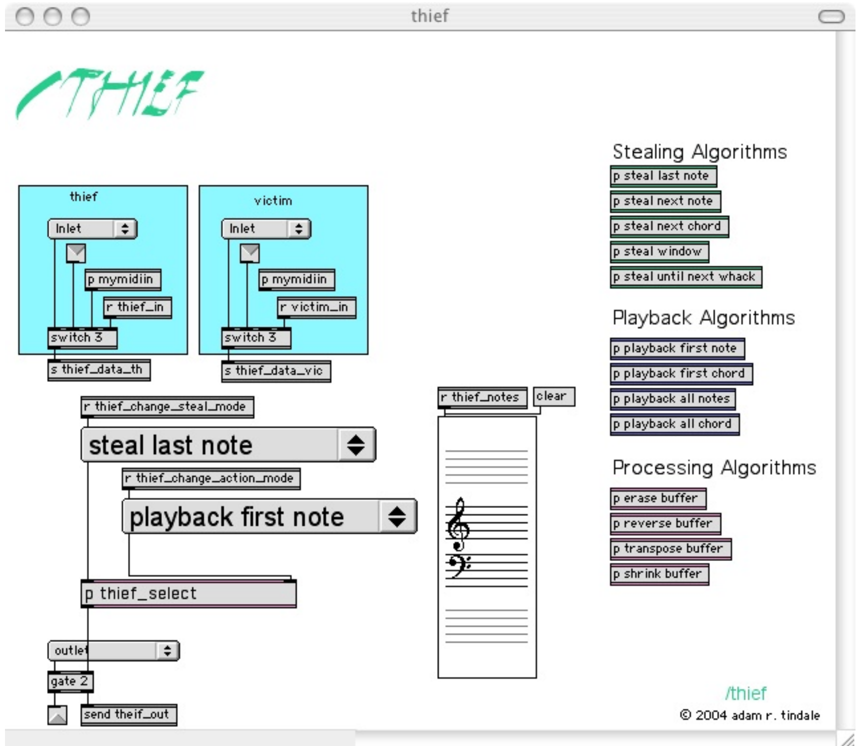


Figure 6.4: Screenshot of the Thief software.

musical ideas in an improvisational context, the Radiodrum performer is provided with the stream which they are free to navigate and manipulate so that the musical material will evolve as ideas are traded between the performers.

Features

Thief is implemented with the Max/MSP programming environment from Cycling '74 [1]. This allows for compatibility with the vast amount of other software developed for performance by Andrew Schloss. Thief does not provide any methods for synthesis, rather it provides macro-level control and processing. Thief is meant to be integrated with other software, either Max/MSP patches or other software that communicates with Max/MSP. It is able to function as a standalone patch that messages can be passed to or it may be run as an abstraction inside of a larger patch.

Thief provides the user with many different algorithms for processing the input

[1] http://www.cycling74.com

MIDI stream. The MIDI stream is captured into a collection of events. The user may:

- Capture the previous event into the collection
- Capture the next event into the collection
- Capture the previous window into the collection
- Capture the next window into the collection
- Capture until the "whack" from the Radiodrum
- Playback only the first event in the collection
- Playback all of the collection
- Erase the collection
- Reverse the collection
- Transpose the collection
- Shrink the collection

Visual feedback in the form of a musical notation is provided to the performer (see Figure 6.4.1). Max/MSP provides a visual object called "noteslider" that displays a musical staff. The next note or notes that are scheduled for playback are presented in this object. This allows the Radiodrum performer the ability to evaluate the choice of the next note.

The framework for thief is designed in such a way that it is relatively simple to expand it with new algorithms. As the software gets tested in real world situations surely new ideas will emerge. These ideas will be integrated into the framework in order to make the software comprehensive for an improvisational environment. Other traditional transformations on music data such as inversions, retrograde, etc. are planned to make a more intuitive interface.

Currently, the framework is designed with the assumption that the “thief” is a Radiodrum player. In the future a more general solution is planned so that these techniques may be used by other computer musicians. A future implementation of the note stealing algorithm will be implemented in a text driven language, such as ChucK.

6.4.2 Strike-A-Tune

Motivation

There are many pieces of software available for organizing music collections and playing it on a computer. With rare exception, these programs utilize the keyboard and mouse for navigating the music collection: devices that are very precise.

For large collections of music, problems arise. Most users tend to navigate their collection in a somewhat similar fashion. Users will more often pick certain artists over others. This can lead to the problem of songs being rarely or never selected, resulting in wasted storage or worse, money wasted on purchased music. The other typical method of navigation is random mode or the iTunes “Party Shuffle.” Random playing does not address the user’s current mood or preferences, and can often lead to frustration and constant track advancement.

Thus a new method of navigation is proposed: *fuzzy navigation* [170]. People are often in the mood for a style, mood, or tempo of music they like, without being concerned about the particular song playing at any moment. Two websites, Pandora.com[2] and last.fm[3], have become very popular by playing music based on user-guided feedback. Users are asked for some general preferences and then asked during each song if they like the current selection. The algorithms used by these websites adapt to play more and more desirable music. Although this method works, it works only with the database of music held by the website, which is subject to copyright. What about a similar fuzzy navigation for the collection already stored on a computer? Furthermore, what if the fuzziness of the navigation was in the

[2]http://www.pandora.com
[3]http://www.last.fm

hands of the user?

Using a musical instrument for an input device provides no foreseeable advantages compared to conventional mouse and keyboard navigation. The imprecision of most acoustic instruments can actually be a disadvantage. This imprecision, however, can be leveraged to promote fuzzy navigation. Another distinct advantage of using a drum interface for music navigation is the tactile feedback provided. The user is naturally aware of any instrument in his or her hands [5]. They are also acutely aware of how hard they hit the drum because haptic or tactile feedback is inherently provided.

This work combines a drum-based physical interface with visualization software to allow users to navigate their collections in a different, yet intuitive way. Utilizing previous work, the drum can provide data about how it was struck enabling it to be used in a similar fashion to a keyboard and mouse. The benefit of the drum is that it provides an accurate device for repeatability, but is not very precise. This lack of precision makes the same output on two separate strikes unlikely. These minor variations mean that songs similar to the desired song are more likely to be played. Songs that the user has not listened to in a while may be rediscovered and the desired style of song is still played. The Strike-A-Tune system, therefore, uses no randomization or induced chaos but rather relies on imprecise motor actions of the user to broaden music searching. Thus, it is possible to increase system precision by participants using more controlled movements.

Previous Work

The MusicTable system using ludic or indirect navigation allows informal groups to collaboratively choose music during a party [160]. This is based on the work of Gaver et al. [48] who wanted to look at navigation based on curiosity, exploration and reflection (ludic navigation) rather than explicit tasks. Although both Strike-A-Tune and MusicTable rely on a lack of precision to promote less task oriented music selection, MusicTable relies on ludic navigation to support collaboration and reduce social tension, while Strike-A-Tune is designed for personal use. The PersonalSoundtrack system plays music selected from a user's personal music library, based on a user's walking pace and gait [34]. Similarly, Oliver et al.[125] used biometric feedback to adjust the music played on a personal MP3 player to adjust an exercise routine through music and to adjust the playlist based on the user's enjoyment [126]. All of these interfaces provide fuzzy music navigation, but none are designed for music re-discovery. Other music navigation systems such as Pandora.com, last.fm, and the Music Rainbow [129] are also neither goal oriented navigation nor a random music selection. The Music Rainbow system, despite also using a circular interface based on song similarity, is designed for music discovery, and permits precision music selection.

Murdoch and Tzanetakis' [119] Radiodrum research looked at using a drum to navigate large music collections. They propose two navigation methods: rhythm based navigation and SVM dimensionality reduction. For dimensionality reduction, three dimensional song positions would be accessed by the three dimenensional position of the drum sticks. They hoped that this would lead to a tangible and intuitive interface for content-aware music browsing.

Several ring or circle based visualization techniques have been proposed in the past. The CircleView system by Keim et al. [88] uses a visually similar approach to ours with segments of a circle representing an attribute, distance from the exterior representing time, and colour representing the average data value for that time. CircleView is used for visualizing multidimensional time-referenced data sets. The visualization is interactive but relies on mouse and keyboard interactions. Appan et al.[7] similarly use the analogy of water ripples to display communication patterns over time. Neither of these approaches are identical to the disk visualization (described in the Visualization section) and the use of a custom designed visualization technique for an unconventional interface is a research contribution.

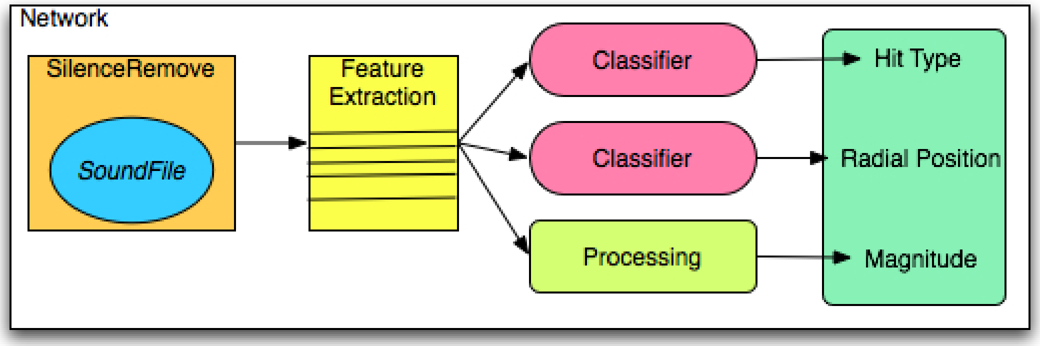


Figure 6.5: Marsyas network used for the drum interface.

Drum Interface

A drum interface is used to control the browsing of music collections. In this section a description of the specific system used for position and gesture tracking is discussed.

The categories of gesture used in the system are Stroke Type (Normal, Flam, Rim), Radial Position(0-1), and Magnitude (0-1). The flam stroke is produced when the user uses two sticks that hit nearly simultaneously. The radial position refers to the point on the drumhead that the drum was struck. Since there is only a single sensor in the middle it is impossible to correlate with another source to find the position in two dimensions.

The drum software interface takes an audio stream and segments it into discrete hits. Once a hit is detected, a fixed window of 4096 samples are captured to be processed. This window is sent to the feature extraction section of the software where the resulting feature vector is delivered to multiple classifiers. The output of the classifiers is collected and sent via Open Sound Control[193].

The features used include: Root Mean Square (RMS), Temporal Centroid, Spectral Centroid, Spectral Kurtosis, Spectral Skewness, Spectral Rolloff, Spectral Flux, Subband analysis, Wavelet Coefficient Averaging and Linear Predictive Coding [74].

The features are collected into a feature vector to be classified by a Guassian classifier provided in Marsyas [177]. Gaussian classifiers are very easy to train and are fast in classification but are not particularly accurate when compared to the performance of other classifiers [31].

In our use of the system, the magnitude metric seemed to be effective for navigating the circle, although it was discovered that it was difficult to reach the first song in the outer wheel. This begs the question of how much precision is actually needed and how much is available. In order to navigate the current interface the radial position must have four regions of accuracy and the stick type and stroke type must be accurate.

When classifying the difference between Normal and Rim the system achieved 89.0% accuracy. The use of the stick gestures to navigate the music collection will be discussed in the next section.

	2	**3**	**4**
Radial Position	**89.0%**	81.2%	73.2%

Table 6.1: Accuracy of Radial Position classification.

The Disk Visualization

The song visualization system for drum based fuzzy music navigation was designed to be intuitive, scalable, appropriate for the drum interface, and require no keyboard or mouse input. Song data needs to be scalable so that an overview of the music library is available while information about individual songs is easily discernible. A disk visualization was chosen for these reasons. The shape of the disk matches the drum interface's shape making radius judgments intuitive. Strike magnitude, although perhaps not intuitive on the first strike, could be quickly acclimatized to.

6.4.3 Visualization Overview

The disk visualization presents rings of song data with the currently playing /chosen song in the center of the disk(see Figure 6.6). Rings closer to the center have fewer songs positioned on the ring, but more song details. Shneiderman's information visualization mantra of "Overview, zoom and filter, details-on-demand" is thus supported [154]. An overview of the music is provided by outside rings, song data

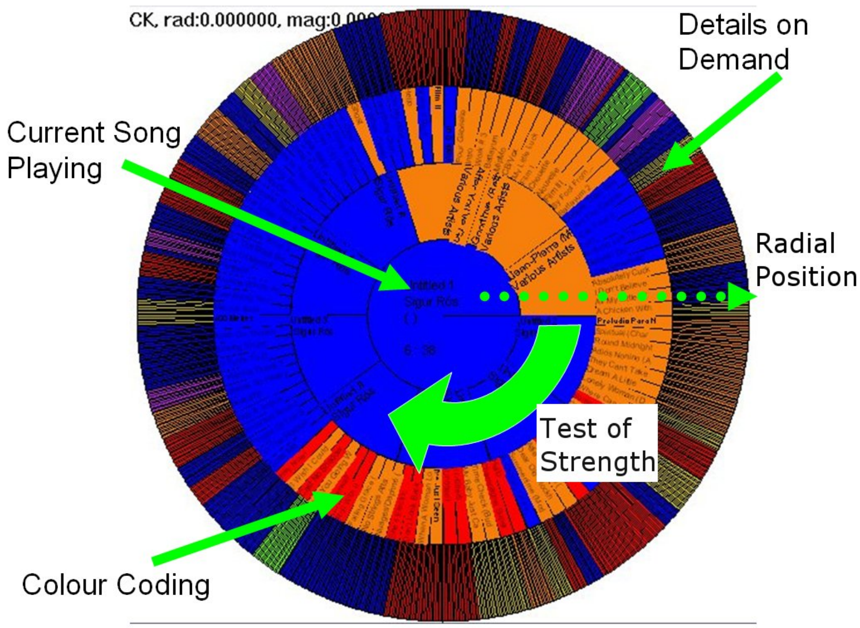


Figure 6.6: Overview of the disk visualization. When songs are highlighted details become available.

is filtered based on the currently playing song, and song details are provided on demand by highlighting the song [88]. Songs on a given ring are currently positioned evenly around the ring in song segments. The current implementation presents 1 song in the center, 10 songs on the first ring, 100 songs on the second ring, and 1000 songs on the outside ring. Ring number thus conveys 1 dimension of information while position on the ring can convey a second dimension. Ring position and segment order can also convey a single variable, like song similarity (used in the current implementation). Currently, songs on a ring are positioned according to similarity to the center song. Songs on rings closer to the center are also more similar to the center song than songs on outer rings. Each song segment is coloured according to a third nominal dimension (such as genre). Strike-A-Tune supports up to three dimensions of song data for each library entry, in addition to traditional text based information.

