VICTORIA UNIVERSITY OF WELLINGTON

DOCTORAL THESIS

**Closing the Loop:
Implementing Real-time Audio Feedback
Systems in Musical Robots**

Author:

Jason LONG

Supervisors:

Dr. Jim W. MURPHY

Dr. Dale A. CARNEGIE

Dr. Ajay KAPUR

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in the

Sonic Engineering Lab for Creative Technology
New Zealand School of Music

2018

“My focus has always been to combine the visual, the acoustical, and the kinetic parts - to aid understanding by making the circle complete.”

Trimpin

Victoria University of Wellington

Abstract

Faculty of Humanities and Social Sciences

New Zealand School of Music

Doctor of Philosophy

Closing the Loop: Implementing Real-time Audio Feedback Systems in Musical Robots

by Jason LONG

A closed-loop control system is any configuration that feeds information about its output back into the control stream. These types of systems have been in use for hundreds of years in various engineering related disciplines to carry out operations such as keeping rooms at the correct temperature, implementing cruise control in cars, and precisely positioning industrial machinery. When a musician performs a piece, a type of biological closed loop is invoked in which the player continuously listens to the sound of their instrument, and adjusts their actions in order to ensure their performance is as desired.

However, most musical robots do not possess this ability, instead relying on open-loop systems without feedback. This results in the need for much manual intervention from the operators of these robots, unintuitive control interfaces for composing and performing music with them, and tuning, timing, dynamics and other issues occurring during performances.

This thesis investigates applying closed-loop audio feedback techniques to the creation of musical robots to equip them with new expressive capabilities, interactive applications, musical accuracy, and greater autonomy. In order to realise these objectives, following an investigation of the history of musical automata and musical robotic control systems, several new robotic musical instruments are developed based on the principals of utilising embedded musical information retrieval techniques to allow the instruments to continuously 'listen' to themselves while they play.

The mechanical and electronic systems and firmware of a closed-loop glockenspiel, a modular unpitched percussion control system, and a robotic chordophone control system are described in detail, utilising new software and hardware created to be accessible to electronic artists. The novel capabilities of the instruments are demonstrated both through quantitative evaluations of the performance of their subsystems, and through composing original musical works specifically for the instruments. This paradigm shift in musical robotic construction paves the way for a new class of robots that are intuitive to use, highly accurate and reliable, and possess a unique level of musical expressiveness.

Acknowledgements

Before the thesis begins, I would like to acknowledge all of the help and support of my friends, family, and colleagues. Without this community of generous, hard-working, and like-minded individuals this research would not have been possible.

Firstly, I'd like to express great appreciation to my supervisory team, Jim Murphy, Dale Carnegie and Ajay Kapur. You have all helped craft the shape of this work in various ways, and I have been honoured to be able to work alongside you for the past three years. I am especially grateful to Jim, for your tireless advocacy and guidance throughout, Dale for your academic rigour and engineering brilliance, and Ajay for your broad perspective and inspiration.

I'd also like to express my thankfulness to the staff of the School of Engineering and Computer Science, the New Zealand School of Music and the wider Victoria University of Wellington. Thank you very much to engineering technical staff Tim Exley, Jason Edwards and Brandon Pothoven for all of your advice and hard work in helping me realise this project. Another big thank you to Roy Carr for all of your assistance in making our concerts the best they could be, and to Fiona Steedman for being the centre cog that makes NZSM keep turning. I'd also like to express great thanks to NZSM teaching staff Michael Norris and Dugal McKinnon for always being there to provide sage advice and support my work. More thanks go out to the staff at Image Services for making my pieces and concerts look beautiful, and to the Cleaning Services staff for your tireless efforts in making the university campus an immaculate and comfortable place to do research.

I'd like to convey my sincerest gratitude and appreciation to my SELECT colleagues and friends, Diana Siwiak, Jon He, Bridget Johnson, Blake Johnston, Mo Zareei, and Kameron Christopher for all of your support and companionship throughout these years. I hope that no matter where in the world we find ourselves, we will keep in touch and find opportunities to collaborate and make music together in the future.

I'd like to thank my family for their undying love and support through thick and thin. I'd be nowhere without the tools and lessons you have provided me from day one. Last but not least, a huge debt of gratitude is due to my wonderful wife Alissa. You are the shining light that has guided me every day throughout this journey.

Contents

Abstract	v
Acknowledgements	vii
Contents	ix
List of Figures	xv
List of Tables	xix
List of Abbreviations	xxi
1 Introduction	1
1.1 Overview	1
1.2 Motivations	2
1.3 Objectives	5
1.4 Document Structure	8
1.5 Summary of Contributions	11
2 Related Works	13
2.1 Overview	13
2.2 A History of Musical Automata	13
2.2.1 Pre-18 th Century	14
2.2.2 The 19 th Century and the Golden Age of Mechanical Musical Instruments	17
2.2.3 The Dawn of Musical Robotics	22
2.2.4 The State of the Field	26
2.3 Defining the Terminology: Musical Robotics	27
2.4 Related Areas of Research	29
2.4.1 Sensors in Musical Instruments	30
2.4.2 Musical Information Retrieval	32
2.4.3 General Robotic Control Systems	33
2.4.4 Pre-Performance Calibration of Musical Robots	35

2.5	Defining Audio-based Closed-loop Control	37
2.6	Contemporary Robots with Audio-based Closed-loop Elements	39
2.7	Opportunities for Improvement	40
2.8	Research Questions	40
3	The Closed-Loop Robotic Glockenspiel	43
3.1	Overview	43
3.2	The Robotic Glockenspiel	44
3.3	Design Considerations	47
3.3.1	Audio Transduction	48
3.3.2	Analogue Signal Conditioning	49
3.3.3	Digital Signal Processing	50
3.4	The Closed-Loop Upgrades	51
3.4.1	The Panel	51
3.4.2	The Mechanics	53
3.4.3	The Electronics	55
	Power	55
	Digital Control	57
	Audio Circuitry	58
3.4.4	The Firmware	63
	MIDI I/O	63
	Analogue I/O	65
	Onset Detection	66
	Amplitude Detection	68
	Continuous Calibration	70
3.4.5	Other Features	73
3.5	Evaluation of Functionality	73
3.5.1	Dynamics	74
	Open-loop Response	74
	Closed-loop Response	78
3.5.2	Latency	82
	Open-loop Response	82
	Closed-loop Response	86
3.6	Practical Issues Encountered	86
3.6.1	Noise	88
3.6.2	Accuracy	88
3.6.3	Programming	89
3.7	Creative Outputs	90

3.7.1	Composition: Broken Mirror	90
3.7.2	Installation: AutoReich	94
3.8	Conclusions	96
4	A New Robotic Drum Kit: Mechanics	99
4.1	Overview	99
4.2	Evaluating Existing Percussion Mechanisms	101
4.2.1	Percussion Mechanism Types	101
	Solenoids	101
	Pneumatics	102
	Servos	103
4.2.2	Test Mechanism Construction	103
	Rotary Solenoid Strikers	103
	Linear Solenoid-based Strikers	104
	Servo Strikers	105
	Pneumatic Strikers	106
4.2.3	The Evaluation Methodology	107
4.2.4	The Evaluations	108
	Latency	109
	Maximum Loudness	111
	Dynamic Consistency	111
	Maximum Repetition Rate	112
4.2.5	Results	113
4.3	Striker Construction	115
4.3.1	Modular Drum Strikers	115
	Snare	117
	Rack Toms	118
	Floor Tom	120
	Kick Drum	121
	Ride Cymbal	124
	Crash Cymbal	125
	Hi-hat	127
4.4	Hi-hat Movement System	127
4.4.1	Hi-hat Movement Mechanics	129
4.4.2	Positioning System	131
4.4.3	Hi-hat Movement Circuitry	133
4.4.4	Signal Conditioning	135
4.4.5	The Power Supply	136

4.5	Summary	137
5	A New Robotic Drum Kit: Control	139
5.1	Overview	139
5.1.1	Design Goals	140
5.2	Robotic Percussion Array Control Unit	142
5.3	Actuator Control System	146
5.3.1	Design Process	146
5.3.2	The Circuit Board	148
5.3.3	Internal Communication	150
5.3.4	Communication with the Outside World	150
5.3.5	The FPGA	151
5.3.6	The Programmer	152
5.3.7	Power	153
5.3.8	Evaluation	154
5.4	Preamplifier Modules	155
5.5	Multi-Channel Embedded MIR Board	159
5.5.1	Design Considerations	160
5.5.2	Part Selection	161
5.5.3	The System's Hardware	164
	Input Stage	165
	Central Processor	166
	Output Stage	166
5.6	Digital Audio Output Board	167
5.7	Power Supplies	172
5.8	MIR Firmware	177
5.8.1	Bootstrapping the DSP	177
5.8.2	Configuring the Peripherals	178
5.8.3	Onset Detection	180
5.8.4	Latency Detection	182
5.8.5	Amplitude Detection	183
5.8.6	Envelope Length Detection	184
5.8.7	Additional Potential Capabilities	187
5.9	Evaluation of Functionality	188
5.9.1	Methodology	189
5.9.2	Dynamics	189
	Open-loop Response	190
	Closed-loop Response	191

5.9.3	Latency	191
	Open-loop Response	191
	Closed-loop Response	192
5.9.4	Hi-hat Height	192
	Open-loop Response	194
	Closed-loop Response	195
5.10	Creative Outputs	196
5.10.1	The Impossible Drum Solo Part One	197
5.10.2	The Impossible Drum Solo Part Two	200
5.11	Conclusions	202
6	Closed-Loop Chordophones	205
6.1	Overview	205
6.2	ServoSlide	207
6.2.1	Picking Mechanism	208
6.2.2	Fretting Mechanism	209
6.2.3	Existing Closed Loops	209
6.3	The DSP Board	210
6.3.1	Board Design	211
6.3.2	Integration with the Robot	213
6.4	Firmware	214
6.4.1	Bootig and Configuration	214
6.4.2	Audio Analysis	215
6.4.3	Envelope Detection	217
6.4.4	Conditioning the Signal	217
6.4.5	Pitch Detection	218
6.4.6	Detecting Other Characteristics	219
6.5	Evaluations	219
	Open-loop Response	220
	Closed-loop Response	221
6.6	Conclusions	224
7	Conclusions	225
7.1	Key Contributions	227
	A Survey of Closed-loop Control in Musical Robotics	227
	The Closed-Loop Robotic Glockenspiel	227
	A Methodology for Evaluating Robotic Percussion Mechanisms	227
	A New Robotic Drum Kit	228

	An Expressive Hi-hat Movement System	228
	A Closed-Loop Modular Robotic Percussion Control Sys- tem	228
	A Highly Accurate FPGA-based Musical Robot Control Board	228
	A New Modular Preamp Board	229
	A 32-Channel Input and Output Embedded DSP Board	229
	A 16-Channel TDM to ADAT Conversion Module	229
	Novel Algorithms for Embedded MIR for Musical Robots	229
	A DSP board for Space-Constrained Closed-Loop Mu- sical Robotics	230
	New Musical Works for Closed-loop Robotic Instruments	230
7.2	Future Work	230
7.3	Research Outputs	231
7.3.1	Awards	232
7.3.2	Peer-Reviewed Publications	232
7.3.3	Musical Performances and Installations	234
7.3.4	Media, Press	235
	Radio Coverage	235
	Internet Coverage	235
7.4	Concluding Remarks	236
A	Track Descriptions	239
A.1	Broken Mirror	239
A.2	The Impossible Drum Solo: Part Two	242
	Bibliography	247

List of Figures

1.1	Robot Ensemble Performance at Tokyo University of the Arts	2
1.2	Open-Loop System	5
1.3	Closed-Loop Diagram	6
1.4	Thesis Structure	9
2.1	Timeline of Early Historical Automatic Instruments	14
2.2	Diagram of a Carillon	16
2.3	Golden Age	17
2.4	Panharmonicon	18
2.5	Early and Modern Piano Rolls	19
2.6	Disc Music Box and Modern Turntable	21
2.7	The Bellenorgel	23
2.8	Timeline of Musical Robotics Participation	25
2.9	The Hyperflute	31
2.10	Control System Architectures	33
3.1	The Robotic Glockenspiel	45
3.2	The Robotic Glockenspiel Calibration Interface	46
3.3	The Closed-Loop Robotic Glockenspiel's Back Panel	52
3.4	The Instrument's Pickup and Striking Mechanisms	53
3.5	The Power Supply	56
3.6	Schematic of the Solenoid Power Supply	57
3.7	Schematic of the Audio Power Supply	58
3.8	The MOSFET Board	59
3.9	The Audio Conditioning Board	60
3.10	Schematic of an Audio Conditioning Channel	61
3.11	Internals of the Closed-Loop Robotic Glockenspiel	62
3.12	The Instrument's Serial Communications	64
3.13	Onset and Amplitude Detection	67
3.14	The Audio Rate Analysis Routine	69
3.15	The Dynamics and Latency Calibration Routines	71
3.16	Loudness vs Pulse Widths Without Calibration - First Trial	75

3.17	Uninterrupted vs Interrupted Strikes	77
3.18	Loudness vs Pulse Widths Without Calibration - Second Trial	79
3.19	Loudness vs Pulse Widths Without Calibration - Third Trial . .	80
3.20	Loudness vs Velocity with Calibration	81
3.21	Latency vs Pulse Width Without Calibration - First Trial	83
3.22	Latency vs Pulse Width Without Calibration - Second Trial . .	84
3.23	Latency vs Pulse Width Without Calibration - Third Trial . . .	85
3.24	Latency vs Velocity with Calibration and Compensation . . .	87
3.25	Extract from the Piano Roll of Broken Mirror	93
3.26	The Software Backend of AutoReich	95
3.27	The Software Interface of AutoReich	96
4.1	Types of Solenoids Commonly Used for Robotic Percussion . .	102
4.2	The Rotary Solenoid-Based Percussion Mechanisms	104
4.3	The Linear Solenoid-Based Percussion Mechanisms	105
4.4	The Servo-Based Percussion Mechanisms	106
4.5	The Pneumatic Percussion Mechanisms	107
4.6	The Various Stages of the System's Latency	108
4.7	The Latencies of the Evaluated Mechanisms	110
4.8	Maximum Loudness Value of Each Striker	112
4.9	Maximum Repetition Rates of Strikers	113
4.10	New Robotic Striker Design	116
4.11	New Robotic Snare Drum	117
4.12	Robotic Tom Drum	119
4.13	Robotic Floor Tom Drum	120
4.14	Robotic Kick Drum Mechanism	122
4.15	The Kick Drum's Internal Mic'ing System	123
4.16	Robotic Ride Cymbal	125
4.17	Robotic Crash Cymbal	126
4.18	Robotic Hi Hat	128
4.19	The Hi Hat Movement System	130
4.20	Internals of the Hi Hat Movement System	132
4.21	Architecture of the Hi Hat Movement System	133
4.22	The Hi Hat Movement Circuitry	134
4.23	The Signal Conditioning Circuits	135
4.24	The Hi Hat Power Supply	136
4.25	Hi Hat Power Supply Schematics	137
5.1	The Robotic Percussion Array System Architecture	141

5.2	The Robotic Percussion Array Control Unit	143
5.3	Internals of the Control Unit	145
5.4	The FPGA-based Musical Robot Controller Board (Render) . .	149
5.5	The FPGA-based Musical Robot Controller Board (Photo) . . .	149
5.6	The Robot Control Board's Signal Flow	150
5.7	The FPGA Programmer Board	153
5.8	The Accuracy of the Output Pulses	154
5.9	The Custom Preamp Module	156
5.10	The Preamp Module's Schematic	158
5.11	Signal Flow of the Embedded MIR System	161
5.12	Custom Blackfin and SHARC Development Boards	163
5.13	Front and Back of the DSP PCB	164
5.14	Annotated Photograph of the System's Circuit Board	165
5.15	Custom ADAT I/O Board (First Iteration)	168
5.16	The ADAT Interfacing Evaluation Boards	170
5.17	The Control Unit's Power Supply Structure	173
5.18	The Main Power Supply Board	174
5.19	5V Power Supply Board	176
5.20	The DSP Board's Audio-Rate Analysis Routine	181
5.21	The Solenoid Control Board's Calibration Routines	185
5.22	Envelope Length Detection	187
5.23	Snare Loudness vs Pulse Width (Open-Loop)	190
5.24	Snare Loudness vs Velocity Values (Closed-Loop)	191
5.25	Snare Latency vs Pulse Width (Open-Loop)	192
5.26	Snare Latency vs Velocity Values (Closed-Loop)	193
5.27	Hi Hat Envelope Lengths vs Power Levels (Open-Loop)	194
5.28	Hi Hat Envelope Lengths vs CC Values (Closed-Loop)	195
5.29	The Robotic Drum Kit at TEDxWellington	196
5.30	First Extract from the Impossible Drum Solo Part One	198
5.31	Second Extract from the Impossible Drum Solo Part One	199
5.32	Third Extract from the Impossible Drum Solo Part One	200
6.1	The ServoSlide Robotic Chordophone	208
6.2	The Embedded DSP Board	210
6.3	Design of the DSP PCB	212
6.4	Pitch Extraction and Calibration Routines	216
6.5	Filtering and Zero-Crossing Extraction	217
6.6	Pitch Response of a Traditional Robotic Chordophone System	221
6.7	Pitch Response of the Closed-Loop System after Manual Tuning	221

6.8	Pitch Response of the Closed-Loop System after Detuning . .	222
6.9	Difference Between Manually Tuned and Detuned Responses	222
6.10	Spectrogram of the First Note in a Detuned Sequence	223
A.1	Screenshot of the <i>Broken Mirror</i> DAW Project	240
A.2	Screenshot of the <i>Impossible Drum Solo Part Two</i> DAW Project .	242

List of Tables

4.1	Dynamic Consistency of Tested Percussion Mechanisms	112

List of Abbreviations

PWM	P ulse W idth M odulation
MIR	M usical I nformation R etrieval
DSP	D igital S ignal P rocessing (or P rocessor)
FFT	F ast F ourier T ransform
RMS	R oot M ean S quare
PID	P roportional I ntegral D erivative
MIDI	M usical I nstrument D igital I nterface
OSC	O pen S ound C ontrol
MEMS	M icro E lectro- M echanical S ystems
ADC	A nalogue to D igital C onverter
UART	U niversal A synchronous R eceiver- T ransmitter
EEPROM	E lectrically E rasable P rogrammable R ead- O nly M emory
GPIO	G eneral P urpose I nterface O utput
DMA	D irect M emory A ccess
I²C	I nter- I ntegrated C ircuit
MOSFET	M etal O xide S emiconductor F ield E ffect T ransistor
DAC	D igital to A nalogue C onverter
SRAM	S tatic R andom A ccess M emory
DAW	D igital A udio W orkstation
FPGA	F ield- P rogrammable G ate A rray
CNC	C omputer N umeric C ontrol
EMI	E lectro- M agnetic I nterference
IC	I ntegrated C ircuit
ADAT	A lesis D igital A udio T ape
TDM	T ime D omain M ultiplexing
SPI	S erial P eripheral I nterface
IDE	I ntegrated D evelopment E nvironment
HCMOS	H igh-density C omplementary M etal O xide S emiconductor
JTAG	J oint T est A ction G roup
FIR	F inite I mpulse R esponse
SHARC	S uper H arvard A rchitecture

MAC	M ultiply A Ccumulate
TWI	T wo W ire I nterface
SPORT	S erial P ort
I²S	I nter- I C S ound
ROM	R ead- O nly M emory
PCB	P rinted C ircuit B oard
MADI	M ultichannel A udio D igital I nterface
AVB	A udio V ideo B ridging
S/PDIF	S ony / P hilips D igital I nterface
NRZI	N on- R eturn to Z ero I nverted
PTC	P ositive T emperature C oefficient
PiP	P lug- i n P ower
PLL	P hase- L ocked L oop
FIFO	F irst I n F irst O ut
CC	C ontrol C hange

Dedicated to my family, friends and teachers...

Chapter 1

Introduction

1.1 Overview

For over a thousand years, people have been building devices that automatically create sound and music. The inspiration to produce such apparatuses often stems from the desire to realise sonic creations that are difficult, impossible, or infeasible for humans to create manually. This desire is reflected in many pieces of modern-day music-making equipment, and is particularly applicable to the field of musical robotics, the primary concern of the research presented in this document.

Robotic musical instruments allow composers and artists to make music and sound art in an entirely different way to conventional methods and provide expressive capabilities that neither humans performing with traditional instruments nor loudspeakers are able to provide. By combining real-world acoustic sonorous objects with precise computer control, musical robots are able to achieve new musical feats that humans cannot, for example performing extraordinarily fast, complex, or long music. They also open up many areas of musical expression that have heretofore belonged exclusively to the realm of electronic music composition, to the use of real-world acoustic sound-sources. Examples of such areas include real-time algorithmic techniques, works which utilise interactive elements and direct audience participation, and gallery and installation works.

Much progress has been made in the field of musical robotic research since the earliest examples from the 1970s, and these developments have allowed practitioners to create very large-scale public installations such as Trimpin's *If VI Were IV* and *Sheng High* [27, 31], sophisticated orchestras of machines

such as those created by Godfried Willem Raes [129], and robots that outperform their human counterparts in a variety of aspects [164]. Each step along this path of progress brings the field greater public awareness and acceptance, and enables musicians and artists to express themselves in ever more compelling ways.

However, it can still be considered that musical robotics, as a musical and research discipline, is in its infancy. Access to machine orchestras is still enjoyed by only a very small number of composers and music schools, and only a few dedicated musicians in the sphere of popular music, such as Björk [8], Pat Metheny [90], Aphex Twin [1], and Squarepusher [153], have been able to make use of them. Just as computer music began as a topic of research in the laboratories of universities and has grown to be a fixture in music departments around the world and accessible to musicians and artists of all levels of technical expertise, musical robotics has the potential to have a much greater presence in our musical lives. By addressing some of the accessibility, reliability, and usability issues of these instruments, this thesis advances the field towards this ideal.

1.2 Motivations

FIGURE 1.1: The Robot Ensemble live in concert at Tokyo University of the Arts in 2014.

In order to realise my own compositional goals, from 2010 to 2014 I designed and built a collection of musical robots, including gamelan players, an array of monophonic and polyphonic percussion instruments, a robotic singing bowl, and a robotically plucked and fretted string instrument called the Robotic Taishogoto [75]. The design, firmware, and construction of these instruments are outlined in the Masters thesis “The Development of an Ensemble of Automatic Musical Instruments and their Possibilities” [74]. Figure 1.1 shows the majority of the ensemble set up in concert in Tokyo in early 2014.