The disk visualization intuitively maps drum strike information to visualization interactions. The radial position from the drum center of a simple “selection” strike

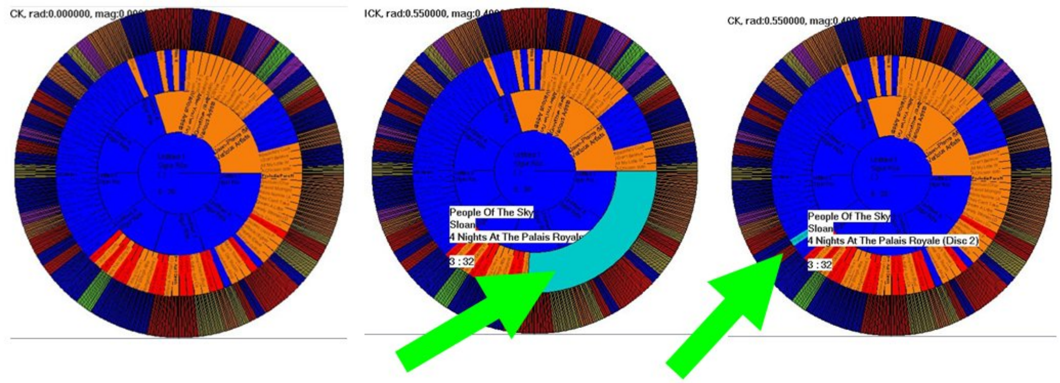


Figure 6.7: Three steps in the song highlighting processing. When a selection strike occurs, the selected song's detailed information is displayed and a blue bar animates around the selected ring until it arrives at the chosen song (ii). The chosen song is then highlighted (iii).

maps to the radius position on the disk. Radius positions (between 0 and 1) are then discretized to a ring number. Strike magnitude (0-1) represents a rotation around the chosen ring between 0 and 360 degrees. A "test of strength" analogy is used to convey how magnitude impacts a selection. A zero degree rotation selects ring segments aligned with the x axis of the center song. Stronger hits will cause an animation to rotate clockwise around the ring until the chosen song is selected . This animation both attracts attention to the selected ring segment (movement is perceptually pre-attentive) and reinforces the "test of strength" analogy (similar to a carnival high striker game). See Figure 6.7 for details.

Stroke type is not currently represented visually but are used for system state control. Drum flams choose/play the currently selected song, rim hits stop the currently playing song and remove all highlighting, and simple strikes (selection strikes) use magnitude and radial position to highlight a song.

Song Similarity

The 3207 song personal music library of one of the authors was chosen because it was relatively large, varied, and convenient to work with. The song library was organized by iTunes and song metadata was output to a tab delimited text file. Each song had 25 different pieces of metadata including song name, artist name, album name, comments, and track length.

Fuzzy music navigation based was computed based upon song similarity. Each metadata category was weighted based on importance and string similarity was either determined by exact string matching or the percentage of characters in common. The weighted similarity values for each of the categories were summed to a value less than or equal to 1.0 (with 1.0 being a perfect match). One minus the similarity metric gave us a euclidean distance between each pair of songs in the library. These $10,284,849$ (3207^2) distance values were then stored in a space separated text file.

Upon program start up, a Song object is created for each line in the metadata text file (Library2.txt) and each song is numbered according to the parsing order. Distance values for each song are then extracted from the distance file (SongDistances3.txt) and stored in an array. This permits real time interactions for large libraries once the files are loaded. Unfortunately, this approach is memory intensive and loading times can take several minutes.

Discussion

The Strike-A-Tune system offers many advantages compared to conventional shuffle or precise mouse and keyboard based music navigation systems (ie. precise navigation). The system is scalable and designed specifically for large libraries. Unlike both precise and random music navigation, the benefits of fuzzy navigation improve as the library size increases. Many songs the user has forgotten about have a chance at being played (unlike using precision navigation), but only songs similar the user's current preference will be played (unlike a randomized playback), reducing user frustration. The disk visualization provides an overview of music data and details for songs in focus, further supporting large libraries. Finally, the intuitive nature of the interface should make interacting with the system as simple as using a traditional mouse and keyboard, but allows us to expand our interface in novel new directions of music exploration.

It is unclear that fuzzy navigation will be appreciated by users. Case studies and experimental testing is required. The lack of precision may frustrate task-oriented users and some users may feel the interface is too error prone. Using metadata for similarity metrics is a naive but fast approach to gathering song distance information. Signal processing of the actual audio signal is likely far more accurate and closer to how humans perceive similarity. Future work will include the investigation of these techniques.

6.4.4 Music Therapy

Music Therapy is a growing field both locally and internationally. For the past year, the UVATT (University of Victoria Assitive Technology Team) and MISTIC (Music Intelligence and Sound Technology Interdisciplinary Centre) groups have been utilizing a 3-Dimensional controller - the Radiodrum - for therapy and musical applications [165]. This joint project aims to bring the needs of the two groups together into a single hardware interface with a common software access point.

The Radiodrum uses capacitance to detect the position of two sticks above a surface. The sticks can be substituted for custom sensors for users with limited facilities, as demonstrated by Prinz [138] and Falkenberg [66]. The data is output as MIDI (Musical Instrument Digital Interface) and transferred to a computer running the ChucK programming language [187].

ChucK is a flexible high-level realtime object-oriented music programming language. ChucK allows the construction of classes that encapsulate the hardware interface of the Radiodrum MIDI functions. This makes it very easy for a programmer to devise a mapping or framework in order to produce musical stimulus for therapy sessions.

To date, there have been two patients to use the system on a weekly basis. The therapist devises strategies appropriate for the patient and their individual needs. Patient 1 had difficulty moving one arm. Patient 2 had little physical difficulty but although mentally handicapped was very musically proficient. Visual aids were provided to help the patient for each mapping.

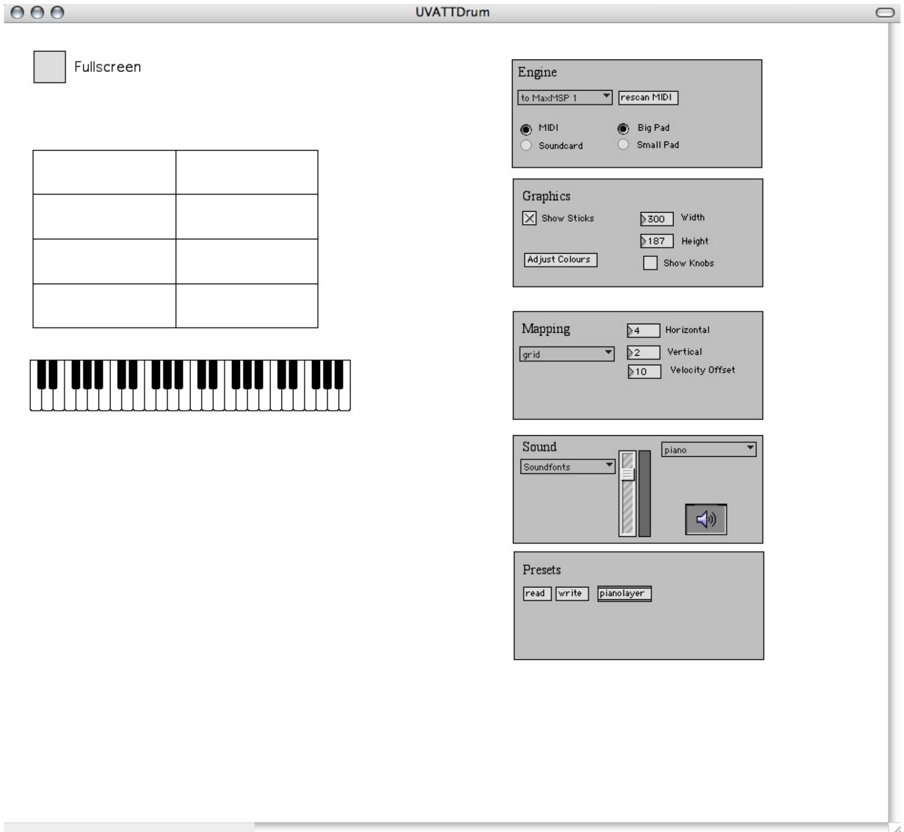


Figure 6.8: A screenshot of the Max/MSP interface of the software.

Figure 6.9: A Radiodrum with a layover on the Radiodrum surface.

The current patients use a simple triggering mapping where the surface is arranged in a grid. Each cell on the grid may be populated with a sound that is triggered when an impact is detected on the surface. Currently, the impact detection is done in hardware and the events are then parsed in software. Future versions will provide an adaptable algorithm for individual patients.

The grid algorithm is also abstracted into a class to allow for different sizes and configurations. (The Radiodrum data is less than perfect and creating a grid that satisfies a users' involved a great deal of trial and error.) The linearization was able to instantly affect all of the mappings that were dependent on the grid class without interrupting regular therapy sessions.

Each patient required a slight variation on the grid analogy. Patient 1 moved the hand with limited mobility in the Z axis to change which sound was triggered when the other hand struck the surface. Patient 2 divided up the grid into 8 zones that contained the 8 notes of a musical scale.

There are two types of events that the patients and their therapists have requested: instantaneous events (eg. Playing a note) and meta-level events (eg. Reconfiguring the layout of the grid). Meta-level events are generated by a footswitch and this hardware is also abstracted for the programmer. The footswitch is used by the therapist to change pitch sets for Patient 2.

The main challenge has been to design mapping strategies that are intuitive enough, so as not to frustrate the patient, yet interesting enough to keep the patient from losing interest. Another significant challenge has been to find a way to communicate the functionality of the device to the patient. The therapists have been in charge of the latter problem while providing ideas for the former.

We have been able to provide an interface for children to play a musical instrument they would otherwise be incapable of playing. The vast flexibility of the Radiodrum has allowed for many different gestures to be captured and translated into musical events. Creating an abstract layer for mapping the Radiodrum has enabled rapid prototyping for patient's needs. Two patients have received extensive, weekly therapy using this system. As the number of users grows, the system will be forced to grow into a more flexible and generalized system centered on patient's needs.

At the inception of the project, therapists would ask for a feature list for a patient and the programmer would deliver the software, usually after one or two weeks. Patients would request an instrument that they would like to play and the therapist would design a control paradigm for that patient. Since the control mechanisms and sounds were hard coded, the therapist was unable to vary their sessions to explore other musical territories. Over the past year, a dynamic software framework was designed enabling therapists to easily customize the mapping of Radiodrum gesture signals to the virtual music space in real-time. Our system provides a therapist with the ability to quickly implement and test musical mapping schemes on-the-fly without low-level knowledge of the system.

While many of the users are able to achieve proficiency with the Radiodrum the cost of the entire setup is prohibitive. Most individuals cannot afford to own one and practice independently of therapy sessions. Currently, the cost of the entire system is above 3000USD per unit. The use of microprocessors and dedicated DSP chips is being investigated to create a complete system that patients may purchase for under 300USD. Although originally created as a 3-Dimensional mouse, the Radiodrum has become widely used in the computer music community. Bresin et al. [66] have adapted the Radiodrum into a DJ interface, virtual bodhran, and a virtual physics controller. Similar projects include using the Radiodrum to browse large music collections in the context of music information retrieval [119]. This novel work is pertinent to the assistive technology community in that it provides more intuitive human computer interaction when searching media. Prinz [138] adapted the Radiodrum to be mounted on a headband for quadriplegic users to interact with a computer.

The software is implemented in Max/MSP/Jitter and combines programming capabilities with visual feedback for both the therapist and the patient. When the software is connected to the Radiodrum hardware two coloured circles are placed within the surface representation area with the size of the circles representing the distance from the tracking board. The main interface (see Figure 6.4.4) provides buttons that allow for quick access to low-level functionality. The piano keyboard can be used as both a display for the notes being output and as a quick input to designate the pitches that the software will output when triggered.

The software package allows events from the Radiodrum to be translated into sounds that can be played or to MIDI messages to be played by other software or hardware devices. The most common event is a hit. Since the Radiodrum is capable of detecting the position of the hit, the surface is often divided into a grid with separate actions for each cell. The cells then emit a separate event that can be further mapped to pitches or actions (eg. single notes, chords, volume control, effects control). More than a dozen application presets have been developed used in therapy sessions. The presets range from playing flute sounds to drums to a theremin. Not only is the software easily programmable by a non-programmer but it allows instantaneous dynamic change. During therapy sessions this has proved to a vital feature because it allows therapists to change the key of Radiodrum output on the spot. In practice, each patient would have a different sized grid for the surface. A laminated layover with a visual aid was created for each patient (see Figure 6.4.4). Pictorial visual aids were added for patients with static presets. Other patients had static layout with varying musical events. One patient was able to achieve such mastery that he was able to select and play a pitch as it was sung by the therapist.

6.5 Summary

The E-Drumset has been used in a number of performances. In these environments there is no tolerance for noise or failure. During years of performance experience the E-Drumset has proven to be a reliable concert instrument. The flexibility of the instrument is illustrated by the ability to improvise in a number of musical situations.

Aspects of the E-Drumset software have been used alternate settings, such as control structures, music therapy, or music browsing. The expansion of the thesis into other areas added validity to the concept. Many of the expansion projects were completed as collaborations. These projects owe greatly to these individuals and their willingness to collaborate and expand this work.