Over the course of designing, building, and exploring the sonic capabilities of this set of instruments, various limitations were observed and a greater understanding of the current issues that prevent wider-scale adoption of musical robots was obtained. Some of the reasons for this un-tapped potential, such as fabrication, cost, distribution, and barrier to entry, are being addressed by 3D printer and laser cutter makers such as Full Spectrum Laser [34], online resources such as Thingiverse [88], and development platforms such as Arduino [3] respectively. However, there remain several ongoing issues that even those with access to these machines frequently encounter, hindering the ability to fully utilise the instruments and maximise their potential:

1. They require frequent calibration and maintenance. While all musical instruments and sound systems take time to set up, sound check, and tune, musical robots generally take much longer than traditional electronic instruments. Most musical robots, like many other one-of-a-kind electronic instruments, are also prone to failure or inaccuracies and require constant human supervision [156]. Nobumichi Tosa of Meiwa Denki mentions in an interview:

The instruments sometimes break, which becomes part of the show. It’s kind of like F1 racing. We try hard to keep them from breaking, but with more than 30 instruments, some kind of trouble always arises. [41]

Shawn Trail makes the following complaints about conducting performances with Guitarbot [143], one of the musical robots included in Pat Metheny’s Orchestrion tour:

The Guitarbot is manually tuned and has no self-sensing technology. It typically goes out of tune during a concert, is cumbersome to transport and difficult to maintenance [sic]. [159]

This is due to the array of moving parts in the instruments, with very small discrepancies in distances and forces resulting in clearly audible errors in the resulting sound. Since human performers are not necessarily involved in a live performance, the robots may fall out of tune or calibration during a performance, with no indication provided to the operators.

2. They require specialised technical knowledge to build. The design and construction of musical robots requires embedded software engineering, electrical engineering, mechanical engineering, audio engineering and electronic musical skills, resulting in a high barrier to entry for potential musical roboticists. As a result, it is common for those with strong musical skills to create robots that lack robust engineering rigor, and those with engineering training to create instruments of little use to musicians. Though it may be unrealistic to expect these machines to be created by those without any of the above skills, measures that lower the barriers to entry will be helpful for the field.
3. They require specialised technical knowledge to play. The majority of musical robots are built with overly complex control systems that make it difficult for anyone other than the creator to naturally make music with them. Very few groups (with a notable exception being Logos [128]) publish guides for potential composers, and this leads to a lack of adoption by musicians without a background in the area. In ideal circumstances, these machines would be as accessible to musicians as current mainstream electronic musical instruments, requiring predominantly musical knowledge to operate. However, current musical robots often require knowledge of actuator positions, pulse-width modulation (PWM) duty cycles and other such electro-mechanical characteristics in order to successfully operate.
4. They lack desired expressive capabilities. In many situations, while musical robots may excel with regards to specific performance criteria, they may not be able to match or exceed human performance ability in others. Examples may be the lack of extended (or indeed standard)

techniques implemented on an instrument, inflexible control over techniques that are implemented, or inability to maintain high acoustic signal to noise ratio and musical dynamics due to extraneous actuation noise and other limitations.

The desire to make improvements with regards to these recurring problems with musical robots is the primary motivation of the work described in this thesis. As these various difficulties are addressed and improved upon, the musical capabilities of these automatons are expanded, barriers to building and making music with them are lowered, and they become much more practical, enjoyable, and inspiring to use. The following section describes in more detail the objectives of this work, and the techniques utilised to facilitate these improvements.

1.3 Objectives

Figure 1.2 illustrates the flow of information in a traditional musical robotics scenario. This ‘open-loop’ configuration shows a linear system where information flows in one direction from the musical controller, through an embedded control system, to an actuator whose movement activates a sounding object. This is the control paradigm utilised in many instruments created by Trimpin, in Eric Singer’s LEMUR project [144], in Kapur’s Karmetik Machine Orchestra [64], and in many other instruments and ensembles.

FIGURE 1.2: The conventional signal flow of an open-loop musical robotics system.

Several studies were carried out analysing the performance of the Robotic Taishogoto [85] and various open-loop robotic percussion systems [79, 80] to better understand the causes of their weaknesses. The studies revealed

various latency, dynamics, and accuracy issues with the range of mechanisms evaluated.

In industrial settings, when more precise results are required by computer controlled machinery, ‘closed-loop’ control systems are often utilised. In these configurations, measurements of the output of the system are fed back into the controller in order to inform it of the results of the actuations as they are happening. This enables the control system to continuously adjust itself towards a verified goal regardless of any disturbances that may have occurred. When applied to musical robots, such a closed-loop system would resemble Figure 1.3. While industrial systems commonly utilise temperature, distance, magnetic, or light sensors to feed information back to the electronic controller, the output of a musical robot consists of acoustic sound pressure waves produced by a sounding body. This means that the sensors utilised in closed-loop musical robotics should respond to this output. Example sensors include piezo-electric discs (contact microphones), magnetic coils (pickups), and electret, dynamic, and condensor microphones.

Closed-Loop System

FIGURE 1.3: The signal flow of a closed-loop musical robotics system. The loop is represented by the red arrows.

Unfortunately, though temperature or distance sensors generally provide sufficient information to close the loop in industrial equipment where the goal is a specific temperature or position, a microphone or a pickup alone is not sufficient for closing the loop with a musical robot, since the desired output is not only the *existence* of sound, but the desired sound with *specific musical characteristics*. In order to inform the controller as to whether or not

it is achieving sounds with the correct characteristics, it is necessary to analyse the signals produced by the audio sensors to deduce relevant musical information such as amplitude (loudness), fundamental frequency (pitch), temporal variation (envelope shape), spectral content (timbre), and latency of onset (timing).

Central to this thesis is the hypothesis that by applying closed-loop techniques to musical robots, improvements can be made to the discipline's shortcomings outlined in the previous section. Each of these aspects of musical robotics performance practice are outlined below, and an explanation provided as to how closed-loop control systems can improve upon these weaknesses.

1. Musical robots need to be robust and reliable.

To meet the demands of a majority of potential users, musical robots must be 'gig-ready', capable of being set up efficiently in music performance environments and being relied upon with the level of trust that is afforded to other electronic musical instruments. By equipping them with closed-loop systems that enable them to 'hear' their own output, they are capable of automatically calibrating their tuning, dynamics, latency, and other characteristics. If their condition changes during a performance, they can dynamically adjust their calibration upon observing the change, just as a human might in a traditional performance.

2. It needs to be easy to make music with and compose for them.

Musical robots must be able to respond in ways that are predictable and easy to understand for musicians who do not possess specialist knowledge about the construction of the robot they are using or the wider field in general. Closed-loop control systems can enable this by allowing users to specify *musical* outcomes, such as note names, dynamics markings, timbral, or temporal characteristics, and having the system automatically internally adjust its own control signals to match these specifications at its output.

3. They need to be musically expressive and capable.

Though many musical robots in the present day possess several parameters for musical expression and entire theses have been dedicated to this topic [23, 28, 93, 102], it is unlikely that musicians will ever be entirely satisfied with the expressive capabilities of their musical robots. Closed-loop control systems have the potential to enable entirely new

capabilities in instruments that require audio-based feedback to be expressive. In addition, the streams of analogue and digital audio signals that are obtained as a part of the process of closing the loop can be utilised in a multitude of ways such as for amplification purposes, hybrid musical robotic / live electronics electro-acoustic performances, live and in-studio recording situations, and interactive musical robotic installations, among others.

Since there exist several families of musical instruments which each present individual and unique challenges both in implementing mechanical robotic systems and closed-loop control systems, in order to meet the above-mentioned objectives in a comprehensive manner, three distinct applications are examined. These are polyphonic pitched percussion instruments, unpitched percussion instrument arrays, and robotic chordophone instruments. By implementing closed-loop audio feedback systems in these three applications and evaluating their performance, a thorough understanding of the potential benefits of applying these techniques to the majority of existing musical robots is achieved.

1.4 Document Structure

So far this introduction chapter has provided an outline of the topic of implementing closed-loop audio feedback systems in musical robots. It has also discussed the motivation for this work and posed a hypothesis concerning the results of the research. This section outlines the structure of the rest of the thesis chapter by chapter in order to orient the reader and provide a guide for how this document may be digested.

The second chapter provides a context for the research of this topic by presenting an outline of the history of automatic musical instruments and sounding objects, which spans over a thousand years. This history then concludes with a survey of the current state of the art of musical robotics and discusses the range of related works in that area. The subsequent sections provide outlines of the related areas of hyper instruments and music information retrieval, and explain how the techniques used in those areas are utilised in this research project. This chapter also clarifies the terminology used, eliminating the many points of ambiguity that exist when describing this area of research, and explains the key points of differentiation that separate these

FIGURE 1.4: Thesis Structure. The arrows represent recommended paths to read the document.

techniques from those used in closely related areas. A brief summary of the types of control systems that are in use in the more general field of robotics is then provided, and the extent to which these types of systems have been implemented in musical robotics thus far will be examined. The chapter finishes by outlining the various opportunities for original research that exist in this area, and research questions are presented that precisely articulate the primary issues that need to be addressed and establish the criteria used for evaluating the success of the project.

The following four chapters consist of studies applying closed-loop audio feedback techniques to three of the most common varieties of robotic musical instruments: polyphonic pitched percussion, unpitched percussion arrays, and chordophones, each with an example instrument. The third chapter describes the design and construction of the Closed-Loop Robotic Glockenspiel, a new robotic, integrated polyphonic percussion instrument. It begins by presenting the mechanical design process of the instrument, followed by detailed descriptions of the electronics and the closed-loop control mechanisms. Both the digital and analogue control circuitry are described, and the firmware of the instrument is explained. The workings of the embedded musical information retrieval (MIR) routines are outlined, and the components of its low-level latency and amplitude detection algorithms are illustrated. The continuous calibration routines are explained, and the various

features and functionality that these novel features provide the instrument are detailed. Evaluations of the characteristics of the instrument with regards to specific performance criteria are presented, and the results are examined and the advantages and disadvantages of the chosen implementation are discussed. The composition of both a performance and an installation work for the new instrument is also outlined. The results of the evaluations and practical experience from public performances are used to improve the design of the systems created in the following chapters.

Chapter Four begins by presenting a study that evaluates a number of different robotic percussion mechanisms categorised by actuator type. The results of these evaluations guide the design of a new robotic drum kit. The striker construction, mounting considerations, and mic'ing techniques utilised in the creation of the kit are then detailed. A novel system for providing increased expression in robotic hi-hats is then presented, complete with diagrams of the mechanical positioning system and schematics of the control and power supply circuitry.

In Chapter Five, a new modular system is created in order to implement the closed-loop audio feedback concept with an array of unpitched percussion instruments, trialling the system with the new robotic drum kit. The design and construction of the various subsystems of the control unit including power supply, robot control board, embedded digital signal processing (DSP) board, preamp, and digital audio converter boards are all outlined, and the actuation hardware for the various percussion instruments is described. The audio transduction methods are illustrated, and the software components of the system are detailed, outlining the algorithms utilised for the specific needs of individual instruments, with particular attention given to the unique requirements of the hi-hat. The capabilities of the system are described and evaluations are undertaken which determine the extent to which closed-loop techniques are beneficial when used with these instruments. The process of creating a two-part composition for the robotic drum kit with the new closed-loop control system is also detailed.

The sixth chapter expands the research into the realm of robotic chordophones: instruments which produce sounds via the vibration of strings. In particular, a portion of the DSP hardware developed in Chapter Five is utilised to apply closed-loop control techniques to an existing robotic chordophone that has been developed by Neill Skelly. The process of augmenting space-constrained existing instruments with closed-loop audio feedback by

integrating additional embedded DSP hardware is examined, and the unique approaches required to implement these algorithms with robotic chordophones are explained. The performance of the augmented system is evaluated and the results of the additions are discussed.

The final chapter summarises the key contributions to the field presented in this thesis and reviews the results of the evaluations conducted in each of the preceding chapters, outlining the various benefits that closed-loop control systems confer to musical robots. Conclusions are made about the effectiveness of these techniques, and reflections are provided that outline how future iterations of these designs could be improved. Lists of research outputs such as publications, public performances, awards, and media coverage are also presented. Lastly, the many avenues for future research that this work has opened up are outlined and the outlook of the field in light of these developments is discussed.

1.5 Summary of Contributions

By realising an improved system of controlling musical robots, this thesis lays the foundation for the widespread creation of a new generation of robotic musical instruments that can be interacted with in a musical manner and afforded a high level of trust in performance situations. This allows musicians, composers, and other practitioners in the field to concentrate on creating compelling musical work rather than focusing on maintaining, fixing, and calibrating machines. Through the process of designing, developing, and evaluating a number of musical automata of various categories to be controlled via this closed-loop system, mechanical designs, source code, schematics, and a workflow for creating these instruments is established. The musical works arising from the creation of these new systems demonstrate their novel capabilities, and act to realise the technological developments of this research in concrete musical performances and installations. An itemised list of specific novel contributions made in this thesis is provided in the Conclusions chapter, Section 7.1.

The twelve peer-reviewed publications, 14 concerts, ten awards, and several features in the media that have occurred as a result of this research highlight the various auxiliary contributions this research has also made to the field.

Chapter 2

Related Works

2.1 Overview

There are many areas of research that influence and are related to musical robotics. Most significantly, the field borrows a large number of its designs, concepts, and inspiration from a rich history of musical automata and other programmable and non-programmable sounding objects. This chapter begins with a brief outline of the history of automatic musical apparatuses from its humble beginnings over a thousand years ago, and traces it through its golden age during the industrial revolution. The field's journey is followed through its transformation into musical robotics in the 20th century, and this history concludes with a description of the current 21st century state of the art.

The second half of the chapter outlines how the fields of hyper-instruments and MIR can provide relevant and useful approaches for improving musical robots, and compares open-loop and closed-loop control system architectures. Following this, an outline of various closed-loop control systems that have been utilised thus far in conjunction with musical robots is provided. The state of the related area of musical robot calibration is then summarised and the gap in current development is established, demonstrating the opportunities for novel research to be undertaken.

2.2 A History of Musical Automata

Many modern-day robotic musical instruments draw inspiration and design ideas from the long history of early, non-computerised devices. Knowledge of the historical context of the creation of these various apparatuses firmly

places our modern incarnation of the field atop a solid foundation, informs future designs, and establishes the field's position within the musical fine arts.

2.2.1 Pre-18th Century

FIGURE 2.1: Time-line of significant pre-19th century automatic musical instruments.

The first occurrences of man-made self-playing musical instruments were apparatuses which were powered by nature. The earliest example of this is something which is still in use today and commonplace around the world: the wind chime. Paleolithic archaeologist April Nowell suggests that early humans may have used spear heads as wind chimes [11], but more concrete archaeological findings put the first known wind chimes, made of various materials such as bamboo, wood, bone, and shells, in South-east Asia circa 3000 BCE [171]. They were also developed independently by the ancient Egyptians, forged in bronze around 2000 BCE [157]. During the Western Zhou Dynasty (c.1100-771 BCE), the Chinese began casting bells called 'yong zhong', which had a more modern, sonorous characteristic and led to the use of this method for creating wind bells. These would be hung at the northern, eastern, southern, and western corners of temples, among other places. Another variation of the wind chime idea was used in ancient Rome around the turn of the common era. Called 'tintinnabulum', these were ornately decorated statuettes with wind-bells hanging from them, often featuring phallic imagery [53].

Another instrument which has ancient origins and is also played by the wind is the aeolian harp, shown in the middle of Figure 2.1. Named after the Greek god of wind, Aeolus, the more-or-less standard design involves mounting

a number of strings to a sound board in such a way that the strings will be excited by the wind, making use of the von Kármán vortex street effect. There are commonly from six to twelve strings of varying diameters, but tuned to the same pitch. The effect is an ever-changing and ethereal sound [137].

The exact origin of the aeolian harp is not known for certain, but according to Stephen Baker's *The History and Organology of the Aeolian Harp*, a possible origin could have involved a dead sheep being left on a post with its dried stretched tendons and intestines acting as strings, and the rest of its body acting as a sort of resonating chamber [10]. Thomas Leroy Hankins and Robert J. Silverman's book *Instruments and the Imagination* [43] also mentions a number of references throughout history to these kinds of instruments playing themselves, but states that the first known reference to the creation of an instrument specifically for this purpose belongs to Athanasius Kircher in his *Musurgia Universalis*, which was written in 1650 [19]. These wind-powered instruments can be considered as the earliest kinetic sound sculptures, likely created with similar motivations to many contemporary sound artists.

While wind was indeed an apt source of power for early automatic sound-making devices, there are also records of a number of devices which used water to drive their mechanics. For example, *Greek and Roman Technology: a Sourcebook* by John William Humphrey provides annotated translations of ancient Greek and Latin texts which describe the workings of a number of such machines. These included contraptions like the 'Singing Bird Automaton', which uses running water flowing through an array of pipes and funnels to move models of birds and push air through pipes at certain intervals to imitate bird-song [48].

Also powered by water and created around the year 875 CE, the Banu Musa water organ, described as 'The Instrument, which plays by itself' in their *Book of Ingenious Devices*, is an important milestone in the history of automatic music. Though the ancient Greek Hydraulis water organ, which dates back to the 3rd century BCE, used water to generate the air pressure for the pipes, the musical sequences were not played automatically. The Banu Musa organ was the first to utilise interchangeable cylinders with raised pins on the surface to program the behaviour of the instrument, a paradigm which has remained relevant over a millennium later [139].

A few hundred years later during the Islamic golden age, Al-Jazari invented a range of water-powered mechanical devices and recorded his work in his *Book of Knowledge of Ingenious Mechanical Devices* [50]. His many contraptions

included musical devices such as a ‘Perpetual Flute’, which is designed to play “as long as the water flows” and is described systematically in great detail [69]. Perhaps Al-Jazari’s most impressive musical automaton, though, is the ‘Boat of Automata’ which is a boat placed on the surface of a large pool of water that carries an automatic flute player and a group of mechanical ‘servant girls’ playing other instruments. Every half hour they take turns performing, and “they do not desist until they have performed about fifteen times”.

FIGURE 2.2: A diagram of the pinned barrel mechanism as used in a carillon.

During the 13th century in Flanders, as the sets of bells in bell towers grew and clock technology advanced from water clocks to mechanical apparatuses powered by weighting systems, the sets of bells were attached to pinned-barrel type mechanisms so they could ring automatically, eventually giving different melodies on the quarter, half, and full hour. This machine was called the carillon, and the basic workings of it are illustrated in Figure 2.2. As they gained popularity and continued to be developed in the Low Countries of western Europe through the following centuries, the carillon became a full-fledged musical instrument. In 1480, the first carillon to be played manually by a performer using a set of batons and pedals arranged in the shape of a keyboard was constructed in Flanders [32].

Not only did the larger carillons get bigger and more sophisticated, but smaller compact versions of these instruments also began to emerge, both as time-keeping systems and as musical playback devices. Pinned-barrel mechanisms were used to activate not only bells, but other instruments as well. Well-known composers such as Beethoven, Haydn, and Mozart wrote music specifically for these automatic instruments [30]. From the 14th to the 18th century, as the town clock was gradually downsized to become the pocket watch, the carillon shrank to become the music box [68]. The music box as we know it today was invented by Antoine Favre in 1796 in Geneva, Switzerland. He coupled a spring-driven mechanism with removable cylinders, and had the cylinders' pins directly activate a comb-like line of fine-tuned steel teeth to create the sound. This design became extremely popular in Switzerland and abroad, and was subsequently used in machines from tiny watches to large furniture-like cabinets. Figure 2.1 provides a visual time-line of the significant mile-stones in early automatic music outlined above. By the dawn of the industrial revolution, a rich groundwork of musical automata had been laid.

2.2.2 The 19th Century and the Golden Age of Mechanical Musical Instruments

FIGURE 2.3: Time line of important developments in automatic music during the 19th and early 20th centuries.

The beginning of the 19th century saw significant developments in the field of musical automata, as shown in Figure 2.3. The most noteworthy development of this time was the Panharmonicon, invented by German musician Johann Nepomuk Mälzel and first exhibited in Vienna in 1804. The machine used the same type of pinned-barrel mechanism that had seen use in the

Banu Musa device almost a thousand years earlier, but they were used to trigger a very large array of orchestral instruments, and each piece of music was sequenced on a barrel which was over one metre wide. Figure 2.4's left picture shows the Panharmonicon being operated, middle picture shows some of the automated instruments in detail, and right picture shows several pieces of music that were sequenced on pinned barrels [89].

Johann Nepomuk Mälzel also happened to be an acquaintance of Ludwig van Beethoven, and Beethoven wrote his *Wellington's Victory* (also known as the Battle Symphony) for the Panharmonicon in 1813. The Panharmonicon is the earliest example of what later became known as orchestrions - large integrated groups of automatically playable instruments, usually contained within a large wooden furniture-like case. Orchestrions often contain a combination of various families of instruments from pipe organs, to pianos, to drums, and even string instruments, all played from one central sequence. Since its inception, the orchestrion has taken many different forms and can be considered an automatic version of the photoplayer and the theatre organ used for accompanying silent films and shows. Though not in common use for this purpose today, modern orchestrions are still being built, a prominent maker being Ken Caulkins of Ragtime West [17]. Another contemporary reference is the studio album and live musical robotics show by jazz guitarist and composer Pat Metheny entitled simply *Orchestrion* [90].

FIGURE 2.4: The Panharmonicon, invented in 1804 by Johann Nepomuk Mälzel (not pictured).

Orchestrions began to be built by many different companies and the metaphor of pinned-barrel-based instruments became more and more prevalent with instruments such as the barrel organ and barrel piano becoming commonplace on the streets of Europe in the early 19th century [26]. The barrel piano had no keyboard and was usually powered by a person turning a crank. The pin mechanism also had no way of dictating the strength of the

strikes and the length of held notes, two defining characteristics of the piano. As a result, during the rest of the 19th century many attempts were made to create a true piano that could perform by itself.

To address some of the above-mentioned limitations, after over a millennium-long run, the pinned cylinder was finally superseded by a roll of perforated paper called a piano roll. This invention was patented in 1883 by Emil Welte [169] and fulfilled a similar purpose to modern transistors – utilising low power signals to switch higher power ones, though using the power of air-flow rather than electrical current flow. Its paradigm is still mimicked today in modern music sequencing software, as shown in Figure 2.5. This mechanical isolation of the programmed instructions from the actuating machinery enabled note lengths to be specified and in some cases various expressive parameters such as dynamics were able to be programmed or manually affected during a performance by controlling the air-flow with pedals.

FIGURE 2.5: Player piano roll (left) and modern digital piano roll (right).

In the late 1800s two primary configurations of automatic piano playing were in use: pianos with internal machinery, called player pianos, and external devices used to play the keys of a standard piano, called piano players. A very early pneumatic player attachment, designed by Merritt Gally of New York, and illustrated in *Scientific American* in June 1879 [37], is pictured in the timeline in Figure 2.3. At first, the piano players were more successful,

since they didn't require people to buy entire new pianos. However, after public confidence in the new technology rose, consumers were more willing to replace their old piano with a new, automatic one. There were countless makes and models of player pianos, but the most notable is the Pianola from the Aeolian Company, to such an extent that the word 'Pianola' has become generally interchangeable with the term 'player piano' itself. A further development on the player piano concept was introduced in 1904. Called the 'reproducing piano', it allowed the great pianists of the day to record their performances in relatively high expressive detail in real time directly to piano rolls, to be reproduced later at will. These recordings remain particularly valuable to music historians to the present day, and have been utilised in Stanford University's Player Piano Project [16].

Music boxes also continued to be developed during the 19th century: adding more springs for extended playback, larger range, and including extra instruments into the sequence all became possible. In 1885, Oscar Paul Lochmann of the Symphonion company introduced an entirely new form of music box. Instead of using a cylinder with pins attached to pluck the comb, he invented a disk version, which used perforated metal disks whose holes would then trigger the mechanism to play the comb's teeth [116].