The E-Drumset has been used in a variety of musical contexts, demonstrating the versatility of the instrument. The E-Drumset is not bound by a single genre of music, just like the acoustic drumset. The instrument was used in high and low volume environments and was successful in both, no problems with balancing, induced noise or feedback were encountered.

In general, the E-Drumset is very successful in performance and is well received by the audience. I often received questions from audience members about the system and how they may obtain a copy of it. With some ensembles I rarely received questions, which generally indicated to me that the audience understood the instrument. With i send data live the quizzical audience member would approach Robin Davies, confirming the clarity to the audience.

Throughout the time with the E-Drumset many things have changed. Initially 6 E-Pedals and multiple expression pedals and sustain pedals were utilized. Currently, only two E-Pedals are utilized because the extras were found to be superfluous. During performance patch changes are not used, therefore eliminating the main purpose of the sustain pedals. Over time, the need for expression pedals was eliminated because the E-Drum software, both synthesis and recognition, improved greatly and offered sound possibilities that required no alternate controls.

Multiple E-Pedals was ultimately abandoned. The idea was to have a variety of bass drum sounds available but was rarely needed, especially with the flexibility offered by one E-Pedal. Although shifting one's foot from pedal to pedal is athletic and satisfying, it became more of a gimmick than a tool and the extra E-Pedals were removed.

As my proficiency with the E-Drumset increased many more opportunities to play with other musicians arose. The initial ensembles were groups that I formed in order to explore the E-Drumset. In more recent times artists would contact me in order to to collaborate. This shift is both flattering and indicative of the success of the E-Drumset as an interface that is most suitable to my abilities and aesthetic choices.

Many instruments are successful in demos and in laboratories. The E-Drumset ha proved to be reliable in a number of real life musical situations, moving it beyond a proof of concept design. The software and configuration of the E-Drumset has proved to be flexible and adaptable, so as to evolve throughout the course of development to be come a true instrument.

CHAPTER

SEVEN

Towards Mastery

We have to get across the idea that the drumhead is simply like a piece of paper to the writer. It's really fairly unimportant. Kids mustn't get the feeling that if you duplicate the drumkit you duplicate the drummer! It's not going to help them to get the same drumheads. What's going to help them is to get past the idea of the machinery and the technique, and get into the c o n c e p t s of the thing, which are much more fun! That's where life starts. I'm not so good about describing drumkits and drumheads endlessly, and talking about paradiddles. I love those subjects, but beyond that is much more interesting.
– Bill Bruford, Modern Drummer, July 1983

7.1 Introduction

This section describes the discoveries of new gestural possibilities on the E-Drumset. These possibilities come in the form of both physical and mental techniques and are outlined throughout this chapter. The utilization of these techniques in a practice regiment is discussed to offer the reader insight into the development of proficiency on the E-Drumset.

Gaining proficiency at a new instrument is a combination of multiple factors.

The initial motivation of a new instrument to deliver a new sound with a given gesture can be realized into practice relatively quickly. Once the initial stage has been completed it is up to the player to discover the other limits, tricks, difficulties and potentials offered that were not originally foreseen. An interface that doesn't reach this point is simply a toy.

How does one create a new language of gesture and pedagogy? I believe that part of the answer is through hard work and experimentation. In a way, this is a simple answer to a complex question. The process followed through the course of developing the E-Drumset was to a combination of experimentation in a solo context and experimentation in a musical context, followed by reflection on new sounds or combinations inspired by varying contexts or ideas. The new ideas would be integrated into exercises.

During my years of teaching drum lessons I found that there were three methods to encourage students to develop new ideas: improvising at the drumset, improvise writing patterns, improvise singing patterns. Improvising at the drumset often leads to situations that are comfortable for the hands to play and I have found that it is good to keep track these tendencies and see how they change over time. Improvising writing patters helps develop one's ability to notate and transcribe as well as leading to patterns that may be easy to write but can be very uncomfortable to play, revealing limits in one's playing and demonstrating difficulties that the instrument possesses. Improvising singing patterns and then realizing those sounds helps to develop a sense of what sounds are most interesting to the player. Singing, debatedly, comes from a more "musical" starting point and can help develop musical sensibilities in the player. Like writing, singing helps to develop patterns that may be uncomfortable to play.

In classical training the method is to overdevelop proficiency so that the music encountered by the player will be relatively simple compared to one's skillset and the player is able to focus on musical expression. In classical music this is an excellent, if time consuming, method of development. For other forms of music where the player has creative input on the resulting musical product this method can run into problems. Some people believe that you are what you practice. Since most drumset pedagogies fall close to the idea of classical training, drummers can fall into the trap of playing overly technical patterns that they have been practicing in a musical context that is not appropriate.

7.2 Exercises

The basic rock beat is the core of the rock drumset repertoire. Exercises for the E-Drumset were devised with the basic rock beat in mind. The basic rock beat consists of a riding pattern on a hihat or ride cymbal, the snare drum playing on beats two and four, and a bass drum pattern usually accenting beats one and three (See Figure 7.2).

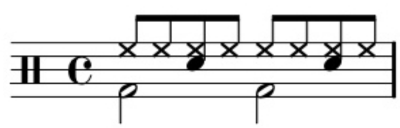


Figure 7.1: Basic Rock Beat.

The first set of exercises explored various methods of fitting four notes within a quarter note (See Figure 7.2). Patterns were generated and the progressively moved through the bar to discover new sounds based on known patterns and techniques. This procedure follows traditional methods of developing fluency. Though interesting patterns were discovered, they were hardly novel.

Figure 7.2: Exploring groupings of four notes within a quarter note.

7.2.1 E-Pedal Position

The E-Pedal foot position allows for control of various synthesis parameters. To exact control over this many exercises were devised. The simplest use of the foot position is to alternate between two positions. The following exercise alternates between the foot all the way forward and then fully retracted.

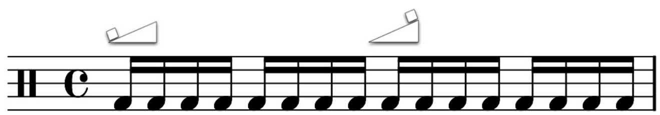


Figure 7.3: Simple exercise changing foot position.

In some situations gradually shifting the footboard position is more appropriate. The following is an example of the notation used for moving the foot position from rear to front.

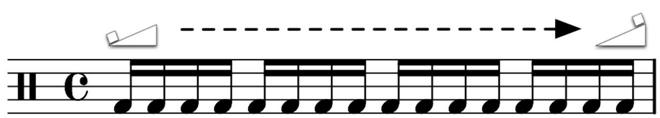


Figure 7.4: Gradually changing foot position.

7.2.2 Polytemporal Approaches

The first exercise that was developed towards mental practicing was a polyrhythm primer that utilized increased groupings but with a different figure at the end of the bar to ensure alternating feet. When there are odd numbers of notes in a bar and the figure is played with single strokes then the lead foot will change every bar. Each bar of the exercise has an odd number of notes for this reason. The set of exercises is meant to proceed from the beginning to end.

Various polytemporal exercises were developed that combined two or more patterns from two to eight. These types of patterns are uncommon but do arise in

Figure 7.5: Alternating leading feet via increasing groupings.

contemporary popular music as sources of rhythmic dissonance. The Swedish technical metal band Meshuggah utilizes polytemporal techniques to assure regularity, as the cycles may be heard, but to create an uncomfortableness because the cycles did not line up until the resolution point. A clear example of a 3/16 with a 4/4 polytemporal pattern is the focal point of the song Bleed by Meshuggah.

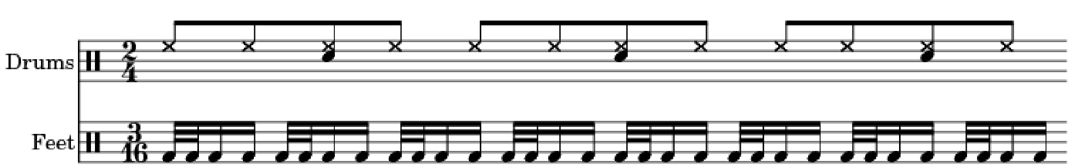


Figure 7.6: Polytemporal pattern in Meshuggah's Bleed.

The following figures were composed for a solo performance with tape (See Figures 7.2.2 and 7.2.2). The figures are applications of the mental practicing developed during the evolution of the E-Drumset. The first figure demonstrates a pattern in seven where 4/4 and 3/16 patterns are laid over top. The resulting pattern takes forty-nine bars to resolve fully.

Figure 7.7: Polymetric patterns played in performance.

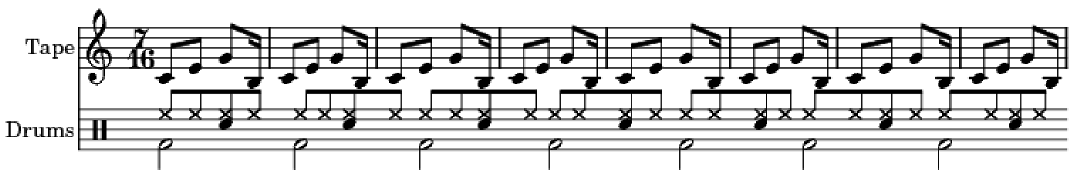


Figure 7.8: Polymetric pattern using a rhythmic illusion of a rock beat.

7.3 Conclusion

The exercises in this chapter represent a sampling of the exercises that were developed during the course of the project. The exercises evolved during the course of the study to become more conceptual and less physical, though equally difficult to perform. As the goals of the exercises changed to be more conceptual, the usefulness to other instruments other than the E-Drumset greatly increased.

Polyrhythms and polytemporal patterns increase the mental capacity of the player. Examples in contemporary popular music demonstrates that these techniques are not only parlor tricks but musical tools when used appropriately.

CHAPTER

EIGHT

Discussion

Not everything that can be counted counts, and not everything that counts can be counted.
– Albert Einstein, (attributed)

8.1 Introduction

The E-Drumset is a new instrument that has been a part of numerous performances. The combination of the new possibilities offered by the E-Drumset along with the author's commitment to improve facility have made the instrument a success. The E-Drumset software has also been applied to other areas of research such as assistive technology and music browsing.

The drum is one of the oldest instruments in human culture. With the E-Drumset drums are making new steps to adapting in an increasingly technological world. As the world continues to change hopefully the E-Drumset will continue to draw upon new technological advances.

8.2 Task Division

Although MIDI is problematic by today's standards almost every device or interface uses it. MIDI brought the idea of a universal communication protocol a standard feature in electronic music contexts. Transmission of data via network protocols, specifically Open Sound Control, is accessible, fast, and easy to implement. Flexible communication protocols such as Open Sound Control make it possible to reuse code and modularize development well. During the course of research Open Sound Control was used to deliver analysis data from Marsyas to ChucK, Pure Data and Max/MSP. During some concerts the analysis layer was not used, and it made no difference to the synthesis patches. Conversely, the analysis software was able to run independently for testing purposes. During the initial phases, software and hardware had to be connected and started in a specific order to ensure correct operation. This process harkens back to the days of SCSI, long over for all but system administrators. During the development of the Strike-A-Tune this separation allowed for rapid development of a prototype, which lead to further refinement.

8.3 Open Source

Open Source software is likely the most popular example of the Open Source philosophy, but it is by no means the only example of such practice. Artists and hardware developers open source parts or the whole of their work for others. The releasing of the components of the research has allowed other researchers to measure the performance of their systems, providing other researchers with existing data sets. The data sets collected for this research have been used by other researchers in the community. You can't win if you don't play. It is often difficult to tell what elements of one's research will yield the most benefit to the community.

8.4 Change is good

The hardware and software components of this research underwent many stages, rewrites, and ports. Technology changes quickly and sometimes it is appropriate to change or adapt by moving code bases from one language to another. Many patches began in one of Max/MSP, Pure-Data, Marsyas, Processing, or ChucK, and were ported to another language. This process illuminates subtle differences in the expression of ideas.

The hardware components changed very dramatically in some cases. The sensor analog to digital converter used in the E-Pedal moved through many stages of development. Initially the Teabox was utilized [91]. The Teabox was replaced with a custom built Atmel controller developed by Mark Marshall [104]. The Atmel system was then replaced with the Arduino platform. Although there was a great deal of reimplementation during this process each iteration added concepts or features due to subtle shifts in thought between different platforms.

There is an inherent danger in shifting to new platforms, programming languages, or technological advancements: stalling. Before a shift, one should evaluate the needs of a project. What are the needs of the project? Does a component need to be faster? More precise? More compatible with current systems? More durable? There will always be a new tool. There will always be better tools. Realizing a project using a new tool may prove to enhance the project, either by nature of the superiority of the new tool, or it may yield a new view on the project that can be implemented within the current tool.

8.5 Simple

When I first conceived of position tracking techniques for percussion controllers I imagined a very complicated system that was highly customized with software and hardware. After years of experiments and "happy accidents" I discovered implicit methods of position tracking via timbre recognition and direct sound manipulation.

With these discoveries, the resulting system became less and less complicated and reliant on custom hardware.

Complexity is a dangerous issue in computer music performance. One can forget the intricacies of mappings during performance and ruin a concert. One can spend a great deal of time between pieces or sections reconfiguring the instrument. These are just a few examples of potential disasters.

Complexity isn't wrong but it often detracts from the intuitiveness of an interface. If the goal of a larger audience is primary for the designer, then preserving intuitiveness should be considered when adding anything to the design. Also, one should take away as many components as possible until the usefulness of the device is affected.

8.6 Future Work

The two most obvious and simple items for future development are to play the E-Drumset more and to improve the software, both synthesis and analysis. The use of the E-Drumset in performance has yielded different ideas for sound and technique development that then translate to more avenues to explore in performance that illuminate more possibilities. The E-Drumset is my instrument for life. It will evolve over time, in configuration, sound, and control mechanisms.

Another item of future work will be to get more people involved with the project. The long term goal will be to have some performer other than myself be the main performer of the instrument. Another similar avenue will be to make it accessible enough to the average drummer so that they feel comfortable adding it as a component to their drum setup.