This new form factor offered a number of benefits over the cylinder. Firstly, the disks were much easier to replace in the machine, just needing to be placed on a spindle in a similar way to modern turntables (as shown in Figure 2.6). Secondly, the disks were a lot cheaper and easier to produce; instead of requiring factory workers to painstakingly attach pins to cylinders one at a time, the patterns of holes on the disks could be stamped out on an assembly line.

For a short time, these disk music boxes were very successful and in the year 1899, the Polyphon Music Box Factory decided to expand to the United States and sent workers and equipment there to form the Regina Music Box Company in Rahway, New Jersey. The Regina music box became well known, and they even developed the first automatic disk-changing machine, the precursor to the jukebox [124].

It is said that from about 1890 until about 1920 was the golden age of automatic musical instruments, as the instruments were built and distributed on an industrial scale and music was beginning to be accessible to the masses [135]. As contemporary music of the early 20th century strove to push

FIGURE 2.6: Disc music box (left) and modern turntable (right).

the boundaries and become more experimental, composers such as Paul Hindemith, Ernst Toch, and Gerhart Münch discovered that by utilising automatic musical instruments, they were able to radically expand the traditional role of the composer and entirely bypass the performer's interpretation of the work, delivering an unadulterated, precise performance of the work exactly as the composer intended [122].

However, the commercial development of the phonograph and broadcast radio in the 1920s saw the sales of automatic pianos and other automatic instruments begin to experience a downturn. By the end of the decade their sales were in a weak position and after the great Wall Street crash of 1929, they practically ceased. Despite this, some composers of contemporary music of the time remained undeterred and continued to write music for the automatic instruments. Examples include *Ballet Mechanique*, written by George Antheil in the late 1920s and the player-piano works of Conlon Nancarrow, starting with *Rhythm Study No. 1*, completed in 1950 [38].

Nancarrow was undoubtedly the player piano's most prolific composer, writing over 50 studies specifically for the instrument from 1950 to 1993. He was also one of the notable inspirations for the creation of a new generation of musical automata later in the century.

2.2.3 The Dawn of Musical Robotics

In the 1970s, a few pioneering artists inspired by traditional musical automata made use of the relatively new transistor and microprocessor technology to revitalise the concept of automated instruments - this time with computer control. Computer controlled instruments created many new compositional possibilities, provided the ability to create previously unheard sounds, and allowed many new modes of interaction that were impossible with traditional musical automata. This new art form has since been dubbed 'musical robotics'.

The most prolific pioneers of this 1970s renaissance were Godfried Willem Raes, Trimpin, and Ken Caulkins. Raes' works have accumulated to become the world's largest machine orchestra, called the Logos M&M Robot Orchestra [87]. His earliest automatic sound sculpture, the Bellenorgel, is shown in Figure 2.7. Over the previous four decades, Trimpin has created some of the most large scale and innovative installation works of musical robotics in the field [27]. Ken Caulkins' musical automata business Ragtime West has provided a very large range of installations and performance instruments internationally for high profile clients such as Disney and Pat Metheny [17]. Each have drawn much inspiration from the breadth of historical musical automata and played their respective parts in seeding the field of musical robotics with their unique artistic aesthetics and technical inventions. All of these pioneers are still very active in the field to the present day and much of the research in this document draws inspiration from their work.

The 1980s saw musical robotics diversify into several new areas as new artists, researchers, and companies entered the field. One of the most notable groups was in the Waseda University Science and Engineering department who unveiled WABOT-2, a humanoid robot that was capable of playing a keyboard [154]. Throughout the 1980s and beyond, the group has continuously developed a range of anthropomorphic musical robots such as a flautist [152], saxophonist [123], and others for "understanding better the human motor control as well as serving as a means for understanding better how robots can transmit sensitive information to the human partners" [150].

The same decade saw Nippon Gakki (now Yamaha) develop the Disklavier [121] in 1982, and in 1985 Bösendorfer developed the 290SE, both computer controlled acoustic pianos and both influenced by Wayne Stanke's

FIGURE 2.7: The Bellenorgel - a pre-microcontroller automatic sound sculpture. Credit: Godfried Willem Raes

work in the late 1970s and early 1980s [101].¹ These robotic pianos and subsequent iterations have been used extensively in commercial, artistic, and academic contexts such as the Mirror Fugue project by Xiao Xiao [172].

The 1990s saw a steady stream of new artists join the established pioneers from the 1970s and 1980s, each contributing novel techniques and new artistic aesthetics to the field. While academic laboratories such as the Takashima lab of Hosei University utilised their engineering expertise to build cutting edge saxophone, trumpet, and trombone playing robots [155], several artists made use of the technology to introduce their own creative style. Meiwa Denki emerged with various novel robotic guitars and percussion instruments imbued with a sense of humour and playfulness [41]. Jay Vance created his somewhat theatrical ‘Captured! By Robots’ robot band complete with automatic guitar, bass, drum, and horn players, and Roland Olbeter presented *Afasia*, a stage production that included robotic violin, guitar, drums and bagpipes [54]. It was also in the late 1990s that QRS music technologies released their own take on the computer controlled acoustic piano, additionally releasing a robotic violin called the ‘Virtuoso Violin’ [35]. These works brought the world of non-loudspeaker electro-acoustic music onto the theatre stage and often merged them with a range of hyper instruments and

¹After developing their CEUS robotic piano technology in 2008, Bösendorfer was acquired as a wholly owned subsidiary of Yamaha Corporation. [173]

other related art forms such as contemporary dance.

The 2000s is the decade where musical robotics and other non-loudspeaker-based electronic music achieved its tipping point and started to show the seemingly exponential growth shown in Figure 2.8. The growing accessibility of microcontroller technology coupled with the broad attainability of information via the internet lowered the previously high technical barrier to entry and enabled artists and musicians with comparatively lower levels of engineering training to participate in and contribute to the area. In *Trimpin: An Interview* [71] Trimpin provides one explanation for the popularity of musical robotics in the 21st century:

In pure computer music, there was always an element missing when performed in public. For most of the audience, it wasn't enough to watch. OK, somebody is just sitting in front of a laptop. However, when you go to a concert, you can see how the whole body is playing an instrument, when you watch a musician or an ensemble or an orchestra. Just watching somebody playing a laptop was always a little difficult.

And when the laptop got another component, a robotic device, suddenly it was bridging over to the next step where the audience or the individual could see where the sound was coming from, what actually happened at the moment when this was moving or when that was turning.

It's not just the laptop, it's also the sound itself. When sound is emitted from any kind of resonant body, or from any kind of source material like metal or wood or whatever, the body reacts differently. It's a different experience because the acoustics work differently and the whole part is resonating. Everything is a part of the music. It's emitting the sound, your body is a part of it, and that makes it a more complete circle for the audience to understand.

New collectives such as Eric Singer's Lemur [144], Kurt Coble's P.A.M. Band [148], the Karmetik Machine Orchestra [64] and EMMI [138] arose, most of them facilitated by academic institutions and each contributing a range of new musical robotic technologies. As musical robotics began to be accepted as a discipline within the wider field of computer music, international music conferences dedicated paper sessions, workshops and keynotes to the

FIGURE 2.8: Approximate entry dates of musical roboticists from 1970 to 2015.

topic [114]. So many new musical roboticists emerged during that decade that a comprehensive list is beyond the scope of this chapter. [56], [106] and [87] provide more detailed information on the developments made during that period, and Figure 2.8 provides a time-line of approximate entry dates of musical roboticists from 1970 to 2015.

2.2.4 The State of the Field

While the last half-decade has seen a continuation of the acceleration of the musical robotics adoption seen in the previous decade, the most significant development is that musical robotics have now penetrated into the sphere of popular music, with artists such as Björk (built by Andy Cavatorta) [8], Aphex Twin (built by Godfried Willem Raes) [1], Squarepusher (built by a team assembled by Zima) [153], and Pat Metheny (built by Eric Singer, Ken Caulkins, and others) [90] all releasing albums or EPs making use of robotic instruments as primary features.

The business world has also begun to take notice of the appeal, and several large corporations have sponsored the creation of robotic musical ensembles. Examples of this include Anheuser-Busch commissioning Andy Cavatorta to create robots that utilise Stella Artois beer glasses to create music [18], Molson Coors Japan commissioning a team of engineers to build the Z-Machines musical robot ensemble to promote their Zima alcoholic beverages [147], Moleskine commissioning Felix Thorn to create the Moleskine Orchestra using their products, and Festo working with Roland Olbeter to create their *Sound Machines*.

With all of the recent success and publicity that the field of musical robotics has garnered with these high profile works of mainstream and commercial nature, it is clear that interest in automatic musical instruments is at its highest level since the golden age of musical automata a century ago. Despite this, much research and development is still necessary in order for these instruments to begin to approach the level of reliability, usability, and capability that many mainstream electronic musical instruments possess.

2.3 Defining the Terminology: Musical Robotics

Within the area of research concerned with the actuation of real-world acoustic sounding objects there are many related areas with similar sets of terminology. This situation creates the potential for ambiguity with regards to the goals of this research and the scope that it takes place within. This section begins by presenting the definition of musical robots used in this thesis, and then compares this with other definitions and delineations that have been used by researchers in the field.

For the purposes of this thesis, ‘musical robotics’ is the area of research that is concerned with utilising moving actuators to activate sonorous objects under digital control. Several related areas and their relationship to these definitions are outlined below.

Transducer instruments are those in which small motors, voice coils, or other types of transducers are affixed to a resonant object and are made to vibrate under remote or local control in order to create acoustic musical effects in installation and performance contexts. These are also often called ‘actuated instruments’ [117], and do not have actuators that move through the air. Some examples of such instruments include the EMDrum [136], The Bistable Resonator Cymbal [125], the Overtone Fiddle [118], the Self-resonating Feedback Cello [24], and the Feedback Lapsteel [44]. The fact that these instruments use permanently fixed vibrating transducers in order to generate their sound separates them from the field of musical robotics, which is predominantly concerned with instruments that have actuators that move through the air, such as solenoids and servos.

A related area is that of *electromagnetic resonance instruments*. Similar to transducer instruments, these instruments possess no moving parts other than their sounding bodies, and instead of utilising physically vibrating transducers to create sound, these instruments apply magnetic fields in order to reinforce, dampen or otherwise shape the sound of metal strings or bars to produce musical results. Some examples of these include the Magnetic Resonator Piano [96], the EMVibe [15], and the Electromagnetically Sustained Rhodes Piano [140]. Though some robotic musical instruments such as Raes’ <Rodo> [130] and EMMI’s MADI [138] include electromagnetic resonance capabilities as part of their instrument, non-robotic electromagnetic resonance instruments are not included under the umbrella of musical robotics for the purposes of this research.

Within the field of musical robotics, there are several sub-categories, each of which have varying research-related and artistic goals. Gopinath separates robotic drumming robots into two categories, ‘Independent Robotic Drummers (IRD)’ which are mounted to or near the drums to be struck, and ‘Body Augmented Robotic Drummers (BARD)’ which are wearable prostheses or supernumerary limbs which augment human performances [23]. The percussion-related research presented in this thesis is primarily concerned with the former category.

Troy Rogers of EMMI separates musical robotics into two categories based on their musical goals:

The contemporary field of musical robotics spans a wide range of research and creative activities. Within this diverse field, we can distinguish between 1) emulative machines that help researchers better understand and/or replicate human performers, and 2) inventive machines developed by composer-builders seeking new vehicles for musical expression. [138]

This distinction between emulative machines and inventive machines is often mirrored in the distinction between anthropomorphic and disembodied forms of musical robots respectively. Examples of the former include the work done at Waseda University [150], Toyota, and Ryukoku University [142]. However there are also significant numbers of musical robots that sit between the two paradigms such as those of the Z-Machines and Compressorhead robot bands, that possess some elements of anthropomorphism but maintain super-human abilities for creative purposes. Other examples of these ‘hybrid’ robots include Weinberg’s Shimon [166] and Haile [168] robots, and McCallum’s Mortimer [92]. These machines maintain vaguely human-like forms in order to conduct research in the area of robot musicianship, that is, the ability for robots to analyse, perform, and compose music based on theoretical foundations in fields such as music theory, computer music, cognition, artificial intelligence, and human–computer interaction [167].

Another delineation based on musical goals and output was proposed by Mo H. Zareei in discussion of his ‘Rasper’ mechatronic noise intoner [175]. Zareei considers ‘musical robotics’ to revolve around “mechanisms that replace the human performer, but not the instrument itself”, which in turn means that these instruments tend to output sounds that are more identifiable as ‘musical’ in the traditional sense. This is contrasted with ‘mechatronic sound-objects’, devices that are conceptually rooted in the tradition

of Luigi Russolo's *intonarumori* (noise-intoners), a category that Rasper adheres to. While the research outlined in this document is primarily concerned with musical robots that belong to Rogers' 'inventive' category and Zareei's definition of 'musical robotics', the systems developed throughout are also equally applicable to works that fall in the 'emulative' and 'mechanical sound-object' categories.

The term 'musical robotics' itself is also not without ambiguity. Ness refers to the area as 'music robotics' [111], while Sheffield prefers to include his work using solenoids to strike resonant bodies under computer control within the category of actuated instruments, choosing to define musical robots as necessarily composing musical material autonomously [141]. To add further confusion, Livingston refers to his 'Orb' devices which are mobile balls containing sensors and loudspeakers as 'musical robots' [72]. The research in this thesis is concerned with utilising actuators that move through the air to activate sonorous objects under digital control, and will henceforth be referred to as 'musical robotics'. The focus of this study is musical robots whose specific purpose is expanding the expressive options afforded to human composers and sound artists.

2.4 Related Areas of Research

Looking back over the history of the field and continuing to the present day, development in musical automata and musical robotics has primarily centred around the following:

- Expanding the timbral repertoire of robotic instruments by way of automating acoustic instruments that have never been automated before [75].
- Inventing new varieties of acoustic instruments and sculptures to be automated [29].
- Improving on existing musical robot designs in ways that increase their capabilities or add supplementary modes of expression [103].

Other, more specialised areas of relevant research include:

- Novel methods of interfacing with musical robots [73].
- Modular audio interfacing between robots [138].

- Robotic musicianship techniques [12, 13, 167].

While this thesis is focused on improving on existing musical robot designs, the research presented is multidisciplinary and touches on each of these different areas to a greater or lesser extent. In order to make the required developments, it is also necessary to draw knowledge and techniques from relevant areas such as those of hyper-instruments, MIR, musical robot calibration, the more general field of robotics, and both local and ensemble level musical robot control systems. These various fields are outlined below, and their relationships to this work are established.

2.4.1 Sensors in Musical Instruments

The implementation of closed-loop systems inside musical robots requires the use of appropriate sensors attached to the instruments, and necessitates the handling, interpretation, and processing of the information that those sensors generate. Fortunately there is a considerable range of literature that deals with the processes involved when utilising sensors with traditional and novel instruments in performance contexts. These sensor-equipped instruments have been dubbed ‘hyper instruments’ and are frequently used in the performance of contemporary electronic music.

As computers of a widely accessible level of cost and of manageable size also became powerful enough to process useful musical data in the 1980s, sensors were used in conjunction with musical instruments to extend musical performances through electronic means. Early pioneers such as Tod Machover and Joe Chung spawned the field of hyper instruments by constructing and modifying an array of musical instruments that doubled as electronic controllers and were linked to personal computers in order to add new expressive parameters to performances [86]. Over the following decades, sensor technology has rapidly progressed and the power of consumer computer technology has enabled the area of hyper instruments to transform the field of computer music performance, composition, and recording.

Many hyper instrument circuits are integrated with traditional musical instruments such as flutes [120], violins [174], and cellos [170] in order to make use of the skills of the majority of musical performers. An example of an augmented flute instrument is shown in Figure 2.9. The traditional musical

parameters of these instruments are augmented with sensors such as gyroscopes, accelerometers, light sensors, and force sensors to provide extra functionality to the performer without inhibiting their existing performance technique. [99] provides an overview of the various types of sensors commonly used in the creation of hyper instruments.

FIGURE 2.9: Cléo Palacio-Quintin playing her Hyperflute.

The outputs of the sensors used with hyper instruments are utilised in various ways. The predominant usage is providing performance data in real time during a musical performance to computers or other processing hardware, which is then used to manipulate the audio output of the system by adding effects or processing to the signal, or triggering and manipulating musical events [49]. Another use for the output of hyper-instrument sensors is in conducting research and providing insight into the methods and techniques of performing musicians. This information can then be used for pedagogical and musicological purposes [47]. There are also a number of recorded instances of musicians utilising hyper instrument sensor data to control musical robots. An example of this is Kapur's work integrating hyper instruments with the Karmetik Machine Orchestra [57].

The research conducted in the field of hyper instruments is useful in informing the design and configuration decisions that must be made when integrating sensor systems with musical instruments. This information becomes

pertinent when implementing such systems with musical robots in order to create closed-loop audio feedback based control mechanisms. The ‘audio transduction’ related areas of chapters 3, 4, 5, and 6 apply some of the principles of hyper-instrument sensor systems to the new musical robot control systems described in their respective chapters of this thesis.

2.4.2 Musical Information Retrieval

Another established area of musical research which is related to this project is that of MIR. MIR is the study of methods of extracting useful musical information primarily from streams of audio data, but also from other auxiliary sources such as sensors, musical scores, MIDI data, and even from mining internet data sources such as social media tags.

The earliest forms of electronic music information retrieval were analogue pitch and envelope follower circuits that received audio from musical instruments, microphones, or other sources and output control voltages that could then be used to control synthesiser and other hardware. Modern forms of MIR use sophisticated digital signal processing algorithms to extract much more detailed and accurate information from audio streams and can be used to achieve results such as extracting individual pitches from polyphonic instrument recordings, identifying genre information from musical recordings [161], and melody identification and matching [40].

The applications of MIR research are far reaching and span from commercial song finding and music library management applications to professional studio recording technology and to academic and contemporary music practice. The International Society of Music Information Retrieval has held an annual conference on the topic since the year 2000 and the Music Information Retrieval Evaluation eXchange (MIREX) has held MIR competitions since 2005. Several tool kits such as Marsyas [160], MIR Toolbox, and MIR_Eval [134] are available for conducting MIR research, and there is a wealth of information published on MIR algorithms such as the series of books by Julius Orion Smith III [146].

Many of the analytical tools and algorithms that are created in the field of MIR are directly applicable for use in the real-time analysis of audio and MIDI information generated by musical robots. Examples include onset detection, Fast Fourier Transform (FFT) analysis, pitch detection algorithms,

Route-Mean-Square (RMS) loudness analysis, and latency analysis among others. The various ways that some musical robotic systems make use of MIR techniques are discussed in sections 2.4.4 and 2.6, and the ways that MIR techniques are integrated within the newly constructed musical robotic systems presented in this thesis are described in the firmware sections of chapters 3, 5, and 6.

2.4.3 General Robotic Control Systems

FIGURE 2.10: Block diagrams of common control system architectures.

In the broader field of robotics, a variety of control systems are utilised to enhance the accuracy, reliability, and broaden the capabilities of robots. The most basic type of system, and the one which is most commonly used by musical robots, is an open-loop system, a block diagram of which is shown in Figure 2.10 (a). This configuration maps an input command directly to an output result, and real-world disturbances that influence the behaviour of the system are unaccounted for at the input stage. The result is that if

unexpected changes occur in the environment in which the robot is working, the output is likely to become inaccurate. In some applications this may not be a significant issue, but in others, may result in expensive or dangerous errors.

The second block diagram in Figure 2.10 (b) shows the signal flow of a feed-forward control system. In this configuration a mathematical model which was either conceived by the engineer creating the system or populated by a machine-learning algorithm predicts disturbances in the system and combines with the control input to produce a more accurate result. Though this is still technically an open-loop configuration, this architecture can still offer benefits over the configuration shown in Figure 2.10 (a), in certain cases justifying the added cost and complexity.

The final block diagram of Figure 2.10 presents the signal flow of a closed-loop feedback control system. In this configuration, the real-world output of the system is observed by one or more sensors, and this information is continuously fed back to the control input where it is used to counteract any disturbances that were introduced into the system. Of the three approaches shown, this is the only one that provides for the unpredictability of real-world environments, and is used by a large number of devices from simple temperature control systems to large industrial and military robotic systems to ensure the accuracy and repeatability of output.

While control systems theory is a broad field with many approaches, a commonly utilised method that is employed to integrate the fed-back sensor data with the incoming control signals in order to minimise output error is the Proportional-Integral-Derivative (PID) algorithm. A PID algorithm makes use of three individual parameters, Proportional (P), which is present error, or difference between the current output and the desired output, Integral (I), which is an accumulation of past errors, and Derivative (D), which is a prediction of future errors by considering the rate of change of error [2]. Various combinations of these parameters can also be utilised to implement P, I, PI or PD algorithms.

PID and related control system algorithms are equally applicable to driving closed-loop musical robots, which may utilise positional and other sensor systems to aid in conducting accurate movements. An example is the robotic hi-hat system described in Chapter 4. Some audio-based closed-loop systems, such as the pitch tracking system described in Chapter 6, can also utilise PID-style logic to improve their performance, while for others, such as

those presented in Chapter 3, algorithms that are more specific to these types of musical robots must be developed for the specific characteristics and limitations of these machines.

2.4.4 Pre-Performance Calibration of Musical Robots

Calibration and tuning are important issues for musical robots and various research papers have been published discussing these topics [105, 110]. There are several reasons why it is necessary for musical robots to be calibrated accurately, efficiently, and continuously. In many robotic installations and performances it is necessary for the devices to be able to work for extended periods of time without human intervention, and stopping for tuning or recalibrating drifting velocity responses and other characteristics is very undesirable. It is also highly beneficial to minimise setup time and reduce the number of mistakes made as a result of human error.

Some traditional robotic instruments such as the ones described in [74] require manual calibration by the operator before a musical performance to achieve satisfactory results. This manual calibration may take the form of altering firmware, sending Musical Instrument Digital Interface (MIDI) messages or other commands from a sequencer to the instrument, or utilising custom software plugins to calibrate the MIDI commands themselves as was previously done with the Robotic Glockenspiel described in section 3.2.

There are, however, a number of musical robots which can be commanded to run various types of pre-programmed calibration routines to initialise actuator position, tuning of the sounding body, or other parameters. One example is MechBass, which upon startup runs each fretter to its extreme position where a tactile switch is triggered and the starting point of the fretting mechanism is initialised [97]. Other robots such as Murphy's Swivel 2 [107] and EMMI's AMI are also able to carry out a pre-performance calibration by connecting an audio feed to an external PC loaded with pitch detection and custom calibration software. The pitch information retrieved during this calibration is then used to populate position tables for the fretting mechanisms.

There are also solutions for calibrating percussion robots such as the one outlined in Murphy's *Better drumming through calibration: Techniques for pre-performance robotic percussion optimization* [105]. A method that adds automatic channel assignment is described by Trail et al. in their paper *Musical*

Information Robotics: Coping Strategies for Musically Challenged Robots [111], where a Karmetik Notomoton [63] triggering several robotic strikers was used with a commercial microphone and custom computer software to automatically assign MIDI output tracks to the correct drums and provide calibration functions. This setup utilised MIR techniques such as onset detection, loudness analysis, timbre analysis and machine learning to implement these functions.

The most scalable example of such pre-performance calibration software is Mathews et al.'s *Tangle* [91]. The system utilises a client/server paradigm using the Open Sound Control (OSC) protocol and wired ethernet hardware, and was written in the multi-platform ChuckK programming language. Among various features, *Tangle* provides for audio to be returned from the robotic instrument to the *Tangle* server for the purposes of carrying out pre-performance automatic tuning, dynamics, and latency calibration routines.