More experiments need to be conceived and carried out with the analysis software. The results of differentiating between left and right hands makes for some interesting potentials. There is definitely a difference in sound between the hands, and the results of the right versus left experiment lead one to the believe that the difference is sound is actually consistent and could potentially be differentiated. What other microcosms of gesture are recognizable? What are the uses? One could conceive of using these as control mechanisms. Capacitance-based percussion trackers are able to track different hands. This possibility on a non restrictive system is exciting. Other micro gestures could be tracked and used for pedagogy. One could track the difference between strokes and aim for the variation to decline over time, demonstrating a more consistent stroke. These results could be recorded to show students their progress over time as encouragement or reinforcement.

There are a number of gestures that have not been explored in the classification section. Flams and dampening are major gestures to be explored in the future. During the study, at no time was there the desire to utilize a flam or a dampening gesture to control sound. If there is no need in performance then there is little need to stress the system by having to classify another gesture. Although these gestures were not used for my performances they may be useful to others, therefore warranting study.

Portable devices now offer significant power that was not available during the beginning of the study. This yields the possibility of porting the software to a mobile platform and leveraging one of the big advantages of acoustic instruments: portability. Being able to perform or practice in the woods or for a limited time in a power outage could provide new and interesting opportunities to the electronic music artist.

New possibilities are on the horizon for this body of research to expand and change. Hopefully, this work will continue to grow and find new applications for liberating artists.

BIBLIOGRAPHY

[1] R. Aimi. New expressive percussion instruments. Master's thesis, Massachusetts Institute of Technology, 2002.

[2] R. Aimi. *Hybrid Percussion: Extending Physical Instruments Using Sampled Acoustics*. PhD thesis, Massachusetts Institute of Technology, 2007.

[3] M. Allen, J. Gluck, K. MacLean, and E. Tang. An initial usability assessment for symbolic haptic rendering of music parameters. In *ICMI '05: Proceedings of the 7th international conference on Multimodal interfaces*, pages 244–251. ACM Press, 2005.

[4] J. Allison and T. Place. Teabox: A sensor data interface system. *Proceedings of the Conference on New Interfaces for Musical Expression*, 2005.

[5] M. B. Alonso and D. V. Keyson. Musiccube: making digital music tangible. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1176–1179. ACM Press, 2005.

[6] X. Amatriain, J. Bonada, A. Loscos, J. Arcos, and V. Verfaille. Content-based transformations. *Journal of New Music Research*, 32(1), 2003.

[7] P. Appan, H. Sundaram, and B. Tseng. Summarization and visualization of communication patterns in a large social network. In *The 10th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2006.

[8] D. Arfib, J. M. Couturier, L. Kessous, and V. Verfaille. Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Org. Sound*, 7(2):127–144, 2002.

[9] S. Bagchee. *Understanding Raga Music.* Ceshwar Business Publications Inc., Mumbai, India, 1998.

[10] J. Beauchamp. Synthesis by spectral amplitude and "brightness" matching of analyzed musical tones. *Journal of the Acoustical Society of America*, 30(6):396–406, 1982.

[11] J. Bello, G. Monti, and M. Sandler. Techniques for automatic music transcription. *International Symposium on Music Information Retrieval*, 2000.

[12] K. Berger. Some factors in the recognition of timbre. *Journal of the Acoustical Society of America*, 36(10):1888–91, 1964.

[13] J. Blades. *Percussion instruments and their history.* Bold Strummer Ltd., Westport, CT, 1997.

[14] J. Blimes. Timing is of the essence: Perceptual and computational techniques for representing, learning, and reproducing expressive timing in percussive rhythm. Master's thesis, MIT, 1993.

[15] B. Boie, M. Mathews, and A. Schloss. The radio drum as a synthesizer controller. In *Proceedings of the International Computer Music Conference*, pages 42–45, 1989.

[16] A. Bregman. *Auditory scene analysis: The perceptual organization of sound.* MIT Press, 1994.

[17] R. Bresin, K. F. Hansen, and S. Dahl. The radio baton as configurable musical instrument and controller. *Proceedings of the Stockholm Music Acoustics Conference*, pages 689–691, 2003.

[18] P. Cano, M. Kaltenbrunner, F. Gouyon, and E. Battle. On the use of fastmap for audio retrieval and browsing. In *Proceedings of the International Symposium on Music Information Retrieval*, Paris, France, 2002.

[19] P. Cano, M. Koppenberger, and N. Wack. An industrial-strength content-based music recommendation system. *SIGIR Forum*, page 673, 08/15 2005.

[20] L. Cohen. Time-frequency distributions: A review. *Proceedings of the IEEE*, 77(7):941–81, 1989.

[21] P. Cook. Principles for designing computer music controllers. In *New Interfaces for Musical Expression*, 2001.

[22] P. Cook. Principles for designing computer music controllers. In *NIME '01: Proceedings of the 2001 conference on New interfaces for musical expression*, pages 1–4, Singapore, Singapore, 2001. National University of Singapore.

[23] S. Dahl. Arm motion and striking force in drumming. *Proceedings of the International Symposium on Musical Acoustics*, 1:293–296, 2001.

[24] S. Dahl. Playing the accent - comparing striking velocity and timing in an ostinato rhythm performed by four drummers. *Acta Acustica united with Acustica*, 90(4), 2004.

[25] I. Debauchies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41:909–96, 1988.

[26] P. Depalle and T. Helie. Extraction of spectral peak parameters using a short time fourier transform modeling and no sidelobe windows. *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1997.

[27] G. Depoli and P. Prandoni. Sonological models for timbre characterization. *Journal of the Acoustical Society of America*, 26(2):170–97, 1997.

[28] N. Derrett. Hickel's law:conclusions from the user interface design of a music appliance- the bassoon. *Personal & Ubiquitous Computing*, 8(3/4):208, 07/01 2004.

[29] T. A. Driscoll. Eigenmodes of isospectral drums. *SIAM Review*, 39:1–17, 1997.

[30] R. Duda, P. Hart, and D. Stork. *Pattern classification*. John Wiley and Sons, Inc., New York, 2000.

[31] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, Inc., New York, 2000.

[32] J. W. Dunn and C. A. Mayer. Variations: A digital music library system at indiana university. In *Proceedings of the Fourth ACM International Conference on Digital Libraries*, 1999.

[33] E. B. Egozy. Deriving musical control features from a real-time timbre analysis of the clarinet. Master's thesis, Massachusetts Institute of Technology, 1995.

[34] G. T. Elliott and B. Tomlinson. PersonalSoundtrack: context-aware playlists that adapt to user pace. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*, pages 736–741. ACM Press, 2006.

[35] G. Essl. The principle of closed wavetrains, resonance and efficiency: The missing link revisited. In *Proceedings of the International Computer Music Conference (ICMC-02*, 2002.

[36] G. Essl and P. R. Cook. Banded waveguides: Towards physical modeling of bar percussion instruments. *Proceedings of the International Computer Music Conference*, pages 321–324, 1999.

[37] M. Fernström and L. Bannon. Multimedia browsing. In *Proc. of the International Conference on Computer Human Interaction (CHI)*, 1997.

[38] M. Fernström and E. Brazil. Sonic browsing: An auditory tool for multimedia

asset management. In *Proceedings of the Internation Conference on Auditory Display*, 2001.

[39] A. Field, P. Hartel, and W. Mooij. Personal dj, an architecture for personalised content delivery. In *WWW '01: Proceedings of the 10th international conference on World Wide Web*, pages 1–7. ACM Press, 2001.

[40] D. FitzGerald, E. Coyle, and B. Lawlor. Sub-band independent subspace analysis for drum transcription. *International Conference on Digital Audio Effects*, pages 65–9, 2002.

[41] N. Fletcher and T. Rossing. *The physics of musical instruments.* Springer-Verlag, 1990.

[42] W. Flür. *I Was a Robot.* Sanctuary Publishing Ltd., London, England, 2005.

[43] E. Frank and M. Hall. A simple approach to ordinal classification, 2001.

[44] A. Fraser and I. Fujinaga. Toward realtime recognition of acoustic musical instruments. *Proceedings of the International Computer Music Conference*, pages 175–7, 1999.

[45] J. Fuernkranz. Round robin rule learning. *Proceedings of the 18th International Conference on Machine Learning*, 2001.

[46] I. Fujinaga. Machine recognition of timbre using steady-state tone of acoustic musical instruments. *Proceedings of the International Computer Music Conference*, pages 207–10, 1998.

[47] I. Fujinaga and K. MacMillan. Toward realtime recognition of orchestral instruments. *Proceedings of the International Computer Music Conference*, pages 175–7, 2000.

[48] W. W. Gaver, J. Bowers, A. Boucher, H. Gellerson, S. Pennington, A. Schmidt, A. Steed, N. Villars, and B. Walker. The drift table: designing for ludic

engagement. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 885–900. ACM Press, 2004.

[49] J. J. Gibson. *The Ecological Approach to Visual Perception.* Lawrence Erlbaum, 1986.

[50] M. Goto. An audio-based real-time beat tracking system for music with or without drum-sounds. *Journal of New Music Research*, pages 159–71, 2001.

[51] M. Goto. An audio-based real-time beat tracking system for music with or without drum-sounds. *Journal of New Music Research*, 30(2):159–171, 2001.

[52] M. Goto and Y. Muraoka. A beat tracking system for acoustic signals of music. *Proceedings of the ACM International Conference on Multimedia*, pages 365–72, 1994.

[53] M. Goto and Y. Muraoka. Real-time rhythm tracking for drumless audio signals: Chord change detection for musical decisions. *Proceedings of the Workshop on Computational Auditory Scene Analysis. International Joint Conference on Artificial Intelligence*, 1997.

[54] M. Gould. Alternative sounds and techniques on drumset. *Percussive Notes*, 92(4):16–24, 2000.

[55] M. Gould and A. Thompson. Alternative sounds and techniques on drumset. *Percussive Notes*, 38(2):16–24, 2000.

[56] F. Gouyon and P. Herrera. Exploration of techniques for automatic labeling of audio drum tracks' instruments. *Proceedings of MOSART: Workshop on Current Directions in Computer Music*, 2001.

[57] F. Gouyon and P. Herrera. Exploration of techniques for automatic labeling of audio drum tracks' instruments. *Proceedings of MOSART: Workshop on Current Directions in Computer Music*, 2001.

[58] F. Gouyon, F. Pachet, and O. Delerue. On the use of zero-crossing rate for an application of classification of percussive sounds. *International Conference on Digital Audio Effects*, 2000.

[59] R. Gray and L. Davisson. *An introduction to statistical signal processing.* Cambridge University Press, Cambridge, 2004.

[60] J. Grey. Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, 61(5):1270–7, 1977.

[61] J. Grey. Timbre discrimination in musical patterns. *Journal of the Acoustical Society of America*, 64(2):467–72, 1987.

[62] J. Grey and J. Gordon. Perceptual effects of spectral modification on musical timbres. *Journal of the Acoustical Society of America*, 63(5):1493–500, 1978.

[63] J. Grey and J. Moorer. Perceptual evaluation of synthesized musical instrument tones. *Journal of the Acoustical Society of America*, 62(2):454–62, 1977.

[64] M. Grimaldi, A. Kokaram, and P. Cunningham. Classifying music by genre using a discrete wavelet transform and a round-robin ensemble. Technical Report TCD-CS-2002-64, Trinity College, University of Dublin, Ireland, 2003.

[65] S. Hainsworth, M. Macleod, and P. Wolfe. Analysis of reassigned spectrograms for musical transcription. *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 23–6, 2001.

[66] K. F. Hansen and R. Bresin. DJ scratching performance techniques: Analysis and synthesis. *Proceedings of the Stockholm Music Acoustics Conference*, pages 693–696, 2003.

[67] T. Heittola and A. Klapuri. Locating segments with drums in music signals. *International Symposium on Music Information Retrieval*, pages 271–2, 2002.

[68] H. Helmholtz. *On the Sesnation of Tone.* Courier Dover Publications, N. Chemsford, MA, 1863.

[69] C. Henzie. *Amplitude and duration characteristics of snare drum tones.* PhD thesis, Indiana University, 1960.

[70] P. Herrera, X. Amatriain, E. Batlle, and X. Serra. Towards instrument segmentation for music content description: A critical review of instrument classification techniques. *International Symposium on Music Information Retrieval*, 2000.

[71] P. Herrera, A. Dehamel, and F. Gouyon. Automatic labeling of unpitched percussion sounds. *Proceedings of the Audio Engineering Society*, pages 1378–91, 2003.

[72] P. Herrera, G. Peeters, and S. Dubnov. Automatic classification of musical instruments. *Journal of New Music Research*, 32(1):3–21, 2003.

[73] P. Herrera, A. Yetarian, and F. Gouyon. Automatic classification of drum sounds: A comparison of feature selection and classification techniques. *Proceedings of Second International Conference on Music and Artificial Intelligence*, pages 79–91, 2002.

[74] P. Herrera-Boyer, X. Amatriain, E. Batlle, and X. Serra. Towards instrument segmentation for music content description: a critical review of instrument classification techniques. In *Proceedings of the 1st International Conference on Music Information Retrieval*, 2000.

[75] G. Holmes, M. Hall, and E. Frank. Generating rule sets from model trees. In *In Proceedings of the 12th Australian Joint Conference on Artifical Intelligence*, pages 1–12. Springer-Verlag, 1999.

[76] R. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11:63–91, 1993.

[77] Y. Hu and J. Hwang, editors. *Handbook of neural network signal processing.* CRC Press, Boca Raton, FL, 2002.

[78] R. Huott. An interface for precise musical control. In *Proceedings of the Conference on New Interfaces for Musical Expression*, 2002.

[79] R. Huott. An interface for precise musical control. *Proceedings of the Conference on New Interfaces for Musical Expression*, pages 1–5, 2002.

[80] S. I. Interaction and T. Blaine. The jam-o-drum interactive music system:. In *In DIS 2000: Proceedings of the conference on Designing interactive systems*, pages 165–173. ACM Press, 2000.