Shawn Trail's STARI robot makes local, pre-performance tuning more integrated than instruments that require a personal computer, by accompanying the robot's electronics with a BeagleBone single-board personal computer running the open-source software suite PureData for pitch detection [159]. Instead of populating a tuning table like Swivel, STARI instead utilises a robotic tuning mechanism to adjust the fundamental frequency of the string in increments. These types of integrated pre-performance tuning technologies are also utilised in commercial products such as the AxCent and Gibson auto-tuning guitars and Trimpin's 'If VI were IX' installation implements tuning by way of modifying off-the-shelf guitar tuners. [109] presents a detailed discussion of the various technologies used in commercial and non-commercial instruments for automatic pre-performance tuning.

The capabilities of systems like *Tangle* and STARI share several common goals with the closed-loop systems proposed in this document: making musical robots more accessible, reliable, and easier to use. This means that many of the techniques offered by the area of pre-performance calibration of musical robots are useful as a starting point when implementing closed-loop audio-based systems. Indeed, the embedded MIR systems described in chapters 3, 5 and 6, can be considered an extension of these pre-performance calibration arrangements, adding continuous real-time analysis to ensure that the pre-performance calibration is maintained throughout the length of the performance.

2.5 Defining Audio-based Closed-loop Control

Since there are various types of audio feedback and closed-loop systems that have been developed for musical robots, it is necessary to define precisely the type of audio-based closed-loop mechanisms that are the focus of this thesis, and delineate these from related techniques. This section begins with such a definition, and compares this with descriptions of related areas so as not to confuse them. Section 2.6 then proceeds to describe instruments that do fit this definition and are directly related to the research presented in this document.

The audio based closed-loop control systems described in this thesis are those that receive audio signals via sensors or transducers, deduce musical information from these signals by way of embedded microcomputers, which then inform the systems that control the robotic musical instruments' actuators.

Electromagnetic resonance instruments such as those described in section 2.3 often utilise a type of closed-loop audio feedback to control their sounding bodies. One of the most prevalent configurations is using audio picked up from the oscillating material via an electro-magnetic pickup or piezo-electric disc, fed through an amplifier, and applied back to the oscillating material via a charged coil in order to reinforce or dampen the vibration of the material [95]. This technique allows instruments with metal strings or tuned bars to sustain oscillations indefinitely and for composers and performers to exert more detailed control over the amplitude of individual notes than is possible with most traditional instruments. Some sophisticated magnetic resonance instruments such as the Magnetic Resonator Piano [96] insert synthesized waveforms into the loop and synchronise their phases with the picked up sound via software.

The EBow² and the Moog Guitar³ are two examples of the commercial use of this technology and this technique has been implemented in several unique instruments such as the Electromagnetically Sustained Rhodes Piano [140], and robotic instruments such as Raes' <Rodo> [130] and EMMI's MADI [138]. These electromagnetic actuation mechanisms are examples of closed-loop *audio feedback* systems and they can provide extra musically expressive features. However, they are not examples of the closed-loop *feedback control* systems for musical robots defined above, which are the focus of

---

²<http://www.ebow.com>

³<https://www.moogmusic.com/products/Moog-Guitars/>

the research documented in this thesis. Transducer or actuated instruments, also defined in Section 2.3, are not included for similar reasons.

Some actuators that have been used with musical robots such as RC Servos have built-in closed-loop feedback mechanisms to allow the operators to command the actuator to rotate the device to a specific point along its arc. Some examples of musical robots that have used such devices include Swivel [104], Sweet Studio's percussion robots, and the author's Automatic Pomade and Singing Bowl Robots [74]. While these systems do aid with and simplify positioning processes, most do not provide the functionality to return the actuator's position to the control electronics or to alter the control scheme's coefficients.

The servo actuators used in more parametrically rich percussion robots such as Murphy's Nudge [103] and Gadd's Beatbot [36] provide their control hardware with information about the current angle of the actuators among other parameters. However, in these implementations, the hardware is not equipped with information about the position of the drum that is intended to be struck, and they also lack any musical information about the results of their actuations. As a result, these systems do not provide adequate information to the controlling hardware to accommodate for varying situations or account for environmental disturbances.

Several musical robots have also been created which make use of closed-loop feedback positioning systems in similar ways to industrial robots, and in fact, some projects such as [52] and [9] have used actual industrial robots. Examples of this include McVay's MechBass [98], Singer's Guitarbot [143], and Raes' <Vibi> [133]. These musical robots use various position sensors such as potentiometers and distance sensors to provide their control hardware with information about the position of moving mechanisms in order to position them precisely and accurately, in many cases using the types of PID algorithms explained in section 2.4.3. The Violina built by Andy Cavatorta also utilises a current sensing mechanism to deduce to what extent the bowing mechanism of the instrument is applying pressure to its glasses. These are all very effective developments that greatly improve the capabilities of these instruments and share the goals of this thesis, and a related system is devised for the positioning of the robotic hi-hat mechanisms introduced in Chapter 4. However, since these types of mechanisms are not continuously sensing the audio output of the instrument and deducing musical information from that output to inform their control systems, these types of systems

are not the primary focus of this thesis.

2.6 Contemporary Robots with Audio-based Closed-loop Elements

The concept of implementing audio-based, closed-loop control in musical robots has been written about for over two decades. Shibuya et al. wrote about their plans to create a ‘violin playing robot with KANSEI information-algorithm’ in 1997 [142]. Translated as ‘perception’ or ‘sensibility’, kansei is the ability for a performer to continuously listen to their own musical output and use that information to adjust their actions in real time in order to bring that output closer to an ideal goal. Shibuya’s lab at Ryukoku University realised at that stage that this was an important capability for their violin playing robot to have in order to more convincingly emulate human players. They undertook several studies of violin playing in order to investigate kansei’s role and application [149]. More recently, Wonse Jo et al. have also been undertaking similar studies, proposing their own audio-feedback based robotic violin system [51].

With similar goals, Solis et al. have also implemented an auditory feedback control system with their anthropomorphic flutist robot [151]. A pitch evaluation system was implemented by feeding audio from the performing robot to the computer controlling it, and conducting Fourier transform based analysis on the signal in order to detect when an incorrect note was performed. They noted that by using the feedback-error learning approach, improved control of their robot’s air pressure was achieved.

While these three examples of existing and proposed audio-feedback based musical robots fit squarely into Rogers’ ‘emulative’ category of musical robots, Barton et al.’s Cyther is a robotic chordophone that includes audio feedback in its control system and is positioned in the ‘inventive’ category. Cyther utilises a tension sensing mechanism to estimate the current pitch of the respective string, with a frequency analysis used to confirm this estimation [6]. This is indeed a promising configuration, but the errors in tuning and the latency of frequency analysis of the instrument demonstrate significant room for improvement.

2.7 Opportunities for Improvement

Several research papers and theses on the topic of musical robotics have listed real-time closed-loop auditory feedback as a future goal in their work [52, 111, 142], demonstrating that the implementation of these techniques is very desirable and an important development in the field. However, the previous section shows that very little has been successfully implemented in this area, as there are multidisciplinary challenges which must be overcome to create a workable system that makes real gains with regards to reliability, ease of use, and musical expression.

Most of the systems that do exist in this area (outlined in the previous section), are implemented as emulative systems with external computers in research environments, making them not appropriate for creative musical performance and installation environments. The inventive robot that does fit this category implements its audio feedback system with 8 bit microcontrollers at low clock speeds and bit depths, resulting in less than satisfactory characteristics. This clearly presents an opportunity to develop novel musical robotics systems with the custom-built hardware required to implement embedded real-time, high quality audio feedback systems in musical robots.

By bridging this gap in the research it is possible to create work that is useful in bringing the field of musical robotics forward. By creating a solution that others will be able to utilise in their own musical robot projects, this research has the potential to have a significant impact on the wider area. When musical robots are equipped with real-time closed-loop auditory feedback systems that focus on their musical output, composers and other musicians can then use them in more intuitive ways and trust the instruments to behave predictably. With this new functionality implemented, musical robots are equipped with the expressive capabilities needed to make the music of the future.

2.8 Research Questions

All significant bodies of research must set out to answer certain questions which have not been answered before. While the general field of research may have countless avenues of enquiry, these questions serve to focus research activities and clarify the concrete goals of the project. As a result of

the survey of automatic and robotic music present in this chapter, the following three questions which articulate the core of this research are posed:

1. Will the implementation of real-time closed-loop feedback based calibration systems for musical robots result in higher accuracy in performance than manually calibrated systems?
2. Can closed-loop musical robotic systems reduce the frequency of calibration and tuning required of open-loop systems?
3. Can real-time closed-loop feedback systems provide robots with added musical capabilities that would otherwise be impossible?

Throughout the course of the research presented in this thesis, these key issues of improving accuracy, robustness, repeatability, user friendliness, efficiency, and expressiveness will form the focal point of the investigations. The hardware constructed and the software written throughout the project aims to make significant improvements compared to traditional open-loop systems in all of these areas. In order to assess the validity of these improvements and answer these research questions in a concrete manner, quantitative analyses are conducted, and the details of these analyses will be outlined in the 'Evaluation' sections of chapters 3, 5, and 6.

The research presented in this chapter is featured in the following publication:

J. Long, J. W. Murphy, A. Kapur, and D.A. Carnegie. "Loudspeakers Optional: A History of Non-Loudspeaker-Based Electro-Acoustic Music". In: *Organised Sound* 22.2 (2017)

Chapter 3

The Closed-Loop Robotic Glockenspiel

3.1 Overview

Pitched percussion is an instrument class that is popular with musical roboticists. Unlike the playing of a cello, which requires several different mechanisms operating in conjunction in order to produce a single note, a single-actuator strike on a percussion instrument is able to produce a satisfying output that resembles a strike by a human. Examples of robotic pitched percussion include instruments in the Karmetik Machine Orchestra [64], Eric Singer’s SingerBots [144], the Logos Robot Orchestra [87], Ken Caulkins’ works [17], and Trimpin’s *Conloninpurple*, *Circumference*, and *Magnitude in C#* installations [31]. Computer-controlled pianos such as the Yamaha Disklavier also fit into this category.

Some of these instruments use mallets to strike keys from above, while others strike directly from below, allowing human performers to play the instrument simultaneously. Almost all designs utilise microcontrollers to trigger an array of linear solenoids. A number of studies such as [110] and [59] have demonstrated the highly non-linear characteristics of solenoids with regards to velocity control in robotic musical instruments. Others, such as [79] and [80], illustrate the extent to which robotic percussion mechanisms exhibit varying degrees of latency at a range of distances and velocities. These issues create timing discrepancies, limit the user-controllable expressiveness of dynamics in musical performances, and reduce the consistency and reliability of the instruments. It is these issues that are addressed in this chapter on the construction of a polyphonic pitched percussion musical robot with an integrated audio based closed-loop control system.

As polyphonic pitched percussion instruments are mechanically some of the simplest musical robots and their fixed pitches enable relatively straightforward algorithms for note detection, it was decided that this class of musical robot would be ideal for exploring the benefits and challenges of implementing the closed-loop MIR control systems described in Section 2.5. The glockenspiel was chosen as the most suitable instrument in this category due to its relatively low note count and cost compared to larger instruments such as vibraphone and marimba, compact size, and tone that closely resembles a sine wave. This sinusoidal waveform simplifies the algorithms utilised to detect pitch in the notes. It was also decided that rather than building a new robotic glockenspiel from the beginning, it would be most efficient to retrofit the existing robotic glockenspiel described in [74].