[81] E. Isaacson. What you see is what you get: On visualizing music. In *Proceedings of the 6th International Conference on Music Information Retrieval*, 2005.

[82] K. Jensen. *Timbre models of musical sounds*. PhD thesis, University of Copenhagen, 1999.

[83] K. Jensen. The timbre model. *Proceedings of MOSART: Workshop on Current Directions in Computer Music*, 2002.

[84] A. Kapur, P. Davidson, P. R. Cook, P. Driessen, and A. Schloss. Digitizing north indian performance. *Proceedings of the International Computer Music Conference*, 2004.

[85] A. Kapur, G. Essl, P. Davidson, and P. R. Cook. The electronic tabla controller. *Journal for New Music Research*, 2002.

[86] M. Karjalainen and C. Erkut. Digital waveguides versus finite difference structures: equivalence and mixed modeling. *EURASIP J. Appl. Signal Process.*, 2004:978–989, 2004.

[87] M. Kashima. The bounded-Q approach to time-varying spectral analysis. Technical Report STAN-M-23, Stanford University, 1985.

[88] D. A. Keim, J. Schneidewind, and M. Sips. Circleview: a new approach for visualizing time-related multidimensional data sets. In *AVI '04: Proceedings*

of the working conference on advanced visual interfaces, pages 179–182. ACM Press, 2004.

[89] R. Kendall. The role of acoustic signal partitions in listener categorization of musical phrases. *Music Perception*, 4(2):185–214, 1986.

[90] R. Kendall and E. Carterette. Difference thresholds for timbre related to spectral centroid. *Proceedings of the International Conference on Music Perception and Cognition*, pages 91–5, 1996.

[91] J. A. Kimberly and J. T. Allison. Teabox: A sensor data interface system. In *In Proceedings, New Interfaces for Musical Expression*, pages 699–701, 2005.

[92] L. Kinsler and A. Frey. *Fundamentals of Acoustics*. John Wiley and Sons, Inc., New York, 1950.

[93] A. Klapuri. Sound onset detection by applying psychoacoustic knowledge. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, 6:3089–92, 1999.

[94] R. Klatzky, D. Pai, and E. Krotkov. Perception of material from contact sounds. *Presence: Teleoperators and Virtual Environments*, 9:399–410, 2000.

[95] V. Krefeld. The hand in the web: An interview with Michel Waiswisz. *Computer Music Journal*, 4(2):28–33, 1990.

[96] H. Lam and T. Munzner. Musicland: Exploratory browsing in music space. In *InfoVis 2005 (Poster), Minneapolis, MN, USA*, 2005.

[97] E. Lank, J. Ruiz, and W. Cowan. Concurrent bimanual stylus interaction: a study of non-preferred hand mode manipulation. In *GI '06: Proceedings of the 2006 conference on Graphics interface*, pages 17–24. Canadian Information Processing Society, 2006.

[98] J. C. Lee. *Projector-Based Location Discovery and Tracking*. PhD thesis, Carnegie Mellon University, 2009.

[99] N. Lee. Interview with Tina Blaine. *Computers in Entertainment*, 5(2), 2007.

[100] R. Lewis and J. Beckford. Measuring tonal characteristics of snare drum batter heads. *Percussive Notes*, 38(3):69–71, 2000.

[101] B. Logan. Mel-frequency cepstrum coefficients for music modeling. *International Symposium on Music Information Retrieval*, 2000.

[102] D. Lubbers. Sonixplorer: Combining visualization and auralization for content-based explorations of music collections.

[103] T. Machover. Hyperinstruments: A progress report 1987-1991. Media lab document, Massachusetts Institute of Technology, 1992.

[104] M. Marshall. *Physical Interface Design for Digital Musical Instruments*. PhD thesis, McGill University, 2009.

[105] M. Marshall, M. Rath, and B. Moynihan. The virtual bodhran - the vodhran. *Proceedings of the New Interfaces for Musical Expression Conference*, 2002.

[106] K. Martin. *Sound-source recognition: A theory and computational model*. PhD thesis, MIT, 1999.

[107] K. Martin and Y. Kim. Musical instrument identification: A pattern-recognition approach. *Journal of the Acoustical Society of America*, 104(3):1768, 1998.

[108] P. Masri and A. Bateman. Improved modelling of attack transients in music analysis-resynthesis. *Proceedings of the International Computer Music Conference*, pages 100–3, 1996.

[109] M. Mathews. The radio baton and conductor program. *Computer Music Journal*, 19(4), 1991.

[110] M. Mathews. The radio baton and conductor program, or: Pitch, the most important and least expressive part of music. *Computer Music Journal*, 15(4):37–46, 1991.

[111] Mathworks. *MATLAB reference guide*. The Mathworks, Inc., Natick, MA, 1992.

[112] R. Mattingly and B. Gatzen. Electronic review - Simmons SDX. *Modern Drummer*, pages 64–67, 1985.

[113] S. McAdams, J. Beauchamp, and S. Meneguzzi. Discrimination of musical instrument sounds resynthesized with simplified spectral parameters. *Journal of the Acoustical Society of America*, 105(2):882–97, 1999.

[114] B. Mcnulty. Mario deciutiis and the evolution of virtual percussion. *Percussive Notes*, 44(2):48–50, 2006.

[115] R. R. Menon. *Discovering Indian Music*. Somaiya Publications PVT. LTD, Mumbai, India, 1974.

[116] K. Micallef. Interview with Bill Bruford. *Modern Drummer*, 2001.

[117] T. Mitchell. *Machine Learning*. McGraw Hill, Columbus, OH, 1997.

[118] S. A. Moeller. *The Moeller Book*. Ludwig Music Publishing, Boca Raton, FL, 1954.

[119] J. Murdoch and G. Tzanetakis. Interactive content-aware music browsing using the radio drum. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2006.

[120] T. Mki-Patola and P. Hmlinen. Latency tolerance for gesture controlled continuous sound instrument without tactile feedback. In *Proc. International Computer Music Conference (ICMC*, pages 409–416, 2004.

[121] B. Neville, P. Driessen, and W. Schloss. A new control paradigm: software-based gesture analysis for music. *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and signal Processing*, 1:360 – 363, 2003.

[122] R. Nosofsky. Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(1):104–14, 1984.

[123] R. Nosofsky. Attention, similarity, and the identification categorization relationship. *Journal of Experimental Psychology: General*, 115(1):39–57, 1986.

[124] M. Notess and N. Minibayeva. Variations2: Toward visual interfaces for digital music libraries. In *Second International Workshop on Visual Interfaces to Digital Libraries*, 2002.

[125] N. Oliver and F. Flores-Mangas. MPTrain: a mobile, music and physiology-based personal trainer. In *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pages 21–28. ACM Press, 2006.

[126] N. Oliver and L. Kreger-Stickles. Papa: Physiology and purpose-aware automatic playlist generation. In *Proceedings of the Internation Conference on Music Information Retrieval*, pages 250–253, 2006.

[127] F. Opolko and J. Wapnick. Mcgill university master samples. Collection of Samples., 1987.

[128] E. Pampalk, S. Dixon, and G. Widmer. Exploring music collections by browsing different views. In *Proceedings of the 4th International Conference on Music Information Retrieval*, 2003.

[129] E. Pampalk and M. Goto. MusicRainbow: A new user inferace to discover artists using audio-based similarity and web-based labeling. *International Conference on Music Information Retieval*, pages 367–370, 2006.

[130] J. Paradiso. The brain opera technology: New instruments and gestural sensors for musical interaction and performance. *Journal of New Music Research*, 28:130–149, 1999.

[131] J. Paradiso. Wearable wireless sensing for interactive media. *First International Workshop on Wearable & Implantable Body Sensor Networks*, 2004.

[132] J. Paradiso, E. Hu, and K. Hsiao. Instrumented footwear for interactive dance. *Colloquium on Musical Informatics*, pages 89–92, 1998.

[133] J. Paradiso, C. Leo, N. Checka, and K. Hsiao. Passive acoustic sensing for tracking knocks atop large interactive displays. *IEEE Sensors Journal*, 17(17):1–7, 2002.

[134] J. Paradiso, C. Leo, N. Checka, and K.Hsiao. Passive acoustic sensing for tracking knocks atop large interactive displays. In *Proceedings of the 2002 IEEE International Conference on Sensors*, 2002.

[135] J. Paulus and A. Klapuri. Model-based event labelling in the transcription of percussive audio signals. *International Conference on Digital Audio Effects*, 2003.

[136] S. Pauws, D. Bouwhuis, and B. Eggen. Programming and enjoying music with your eyes closed. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 376–383. ACM Press, 2000.

[137] M. Peinado, B. Herbelin, M. Wanderley, B. L. Callennec, R. Boulic, and D. Thalmann. Towards configurable motion capture with prioritized inverse kinematics. In *Proceedings of the Third International Workshop on Virtual Rehabilitation*, 2004.

[138] R. Prinz. Radio drum mouse controller. Technical report, University of Victoria, 2004.

[139] M. S. Puckette, , T. Apel, and D. D. Zicarelli. Real-time audio analysis tools for pd and msp. In *Proceedings of the International Conference on Computer Music*, pages 109–112, 1998.

[140] W. Putnam and R. Knapp. Input/data acquisition system design for human computer interfacing. 1996.

[141] J. C. Risset. *An Introductory Catalog of Computer-synthesized Sounds*. Bell Laboratories, 1969.

[142] F. Rossenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psycological Review*, 65:386–408, 1958.

[143] T. Rossing. The acoustics of percussion instruments. *Percussionist*, 19(3):6–83, 1982.

[144] T. Rossing. *The science of percussion instruments*. World Scientific, River Edge, NJ, 2000.

[145] T. Rossing. *The science of sound*. Addison-Wesley Publications Company, San Francisco, 2002.

[146] G. Rule. Keyboard reports: Korg wavedrum. *Keyboard*, 21(3), 1995.

[147] E. Saldanha and J. Corso. Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, 36(11):2021–6, 1964.

[148] G. Sandell. *SHARC timbre database*. http://www.parmly.luc.edu.sharc/, 1994.

[149] E. Scheirer. Tempo and beat analysis of acoustic musical signals. *Journal of the Acoustical Society of America*, 103(1):588–601, 1998.

[150] S. Schertenleib, M. GutiA(C)rrez, V. FrA(C)dA(C)ric, and D. Thalmann. Conducting a virtual orchestra. *IEEE Multimedia*, 11(3):40, 07/01 2004.

[151] A. Schloss. *On the automatic transcription of percussive music: From acoustic signal to high-level analysis*. PhD thesis, CCRMA, Stanford University, 1985.

[152] A. Schloss. Using contemporary technology in live performance: The dilemma of the performer. *Journal of New Music Research*, 32(3), 2003.

[153] S. Sharma. *Comparative Study of Evolution of Music in India & the West*. Pratibha Prakashan, New Delhi, India, 1997.

[154] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *VL '96: Proceedings of the 1996 IEEE Symposium on Visual Languages*, page 336. IEEE Computer Society, 1996.

[155] J. Silpanpää. Drum stroke recognition. Technical report, Tampere University of Technology, Tampere, Finland, 2000.

[156] J. Silpanpää, A. Klapuri, J. Seppänen, and T. Virtanen. Recognition of acoustic noise mixtures by combined bottom-up and top-down processing. *Proceedings of the European Signal Processing Conference*, 2000.

[157] J. Smith. Physical modeling using digital waveguides. *Computer Music Journal*, 16(4):74–91, 1992.

[158] J. O. Smith. Physical modeling synthesis update, 2000.

[159] A. Srinivasan, D. Sullivan, and I. Fujinaga. Recognition of isolated instrument tones by conservatory students. *Proceedings of the 7th International Conference on Music Perception and Cognition*, pages 720–3, 2002.

[160] I. Stavness, J. Gluck, L. Vilhan, and S. Fels. The MUSICtable: A map-based ubiquitous system for social Interaction with a digital music collection. In *International Conference on Entertainment Computing*, pages 291–302, 2005.

[161] L. Stevens. *Method of movement.* Marimba Productions, Asbury Grove, NJ, 1979.

[162] A. Tanaka. Musical performance practice on sensor-based instruments. *Trends in Gestural Control of Music*, pages 389–406, 2000.

[163] A. Tindale. Classification of snare drum sounds with neural networks. Master's thesis, McGill University, 2004.

[164] A. Tindale. The e-drumset. *West Coast Composers Seminar*, 2005.

[165] A. Tindale, M. Benning, and N. Livingston. Virtual instrument control for music therapy. *In Proceedings of the European Conference for the Advancement of Assistive Technology*, pages 414 – 419, 2007.

[166] A. Tindale and R. Davies. i send data live. *Interactive Futures*, 2006.

[167] A. Tindale, A. Kapur, and G. Tzanetakis. ining surrogate sensors in musical gesture acquisition systems. *Submitted to the IEEE Sensors Journal*, 2009.

[168] A. Tindale, A. Kapur, G. Tzanetakis, and I. Fujinaga. Retrieval of percussion gestures using timbre lassification techniques. *International Symposium on Music Information Retrieval*, 2004.

[169] A. Tindale, A. Kapur, G. Tzanetakis, and A. Schloss. Timbre-recognition based instruments. *Conference of Interdisciplinary Musicology*, 2005.

[170] A. Tindale, D. Sprague, and G. Tzanetakis. Strike-a-tune: Fuzzy music navigation using a drum interface. *In Proceedings of the International Conference of Music Information Retrieval*, 2007.

[171] A. R. Tindale, A. Kapur, G. Tzanetakis, and I. Fujinaga. Retrieval of percussion gestures using timbre classification techniques. In *Proceedings of the 5th International Conference on Music Information Retrieval*, pages 541–544, 2004.

[172] T. Tolonen and M. M. Karjalainen. A computationally efficient multipitch analysis model. *IEEE Transactions on Speech and Audio Processing*, 8:708–716, 2000.

[173] D. Tomlinson. *World of Flams.* Alfred Publishing Company, Los Angeles, CA, 1999.

[174] M. Torrens, P. Hetzog, and J.-L. Arcos. Visualizing and exploring personal music libraries. In *Proceedings of the 5th International Conference on Music Information Retrieval*, 2004.

[175] C. Traube and P. Depalle. Deriving the plucking point location along a guitar string from the least-square estimation of a comb filter delay. *Proceedings of the Canadian Conference on Electrical and Computer Engineering*, 2003.

[176] C. Traube, P. Depalle, and M. Wanderley. Indirect acquisition of instrumental gestures based on signal, physical and perceptual information. *Proceedings of the Conference on New Musical Interfaces for Musical Expression*, pages 42–7, 2003.

[177] G. Tzanetakis and P. Cook. Marsyas: A framework for audio analysis. *Organized Sound*, 4(3), 2000.

[178] G. Tzanetakis and P. Cook. Marsyas3d: A prototype audio browser-editor using a large scale immersive visual and audio display. In *Proc. of the International Conference on Auditory Display(ICAD)*, 2001.

[179] G. Tzanetakis and P. Cook. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5), 2002.

[180] G. Tzanetakis, G. Essl, and P. Cook. Audio analysis using the discrete wavelet transform. *Proceedings of Conference in Music Theory Applications*, 2001.

[181] G. Tzanetakis, G. Essl, and P. Cook. Automatic musical genre classification of audio signals. *Proceedings of the International Conference on Music Information Retrieval*, pages 2005–210, 2001.

[182] R. van Gulik and F. Vignoli. Visual playlist generation on the artist map. In *Proceedings of the 6th International Conference on Music Information Retrieval*, 2005.

[183] R. van Gulik, F. Vignoli, and H. van de Wetering. Mapping music in the palm of your hand, explore and discover your collection. In *Proc. of the 5th International Conference on Music Information Retrieval*, 2004.

[184] T. Verderosa. *The Techno Primer.* Hal Leonard, Milwaukee, WI, 2002.

[185] A. Voida, R. E. Grinter, N. Ducheneaut, W. K. Edwards, and M. W. Newman. Listening in: practices surrounding itunes music sharing. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 191–200. ACM Press, 2005.

[186] M. M. Wanderley and N. Orio. Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal*, 26(3):62–76, 2002.

[187] G. Wang and P. R. Cook. ChucK: A concurrent, on-the-fly audio programming language. *Proceedings of the International Computer Music Conference*, pages 219–226, 2003.

[188] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., 2004.

[189] N. Weinberg. *The Electronic Drummer*. Hal Leonard, Milwaukee, WI, 1989.

[190] N. Weinberg. *Guide to Standardized Drumset Notation*. Percussive Arts Society Press, Lawton, OK, 1998.

[191] D. Wheeler. Focus on research: Some experiments concerning the effect of snares on the snare drum sound. *Percussive Notes*, 27(3):48–52, 1989.

[192] I. Witten, E. Frank, and M. Kaufmann. *Data mining: Practical machine learning tools with Java implementations*. Addison-Wesley Publications Company, San Francisco, 2000.

[193] M. Wright, A. Freed, and A. Momeni. Opensound control: State of the art 2003. In *New Interfaces for Musical Expression*, pages 153–159, 2003.

[194] D. Young and I. Fujinaga. Aobachi: A new interface for japanese drumming. *Proceedings of the Conference on New Musical Interfaces for Musical Expression*, 2004.

[195] M. Zadel and G. Scavone. Laptop performance: Techniques, tools, and a new interface design. *Proceedings of the International Computer Music Conference*, pages 643–648, 2006.

[196] H. Zhao. Acoustics of snare drums: An experimental study of the modes of vibration, mode coupling and sound radiation patterns. Master's thesis, Northern Illinois University, 1990.

[197] J. Zhu and L. Lu. Perceptual visualization of a music collection. In *Multimedia and Expo, 2005. (ICME 2005).*, pages 1058–1061, July 2005.

[198] U. Zoelzer, editor. *Dafx: Digital Audio Effects.* John Wiley & Sons, Inc., New York, NY, USA, 2002.

APPENDIX

A

Media Assets

Writing about the sound of a drum is difficult. Sometimes it is necessary to hear the drum in order to evaluate it properly. Sound and video examples have been prepared to accompany this document in order to enhance or reinforce statements made within. All of the media assets are available at my website (http://www.adamtindale.com/).

The audio contained on the website includes pieces by i send data live and individual samples of the E-Drumset illustrating the various synthesis methods available. The video assets include a duo performance of me with Chris Chafe (Director of CCRMA, Stanford University) playing his electric cello and a demo video showing response use of brushes, strike location, and a machine learning patch on the rim to change presets.

APPENDIX

B

Glossary

- **Acoustically excited physical models:** A synthesis technique that extends traditional physical modeling whereby the excitation is acquired from a microphone or other acoustic source. (pg. 72)
- **Affordances:** Affordances refers to the ability of the interface to allow for flexibility beyond its original purpose or primary function. This is the original meaning of the term as proposed by James Gibson. The Human-Computer Interaction (HCI) community uses the term to mean the inviting of an action due to the object design. (pg. 91)
- **Crosstalk:** The interference caused when a signal is received by an adjacent piece of circuitry. In electronic drums the vibrations caused when hitting a drum resonate through the hardware and are picked up by neighboring pickups. (pg. 30)
- **Descending pitch model:** The toms of a drumset are typically arranged from smallest to largest around the player. The result is that as the player moves from one to the next the pitch descends. (pg. 87)
- **Drum Machine:** A drum machine is a drum synthesizer and sequencer in-

tegrated into a single unit. The synthesizer can be any number of sound generating devices but samplers and simple synthesis algorithms are popular. (pg. 11)

- **Drum Brain:** Converts electronic drum pad inputs into MIDI signals or sound by triggering samples contained in the unit. (pg. 34)

- **Drumset:** (alternate: drumkit, traps) A canonized collection of instruments consisting of a battery of drums and cymbals. The standard configuration varies for each genre of music but a hihat, bass drum and snare drum are in nearly every configuration. (pg. 1)

- **Hyper Instrument:** A concept of instrument building from MIT where acoustic instruments are extended by attaching sensors. The main idea of using an acoustic instrument is that in the event of a technology failure the player will still have a functional instrument in order to finish the performance. (pg. 74)

- **Marsyas** Marsyas (Music Analysis, Retrieval and Synthesis for Audio Signals) is an open source software framework for audio processing with specific emphasis on Music Information Retrieval applications. (pg. 81)

- **Mapping:** The translation of data from one form to another to make it optimally appropriate for the receiving side. Sensor data is often mapped to different values or ranges in order to create musical control. (pg. 5)

- **Meta Instrument:** An instrument whose function is to control other instruments or controllers. (pg. 74)

- **MIDI:** Short for Musical Instrument Digital Interface. MIDI is a serial protocol that has become a standard connection on most hardware synthesizers and control interfaces. (pg. 2)

- **Modal Synthesis:** A physical modelling technique where the modes of a sound are analyzed and then a synthesis algorithm reproduces the modes to emulate the original sound. (pg. 87)

- **Ordinal Regression:** A machine learning technique for classification problems where the classes are not discrete or unrelated classes but a gradient from one to another. (pg. 97)

- **OSC:** Short for Open Sound Control. OSC is a network protocol based on UDP for transmitting programmer defined information across network connected devices. (pg. 2)

- **Overfitting:** Machine learning algorithms aim to generalize their ability to classify by examining a training set based on future unknown input. Overfitting occurs when the algorithm fits too closely to the training set and fails to generalize well. (pg. 101)

- **RTAudio:** RTAudio is a software package developed by Gary Scavone at McGill University. The software is works across many operating systems to provide a unified API for interfacing with audio and MIDI hardware. (pg. 77)

- **Supercollider:** A software package originally developed by James McCartney that was further released as open source to the community. (pg. 85)

- **Surrogate Sensor:** A technique to use a sensor to train a machine learning algorithm on another sensor. The technique allows for lower cost sensors to be deployed after correlating and calibrating using a higher cost sensor. (pg. 7)

- **Timbre Recognition based instrument:** A instrument that uses machine learning techniques to determine the timbre produced by a controller and use that label as control information. (pg. 2)

- **Weka:** Weka is a Java package for machine learning and data mining developed at the University of Waikato. (pg. 96)

APPENDIX

C

Discography

This list provides a brief musical history of electronic percussion. Though not exhaustive, it provides a list of major recordings that use electronic percussion, as well as recordings that were influential in the development of my style of playing.

- John Cage – Imaginary Landscapes No. 1. 1960. Avakian 1.
- Edgar Varese – Arcana/Deserts/Offrandes, Columbia Symphony with Robert Craft. Columbia Records. 1962.
- Kraftwerk – Kraftwerk. 1973. RCA Victor.
- Emerson, Lake & Palmer – Brain Salad Surgery. 1973. Manticore Records.
- Gil Evans – There comes a time. 1975. RCA.
- Béla Fleck and the Flecktones – Béla Fleck and the Flecktones. 1980. Warner Bros.
- Phil Collins. Face Value. 1981. Virgin Records.
- King Crimson – Discipline. 1981. Virgin Records.
- King Crimson – Beat. 1982. Virgin Records.

- King Crimson – Three of a Perfect Pair. 1984. Virgin Records.
- Laurie Anderson – United States Live. 1984. Warner.
- Laurie Anderson – Mister Heartbreak. 1984. Warner.
- Akira Jimbo – Cotton. 1986. wint.
- Earthworks – Earthworks. 1987. EG Records.
- Earthworks – Dig?. 1989. EG Records.
- Akira Jimbo – Jimbo. 1990. wint.
- Béla Fleck and the Flecktones – Flight of the Cosmic Hippo. 1991. Warner Bros.
- Earthworks – All heaven broke loose. 1991. EG Records.
- Tom Waits – Bone Machine. 1992. Anti.
- Tom Waits – The Black Rider. 1993. Anti.
- Tony Verdersa – Beatnik Rebel Science. 1993. Self-Published.
- Earthworks – Stamping Grounds. 1994. EG Records.
- King Crimson – THRAK. 1995. Virgin Records.
- King Crimson – Thrak Attack. 1996. Virgin Records.
- Tool – Aenima. 1996. Zoo Entertainment.
- Terry Bozzio – Drawing the Circle. 1998. Self-Published.
- Béla Fleck and the Flecktones – Left of Cool. 1998. Warner Bros.
- Tom Waits – Mule Variations. 1999. Anti.
- Tool – Lateralus. 2001. Volcano Records.
- Tom Waits – Blood Money. 2002. Anti.

- Terry Bozzio – Solo Drum Music I/II & III. 2003. Rock Empire.
- Tool – 10,000 Days. 2006. Volcano Records.

APPENDIX

D

Breadth of Gesture

D.1 Introduction

Before designing a new interface based on percussive gestures a thorough survey of the already available gestures was conducted. This chapter outlines the available standard and extended drumset gestures. The repertoire is a collection of stick techniques, instrument specific techniques, and extended gestures.

D.2 Standard Stick Techniques

Drummers and percussionists have been devising formalized methods of technique for approximately 100 years. Though there has been formal percussion for a very long time, until recently each player would devise their own method of gripping the stick and associated movements. The goal of these systems is to maximize efficiency, power, and speed.

The Gladstone Technique of stick motion was invented by drum instructor Billy Gladstone (1893-1961). Gladstone propagated his ideas through high profile students, such as Joe Morello and Buddy Rich. To this day, it is rarely written down and is mainly transmitted from instructor to student. The technique breaks down all drumming motions into four strokes: full stroke, down stroke, up stroke, and the tap. The strokes revolve around the idea that there are two main positions when playing drums: high and low. These positions correlate to heights above the drum, usually 2 inches for the low position and 9 inches for the high position.

- Full Stroke: Start from the high position and strike the drum with full force. Return to high position.
- Down Stroke: Start from a high position and strike the drum with full force. End in the low position.
- Up Stroke: Start in the low position and strike the drum with light force. End in the high position.
- Tap: Start in the low position and strike the drum with light force. Return to the low position.

Moeller Technique is another method of stick motion named for its' inventor: Sanford Moeller. The technique involves a whipping motion that allows drummers to produce a stroke while moving the stick upwards towards the preparatory position. The technique and theory are outlined in The Moeller Book [118]. The technique was made popular by Jim Chapin and more recently by Jojo Mayer and Dave Weckl.

Leigh Howard Stevens in his book Method of Movement [161] provides a modified Gladstone technique for the marimba. The technique involves not only preparing the stick for the next stroke on the marimba but also preparing the arm as the mechanism for selecting the drum to strike. Stevens identifies that percussionists often move just in time rather than properly preparing. The prep stroke (or preparatory stroke) is a common mistake in drummers and percussionists where the stick often rests at a comfortable position (approximately 6 inches from the drum) and when it is time to perform a full stroke the player moves their stick up to proper position and then down. This is not only a waste of motion but may also lead to accuracy problems. Many people have adapted these ideas from the marimba to other percussion instruments that have multiple drums.

D.2.1 Rolls

The drum roll is the percussion method of creating long notes. Percussion instruments have an attack and a decay section to their envelope. A percussionist may strike an instrument loudly in order to increase the time that the decay section remains audible but otherwise has few options to play a note longer than the decay of the instrument.

The drum roll is performed by playing notes in rapid succession so that the decay portion of the instrument is not fully decayed when the next note occurs. The two most common methods for producing a drum roll are the double stroke roll (or open roll) and the buzz roll (or closed roll). The open roll consists of each hand performing exactly two strokes that are evenly spaced. The closed roll consists of each hand playing a multiple bounce stroke. The multiple bounce stroke is performed by the player dropping their stick onto the drum and allowing it to bounce. When pressure is exerted on the stroke it may be played at a faster rate.

The drum roll generally has a frequency between 15 and 30 hertz. Though the roll is very close to beginning of the perception of pitch it is generally not perceived as a pitch. If the player uses a drum with a short decay the effect is more audible as a series of notes and the perceptual result is very different.

D.2.2 Muting

Muting occurs when the duration of the drum is shortened by obstructing vibrations. This can be achieved by putting an object on the drum, applying pressure to the shell, or by installing a mechanism onto the drum. A special case of muting is discussed in Section D.2.3.

There are various mechanisms that can be attached to a drum in order to mute it; some are permanent while others allow for easy manipulation or quick removal. Drumset toms usually come equipped with a built-in dampener that is a simple arm with a piece of felt attached that pushes on the batter head. The pressure of the arm on the drumhead may be adjusted with a knob on the exterior of the drum. These dampeners are permanent modifications to the drum, though the mechanism may be removed the shell must pierced, altering the vibrational properties of the drum.

Percussionists often utilize non-permanent solutions to mute a drum such as: tape, adhesive gel, paper, or my favourite: credit cards. These items can be moved around the surface of the drum to vary the timbral effect of the muting. Light objects, such as credit cards, are useful since they are often on hand, are useful for purposes other than muting, and when the drum is hit hard enough they rattle much like the Cagean prepared piano.

Perhaps the most obvious way to mute a drum is by using one's hand. Hand drum players use this technique to greatly vary the timbral possibilities of the instrument.

D.2.3 Dead Stroke

The dead stroke is performed by striking the drum and then leaving the stick on the surface, and possibly applying pressure. The dead stroke functions with opposite effect as the roll: to shorten the decay of a percussion instrument.

Closely related to the dead stroke is dampening (or muting). Dampening is achieved when something is placed on the instrument during the stroke or after the stroke. This technique is common with cymbals since the motion of the instrument nullifies the effect of the dead stroke.

D.2.4 Grace Notes

Grace notes are performed by placing a small stroke before a primary stroke. When there is only one grace note the technique is usually referred to as "flams". The spacing of the two notes in time is widely variable. A flat flam is when both notes strike at the same time, while a fat flam has a greatly exaggerated space between the strokes. Dan Tomlinson outlines 23 variations for the basic flam [173]. Drumset players have made major advances in flam performance by splitting the parts of flams between different drums and different limbs.

Multiple grace notes may precede the primary note. When there are multiple grace notes the technique is often referred to as a “ruff”. While there are no limits to the number of strokes for a ruff, at some point the perceptual effect switches from a ruff to a roll.

D.3 Standard Hihat Techniques

The hihat is a standard drumset instrument. The high hat is made up of two cymbals on a stand that are moved together by pressing down on a foot pedal. The cymbals may also be struck with a stick while the cymbals are in various states of opening and closing. The hihat is usually placed on the weak side of the drumset (left side for right handed players and vice versa). The hihat is the only component that has a single playing point that allows for the control of duration.

D.3.1 Chick

The chick sound is achieved entirely through the pedal mechanism of the hihat. The player frees the mechanism, thereby opening the cymbals, and then depresses the pedal fully. This forces the cymbals together and the player sustains pressure to insure a short contact.

D.3.2 Splash

The splash sound is also achieved entirely through the pedal mechanism of the hihat. The opening motion is the same as the chick but instead of securing the cymbals together the pedal is the released again to allow the cymbals to ring. The nature of the ringing can be changed by varying the time the cymbals are in contact. Some players allow for the cymbals to rest slightly against each other, resulting in a sizzle like sound.

D.3.3 Bark

The bark sound is achieved by striking the top cymbal of the hihat while the cymbals are free. The mechanism is closed afterwards creating a barking sound. The length of the bark can be varied by altering the time between strike and pedal depress. This is one of the few drumset gestures where the duration of the event is contained within the gesture.

D.4 Standard Bass Drum Techniques

The bass drum is played with the foot via a pedal attached to the bass drum. There are many mechanisms available for the drumset

The bass drum beater can be left on the bass drum head to perform a damp stroke with the bass drum pedal. To perform the stroke the player simply depresses the pedal and then uses the weight of the leg to keep the beater in contact with the batter head of the bass drum. This stroke allows for the drummer to be able to perform a standard stroke and a dead stroke. Like the piano, because there is a mechanical intermediary, there are no other strokes possible with standard techniques.

There are three main methods of using the the bass drum pedals: heel up, heel down, and heel-toe. The heel up and heel down positions refer to the position of the heel relative to the baseplate of the bass drum pedal. Heel up players utilize their whole legs while heel down players utilize the lower leg while their heel rests on the pedal, allowing for greater stability on the drumset. The heel-toe technique was developed by Steve Gadd in order to achieve fast double strokes with the bass drum. The stroke is performed by bringing a raised foot down on the pedal so that only the heel strikes the pedal, activating the mechanism, then the toe snaps forward performing a second stroke.

D.5 Extended Techniques

There are numerous extended techniques available for drummers: some involve modifying the drum itself, others involve alternative uses of sticks, some with hands, some involve using unconventional objects to excite the drum. Drumset has been a largely individual pursuit with a great deal of oral tradition, thanks in part to recording technology being available during its entire history. Because there are so many techniques and so many individuals, there is not a large collection of notation that accompanies the practice. If drummers are to read charts, usually there is a legend provided at the beginning of the score. In 1998, after nearly 100 years of common practice on the drumset, Norman Weinberg published a book standardizing the drumset notation [190].

Michael Gould published an article defining and notating some of the common extended drumset gestures [55]. The article explores various methods of striking cymbals and various methods of using the rim of the drum in new ways. The problem with this material is that most software is unable to represent the new symbols for the new gestures.

D.5.1 Gimmicks

When Gladstone formalized the idea of a stroke being made of the upwards motion and the downwards motion many drummers started to explore how to create a strike during the upwards portion of a stroke. If this could be achieved then one could double their playing speed, assuming the method of achieving the second hit did not impede the speed of the overall motion.

Steve Gadd began extensively utilizing a technique where he would put his stick inside of a cowbell and move it up and down, thereby achieving the double rate motion. Billy Cobham began to use the same motion between a ride cymbal and a crash cymbal for loud bursts of noise.

Various bass drum pedal manufacturers have created pedals that allow for the double motion. The Dualist is a pedal that has an additional beater that is attached so that the negative phase of the footboard activates it, in addition to the normal beater. A recent venture had the drummer put a foot into a holder that moved a beater back and forth and electronic pads were put on either side of the beater allowing for double motion.

D.6 Temporal Techniques

D.6.1 Melodic Drumming

Jazz drummers evolved the drumset from a time keeping role, to a solo instrument, and eventually to be recognized as a melody instrument. Max Roach's style of drumming is often referred to as melodic and thus began the term melodic drumming. Drummers may play melodically by utilizing the instruments of the drums to create melodic elements. The toms of a drumset my be tuned to intervals that lend themselves to melodies. Cymbals my also have fewer overtones and create a pitched sound. Terry Bozzio is able to play an extended acoustic drumset with enough elements to create full compositions.

D.6.2 Textural Drumming

Textural drumming is a term that is used in association with the Alternate Mode Drumkat and associated products. Textural drumming is the use of electronic percussion instruments to create pitched textures. Most textural drumming consists of a solo performer playing a pre-composed piece using a sophisticated mapping of the controller that is highly context dependent.

D.6.3 Polyrhythms

Polyrhythms occur when two rhythms occur at the same time. In practice, the term polyrhythm refers to rhythms where the components are indivisible by an integer (3 in the space of 2 is a classic and simple example). Polyrhythms rarely persist for extended periods of time due to their complexity.

D.6.4 Free Playing

Drumset drummers are most often required to keep time for an ensemble using standard beat patterns. Humans tend to rebel and free music evolved as a method of reacting to strict time. Originally classified as Free Jazz, free music has evolved aurally and conceptually. The full realization of free playing in the drumming world was achieved by players such as Paul Motion and Jack DeJohnette, who slip into sections of free playing while other members of the ensemble are still playing the original metered music.

D.6.5 Cross rhythms

Cross rhythms occur when a rhythmic element is inserted into an ongoing rhythmic structure where the inserted material is of a differing grouping. The effect can also be referred to as a local polymetric sequence. The difference between a cross rhythm and polymetric music is that cross rhythms are usually an effect inserted into music, whereas polymetric music is often through composed with the meters remaining constant.

APPENDIX

E

Signal Processing and Machine Learning Techniques

The software designed for this project employs standard machine learning and signal processing techniques in order to achieve the end results of timbre recognition of the input instrument. This software uses standard feature extraction and classification methods. Features are extracted from signals recorded during the experimentation stage and the resultant feature matrix is input into the classifier for classification.

An exemplar-based classifier was used because the data had to be sorted into predetermined classes (as opposed to classes determined by the computer derived from classifiers, such as a self-organizing map). Neural networks were selected because of the availability in established software packages and their conformity to the exemplar-based learning preference.

Two different software packages were employed for this project: Weka and Marsyas . During the first phase of experimentation Matlab was used for both feature extraction and classification [111]. During the second phase of experimentation Marsyas was used for feature extraction and the output was written to the Weka file format [192] for classification with Weka.

After the raw data was collected from the experiments the data had to be prepared to be processed by the software. A preprocessing section was added to insure the integrity of the data in the system.

The software can be broken down into three main parts: Preprocessing, Feature Extraction, and Classification (see Figure E.1). Each one of these components will be explained in their own section in this chapter.

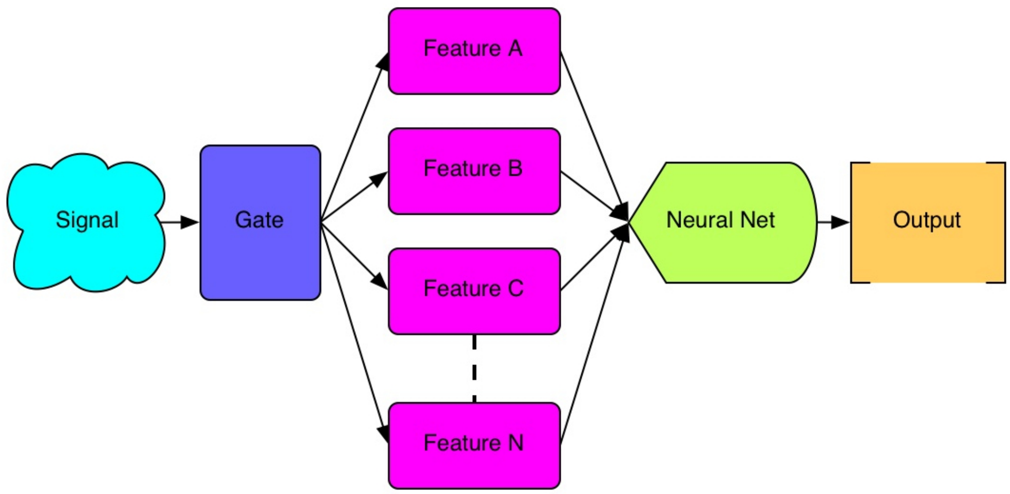


Figure E.1: Diagram of software layout.

E.1 Preprocessing

All of the samples were segmented automatically, via the surrogate sensor method, and then normalized for input into the system. The samples begin with a short silence and a gate algorithm is used so that all of the files are similar. The gate function has a threshold of -60dB. The signal is then examined for its peak amplitude and the zero-crossing previous to the index returned by the gate function.

The final size of the file that was kept for feature extraction was variable. Four different window sizes were used: 512, 1024, 2048, and “Attack Window.” All windows began at the index given by the gate function. The fixed width windows, 512, 1024, and 2048, consisted of the number of corresponding samples. The “Attack Window” is a window of variable size. The window ends when the sound reaches its peak amplitude.

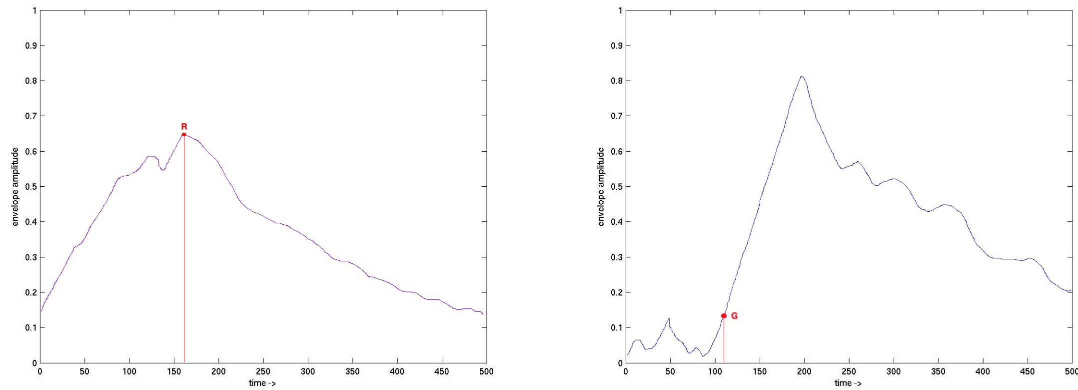


Figure E.2: Diagram of attack window and gate function.

E.2 Feature Extraction

Feature extraction was accomplished with Marsyas by creating small functions that performed some operation on the signal. A feature vector for each sample was produced for each signal that was analyzed. The feature vectors were then combined into a feature matrix to be fed to the classifier. The features used will be briefly described below.

E.2.1 Zero-Crossings

Zero-crossings is a simple technique that counts the number of times that the signal crosses zero. This can be useful in giving a rough estimate of the pitch, or in the case of the drum, a rough estimate of the major mode of vibration. The result was normalized for the variable width “Attack Window.”

E.2.2 Attack Time

Attack time is the size, in samples, of the onset section of the signal. The algorithm used looks for the first maximum value in the signal and returns its index. For the “Attack Window” the index was always the last value.

E.2.3 Temporal Centroid

Temporal Centroid is the center of gravity of the time domain representation of the signal as given by:

$$Tc = \frac{\sum_{i=1}^{n} |x_i|}{n} \tag{E.1}$$

x = the signal to be evaluated

n = the number of samples

i = the index of the current sample being computed

E.2.4 RMS

RMS, root mean squared, is a measurement of amplitude that returns the value as given by:

$$RMS = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \tag{E.2}$$

See Temporal Centroid for explanation of symbols.

E.2.5 Subband Analysis

Subband analysis techniques were used to determine energy in specific frequency bands. RMS was measured in four bands: 0–200Hz, 200–1000Hz, 1000–3000Hz, 3000–20,000Hz. A five-point IIR filter was used in order to separate the different bands. This measurement gives another rough estimate of which bands are being excited.

E.2.6 Spectral Centroid

Spectral Centroid returns the center of gravity of the spectrum as given by:

$$SC = \frac{\sum_{k=0}^{N/2} k * |X(k)|}{\sum_{k=0}^{N/2} |X(k)|} \tag{E.3}$$

$X(k)$ = the spectrum of the signal given by an FFT calculation

N = number of analysis frames (determined by FFT size)

The FFT size used for all spectral features was 256 samples. The algorithms were run on each FFT block and then the results were averaged into one value for all spectral features with multiple windows.

E.2.7 Spectral Flux

Spectral Flux measures the amount of local change over time in the frequency domain. It is defined by squaring the difference between normalized magnitudes in the frequency domain of frame t and $t-1$. If $N_t[n]$ and $N_t[n-1]$ are defined by the normalized magnitude of frame t and $t-1$, then the spectral flux F_t is given by:

$$F_t = \sum_{n=1}^{N} (N_t[n] - N_{t-1}[n])^2 \tag{E.4}$$

It should be noted that magnitudes are normalized by dividing each value in every frame by the RMS value of that frame [179]. F_t is calculated for each frame and then averaged over time in order to yield one value for spectral flux.

E.2.8 Spectral Rolloff

Spectral rolloff is another feature that describes the spectral shape [179]. It is defined as the frequency R_t below which 85% of the magnitude of the spectrum is concentrated. If $M_t[n]$ is the magnitude of the spectrum then the spectral rolloff is given by:

$$\sum_{n=1}^{R_t}(M_t[n]) = .85 * \sum_{n=1}^{N}(M_t[n]) \tag{E.5}$$

E.2.9 Spectral Skewness

Spectral Skewness is a third-order moment that returns the skewness of the spectrum as given by:

$$Sk = \frac{\sum (x-u)^3}{\sigma^3} \tag{E.6}$$

$x =$ the magnitude of the spectrum of the signal

$u =$ the mean of the signal

$\sigma =$ the spectrum distribution standard deviation

E.2.10 Spectral Kurtosis

Spectral Kurtosis is a fourth-order moment that examines the number and strength of outliers in the spectrum. A spectrum with normal distribution will have a spectral kurtosis of 3. The function in this experiment conforms to the convention where three is subtracted from the kurtosis so that a spectrum with normal distribution will have a spectral kurtosis of 0.

$$K = \left(\frac{\sum (x-u)^4}{\sigma^4}\right) - 3 \tag{E.7}$$

See Spectral Skewness for explanation of symbols.

E.2.11 Linear Predictive Coding

Linear Predictive Coding (LPC) produces an estimation $\overline{x}(n)$ for a sample value $x(n)$ as a linear combination of previous sample values. This is shown by:

$$\overline{y}[n] = \sum_{k=1}^{L} a_k x(n-k) \tag{E.8}$$

where a_k are the coefficients of the predictor. The z-transform of this equation is given by:

$$\overline{X}(z) = \sum_{k=1}^{L} a_k z^{-k} X(z) \tag{E.9}$$

Thus, using LPC coefficients a sound can be represented in terms of coefficients to an IIR filter. Ten LPC coefficients were used for the feature data.

E.2.12 Mel-Frequency Cepstrum Coefficients

Mel-frequency Cepstrum Coefficients (MFCC) are a product of two distinct stages of operations. First, the cepstrum of the signal is calculated, which is given by taking the log of the magnitude spectrum. This effectively smooths the spectral content of the signal. Second, the spectrum is divided into thirteen bands based on the mel scale, which is a scale based on human perception of pitch [101].

This feature returned a set of coefficients for each FFT frame of the signal that was analyzed. A 256-point FFT size was used providing thirteen coefficients for each FFT frame. The values for each frame were averaged together to provide only thirteen values for the feature data.

E.3 Classification

Many classification strategies are available to classify data. This study has chosen the artificial neural network as the main classifier. Two other classifiers were used in order to evaluate the performance of the neural network: Support Vector Machines, k-Nearest Neighbour.

Throughout this thesis the term Neural Network will be employed. There are many forms of Neural Networks that can be used. The following section is intended to provide details of the type of network employed in this study.

E.3.1 Neural Networks

The human brain is comprised of a vast network of neurons. Neurons contain a small mass of tissue, the nucleus, with two types of connecting bodies: the axon and dendrites. Dendrites bring electrical pulses into the nucleus. Based upon the strength and the source of the impulse the neuron may or may not fire an impulse out of its axon. "Learning" is believed to occur when the variables needed to fire the neuron are changed, thus a different input will be needed in order to get the same response.

E.3.2 The Perceptron

The perceptron model was first proposed by Rosenblatt (1958). The perceptron model is based on the idea of a single neuron of the human brain. The inputs to the perceptron are an n-dimensional feature space where $0 \leq x_i \leq 1$ and the input vector $\underline{x}$ is the set $\underline{x} = (x_1, x_2, ..., x_n)$. The perceptron collects inputs, $\sum_{i=1}^{n} w_i x_i$, as a weighted sum (w_i being the weights) as demonstrated below. The weighted input to the transfer function is the output $u(\underline{x})$, as demonstrated in the following equation.

$$u(\underline{x}) = w_O + \sum_{i=1}^{n} w_i x_i \tag{E.10}$$

The output $y(\underline{x})$ is determined by the threshold function below, which demonstrates a binary output. The formula demonstrates that if the input is greater than 0 then the perceptron fires. The learning process changes the weight coefficients so that only under certain circumstances will the perceptron fire.

$$y(\underline{x}) = \begin{cases} 1 & u(\underline{x}) \geq 0 \\ 0 & u(\underline{x}) > 0 \end{cases} \tag{E.11}$$

The weights $\{\underline{w}\}$ determine how much a particular feature affects the decision of the perceptron to fire. The activation threshold is given by W_O. The following diagram graphically demonstrates these formulae:

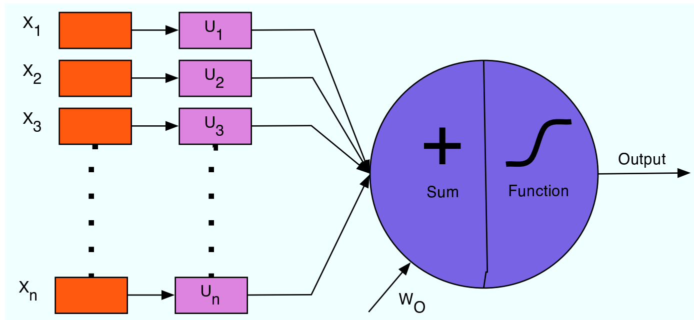


Figure E.3: Diagram of a perceptron.

E.3.3 Multi-Layer Perceptron

The multi-layer perceptron is simply a collection of perceptrons organized into layers. One major feature of this type of organization is that not all of the perceptrons will receive the inputs of the feature vector. There are three types of layers: input layer, output layer, and hidden layer. The types of layers differ based on how their inputs and outputs are connected.

Input Layer The inputs are received from the feature vector and the outputs are connected to the inputs of the hidden layer.

Hidden Layer The inputs are received from the outputs of the input layer and the outputs are connected to the inputs of the output layer.

Output Layer The inputs are received from the outputs of the hidden layer and the outputs are the output of the system.

The following diagram graphically demonstrates the connection structure of a neural network.

There are two different ways to connect the different neurons together: feedforward and feedback. A feedforward connection structure is when the output of all neurons are connected only to the inputs of the following layer, or the output in the case of the output neuron of the structure. A feedback connection structure is when

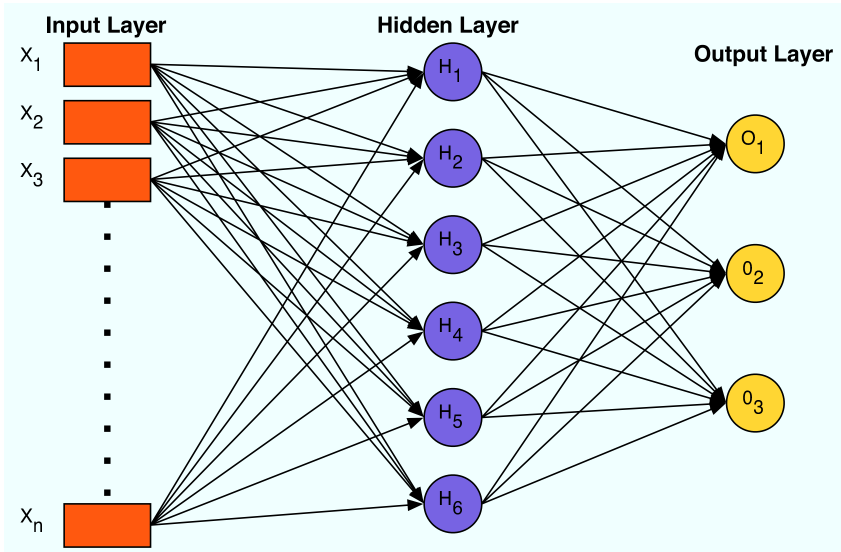


Figure E.4: Diagram of an artificial Neural Network.

the output of a neuron is connected to either its following layer but it may also have a connection to a neuron in its own layer or a previous layer.

There are many different ways to implement the learning process in the perceptron. The backpropogation method is a supervised learning strategy where the output of the perceptron is compared with the desired output and the weights are then adjusted. The learning rate determines how much the weights are adjusted. Each time the process is repeated it is called an epoch. The learning rate and number of epochs are important factors in the accuracy of performance when constructing a neural network.

E.3.4 k-Nearest Neighbor

The k-Nearest Neighbor algorithm, also referred to as kNN or k*, is a simple classifier that puts training data into a feature space. In the simplest case the training phase is simply the storage of the training examples within the n-dimensional feature space. In classification mode the query is placed in the feature space and the k nearest neighbors in the space are computed. For kNN to work properly k must always be odd so that there is always a majority of a given class for the training data points that are retrieved and the class with the greatest representation in this set is returned as the classification for the query.

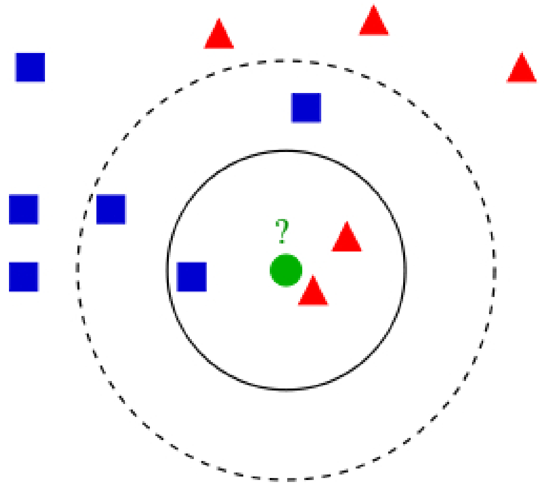


Figure E.5: Example of a kNN classifier showing the query in green and the area of k being equal to 3.

kNN is a very simple algorithm but does allow for optimizations for given contexts. For the simple case, computation for the training phase is very simple. If the application permits then during training phase a sorting algorithm can be applied to the training set to allow for faster retrieval. The distance is often computed as Euclidean, but other metrics such as Hamming distance or Manhattan distance may also be applied.

E.3.5 Support Vector Machines

The Support Vector Machine algorithm (or SVM) is similar to the kNN algorithm in that it collects instances of training data into an n-dimensional feature space. When the training has been collected the algorithm attempts to place a divisor, in the form of a hyperplane in the feature space, between the classes of the training set. For classification all that has to be done is to compute which side of the divisor the new instance sits and then declare the query as belonging to the class that occupies that side of the divisor.

SVM can allow the user to determine the quality of their feature set by examining the size of the margins - distance from the closest instance of a given class to the divisor. The larger the margins, the more clear the feature space and therefore the better the feature set. Optimizations to the SVM algorithm allow for the divisor to take on other shapes than a line, including the concept of a "Soft Margin" for cases where the divisor is unable to clearly separate between classes.

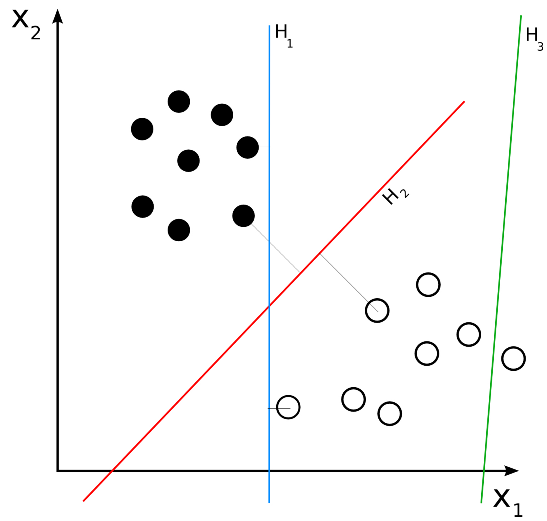


Figure E.6: Example of a SVM classifier showing the divisor and the margins.

E.3.6 Bayesian Classifier

A Bayesian classifier (or Naive Bayes Classifier) is based upon Bayes' theorem where the features have a strong independence. The classifier is derived Bayes' Rule as expressed below:

$$p(A|B) = \frac{p(A)p(B|A)}{p(B)} \tag{E.12}$$

The classifier determines the probability of a feature value being represented in a class and proceeds to make an estimation model. Input is measured against a series of likelihood ratios to determine the most probable class. In theory the assumption of the classifier that the variables are completely independent is problematic because because most features collected are related to each other in some manner. In practice the classifier is able to perform well in noisy environments because of this correlation. Bayesian classification for the purpose of filtering is common in email applications for determining unwanted messages.

E.3.7 ZeroR Classifier

A ZeroR classifier is a ground test classifier that examines the training set and selects the label with the most number of instances in the training set. All incoming unknown examples will be labelled according to whichever label was most prevalent in training.