This chapter begins with a brief description of the original Robotic Glockenspiel, outlining the various characteristics of the instrument that cause issues in performance and can be improved upon by adding closed-loop control. Following that, the various approaches for implementing the new system are discussed, and the advantages and disadvantages of different electronic hardware options are outlined. The selected design options are presented, and the design and construction of the subsystems of the upgraded instrument are described in detail. The firmware, including the MIR algorithms used are illustrated, the methods of controlling the instrument are outlined, and the additional abilities that the upgrades grant the instrument are described. Evaluations of the effectiveness of the self-calibration algorithms are undertaken, comparing the original Robotic Glockenspiel with the new Closed-Loop Robotic Glockenspiel. Finally, the composition of *Broken Mirror*, a work written for the Closed-Loop Robotic Glockenspiel and surround sound live electronics, and the creation of the installation work *AutoReich* are discussed, and practical performance considerations for the instrument in these concert and installation settings are summarised.

3.2 The Robotic Glockenspiel

The pre-existing Robotic Glockenspiel design utilised a simple open-loop configuration and is pictured on the top left of Figure 3.1. The control box, pictured on the top right of Figure 3.1, received MIDI commands via either a panel mounted 5-pin DIN connector or USB. It produced higher voltage pulses which were sent to the instrument via a standard DB25 cable in order

to control the solenoids mounted under the glockenspiel keys. The width of these pulses of voltage dictate the strength of each strike of a key, with longer pulses resulting in louder hits.

The instrument was intended for contemporary music and installation performance and operated with very low levels of extraneous acoustic noise and comparatively low levels of latency (between 10 ms and 25 ms). This made it suitable for real-time interactive applications [73].

FIGURE 3.1: Top Left: The Robotic Glockenspiel. Top Right: The control box necessary for operating it. Bottom: An exploded rendering showing the inside of the instrument.

However, in performance it was noticed that a straight mapping of MIDI velocities to pulse widths that was identical for each key resulted in inconsistencies among the keys' volumes. Though manually editing the firmware of the control box with customised pulse-width ranges for each key provided

improvements in volume consistency, it was observed that over time and especially when the instrument was transported between venues, the manually input values would no longer provide satisfactory performance. In short, the act of using this instrument in a performance context resulted in the introduction of unacceptable inconsistencies.

To address these inconsistencies, the software interface shown in Figure 3.2 was created in the Reaktor visual programming environment. The plugin allowed the user to specify minimum and maximum velocities for each note of the Robotic Glockenspiel, and when inserted into the MIDI stream between the original signal and the instrument, it would map the input values between these limits.

FIGURE 3.2: The external software interface used to manually calibrate the Robotic Glockenspiel.

This open-loop method added convenience when calibrating the instrument, but the manual calibration method was still time consuming, and mapping the MIDI velocities to a small range decreased the resolution of the control signal. In addition, since even slight changes occurring during a performance with the instrument could have a significant effect on the velocity response, this pre-performance calibration was deemed insufficient. The lack of reliability meant that in order to prevent the missing of notes during very quiet passages due to slightly short pulse widths being sent, conservative minimum values had to be set. This had a significantly negative effect on the dynamic expressiveness of the instrument's performance. In order to solve these problems, it is decided to implement an embedded closed-loop control

system that is capable of continuously monitoring the performance of the instrument and autonomously making adjustments when required.

3.3 Design Considerations

Before beginning a retrofit of the instrument, the parameters and required specifications of the system must be decided upon and used to inform each stage of the design process. In line with the goals articulated in Section 1.3 that specify that robotic musical instruments should be more robust, reliable, and easier to use, the instrument should be as integrated as possible, without requiring separate controller hardware with extra cabling and connections. To enable this, all of the necessary equipment should be embedded within the enclosure of the instrument itself.

To make the instrument's calibration more robust and the performance of the instrument more reliable, a continuous automatic calibration capability must be included. The primary performance characteristics of velocity and timing must be made as consistent as practical. This consistency should result in enhancing the expressiveness of the instrument, allowing the composer writing for the instrument to program a larger range of dynamics safely. Other potential modes of expression that utilise the newly implemented hardware in the Closed-Loop Robotic Glockenspiel, such as cooperative play and creative use of the bleedless audio output of the instrument must also be accommodated.

In order to implement a closed-loop audio feedback control system within the instrument that allows it to both 'listen to' and 'understand' what it hears and continuously calibrate itself, several subsystems are required. A list following the signal flow from an arbitrary point in the loop is itemised below.

1. Transducers that convert the sound of the keys to analogue voltages.
2. A method of conditioning those voltages to impedances, bias levels and amplitudes suitable for analogue to digital conversion.
3. Analogue to digital conversion.
4. Digital signal processing hardware with firmware that allows the machine to deduce musical information from the signal.

5. A system capable of utilising that information to inform the performance parameters of the instrument and control the solenoid actuators used to strike the keys.

The options and considerations for carrying out these tasks are outlined in the following three subsections.

3.3.1 Audio Transduction

There are several potential methods of capturing acoustic audio from the instrument and converting it to analogue voltages. The most common way to record the sound of an instrument is simply to point either a dynamic or condensor microphone at it. However, it is necessary for the Closed-Loop Robotic Glockenspiel to analyse and calibrate the performance characteristics of each key of the instrument individually while a polyphonic piece is being played, and a microphone in a single position would be closer to some notes of the instrument than others and subject to bleed from surrounding sound sources. While some algorithms do exist that are capable of isolating individual notes of a single-track recording of polyphonic music, at this point in time, achieving that in real time at low latency with embedded electronics would require resources beyond the scope of this project.

Instead, rather than a single microphone, it is necessary to utilise individual transducers for each bar of the glockenspiel, and the fact that the instrument's keys are 20 mm wide, with 5 mm gaps between them, dictates that these transducers must be smaller than 25 mm in diameter. In small form-factor projects that require audio transduction, commonly utilised devices include electret microphones, MEMS (Micro Electro-Mechanical Systems) microphones, and piezo-electric discs, also known as contact microphones. While electret and MEMS microphones would likely be an improvement over a single microphone approach, they would still receive bleed from adjacent notes, increasing the possibility of erroneous analyses taking place. It is for this reason that commercial vibraphone pickup products like the K&K Sound and vanderPlas systems, and other experimental instruments such as the DSRmarimba [158] utilise piezo-electric discs attached to the nodal points of each key.

The piezo-electric method was trialled with the Robotic Glockenspiel, but due to the much smaller size of the keys in comparison to a vibraphone or

marimba, the attachment of even very small piezos to the nodal points of the keys resulted in noticeable dampening of their vibrations. The solution to this issue was inspired by electromagnetic resonance instruments like the EMvibe [15], which uses neodymium magnets attached to the instrument's keys, and makes use of a coil positioned underneath the key to induce motion in the magnet and therefore actuate the key itself. By reversing this process and instead using the coil to pick up sound from the neodymium magnet, it is possible to receive an audio feed of the glockenspiel note without interfering with the movement of the key, and with very little bleed.

After trialling several different coils, the one with the most suitable output was a 500 mH thru-hole inductor coil as used by Keisuke Kimoto in his *Are You Experienced II* exhibition [67]. This is utilised in conjunction with 1 mm thick, 4 mm diameter neodymium magnets attached to the glockenspiel keys. These small magnets provide adequate sound level and quality without dampening or detuning the glockenspiel keys by perceivable amounts.

3.3.2 Analogue Signal Conditioning

Before the analogue signals received from the transducers can be digitised for analysis, it is necessary to condition them in several ways. Firstly, the gain of the signal should be set so that the maximum expected amplitude will not exceed the maximum level of the analogue to digital converter (ADC). In addition, the signal should be AC coupled to eliminate sub-sonic frequencies and it should be low-pass filtered to remove ultrasonic frequencies that may be present in the signal. It should also be biased so that the waveform revolves around the mid point of the receiving ADC.

In the context of this project, this circuitry must be replicated 21 times, once for each note, and it must be small enough to fit into the instrument's enclosure. It is also necessary for the gain to be adjustable for each channel for initial calibration, and because this is the first example of such an instrument, it would be beneficial if the circuitry was physically accessible, with potential for modification during the development process. An additional benefit of the sound being captured by the instrument is that it is also possible to include an analogue audio output from the unit for the purposes of sound reinforcement, recording or live effects. To implement this, a 21 to 1 channel

mixer circuit is included in the analogue signal conditioning circuitry design to feed this output.

As there is no circuit board commercially available that is capable of fulfilling all of these requirements while fitting inside the Robotic Glockenspiel's enclosure, it is necessary to create a custom circuit board to carry out these tasks. The design of the schematic and layout of the board is discussed in Section 3.4.3.

3.3.3 Digital Signal Processing

At the core of the implementation of the instrument's audio based closed-loop mechanism is a digital signal processing system. This must be capable of receiving 21 channels of audio, conducting MIR analyses such as onset detection, latency calculation, and amplitude detection on all channels simultaneously and continuously in real time, and with low latency. Due to space constraints, the system should be multi-functional with a suitable range of peripherals. It must be equipped with a minimum of 21 channels of analogue to digital conversion at a sufficient resolution and sample rate on board, have multiple Universal Asynchronous Receiver Transmitters (UARTs) for sending and receiving MIDI and USB-MIDI commands, a bit-addressable Electrically Erasable Programmable Read-Only Memory (EEPROM) for saving calibration data between performances, and at least 21 GPIOs (General Purpose Input Output pins) for triggering the actuators. The CD quality specification of 44.1 kHz sample rate and 16 bit resolution is a suitable goal for the system's audio. It must also run at a high enough clock speed to carry out the required MIR analyses.

As this application requires a general purpose board that is capable of fulfilling several different requirements in a small space, the natural choice is some variety of microcontroller-based solution. The most commonly utilised boards in musical robotics are those of the Arduino series, particularly the Arduino Uno, which is equipped with an Atmel Atmega328 microcontroller. When greater GPIO and peripheral counts are required, users often opt for the Arduino Mega boards based on the Atmega2560. This chip is equipped with many of the required capabilities for this project, but runs at a clock speed too slow to carry out the MIR functions.

The Arduino Due board is similar to the Mega, but utilises an Atmel SAM3X8E microcontroller which is based on an ARM Cortex-M3 core running at 84 MHz. The Due has sufficient UARTs and GPIOs, DMA (Direct Memory Access), and a 12 bit ADC that can be clocked up to 1 MSPS (MegaSample Per Second), more than enough to sample 21 channels of audio at 44.1 kHz. Though 12 bit resolution is below that of standard 16 bit ‘CD quality’ audio, since the digitised signal is purely for analytical purposes and will not be output for listening, 12 bits is deemed sufficient. The onboard ADC is multiplexed to 16 channels, with twelve of those channels accessible on the Due.

The Arduino Due fulfills most of the specifications of the instrument, except EEPROM and ADC channel count. However, the form factor of the Due makes it convenient to mount multiple Dues vertically while still maintaining access to the onboard headers. This means that by basing the system on a pair of Arduino Dues, each processing ten or eleven channels of audio, and attaching external Inter-Integrated Circuit (I²C) EEPROMs, the system meets the requirements set.

3.4 The Closed-Loop Upgrades

Having surveyed the available hardware options for upgrading the Robotic Glockenspiel with closed-loop control and selected solutions that deliver satisfactory specifications, the process of designing the upgrades begins. This section illustrates the layout of the various subsystems of the instrument starting with the external componentry and proceeding to the internal circuitry. Following that, it describes in detail the design of the schematics, custom circuit boards, and other elements of the instrument.

3.4.1 The Panel

To achieve the previously stated goals of making the instrument easy to use and efficient to set up in performance environments, it is important to ensure that it is not complex to operate. It should have an interface that is comparable to electronic instruments such as commercial desktop synthesisers, with which the intended userbase will be familiar. For this reason, the original back panel of the instrument was replaced with an entirely new black

laser-cut panel as shown in Figure 3.3, and contains very similar interfacing componentry to a modern piece of audio equipment.

FIGURE 3.3: The Closed-Loop Robotic Glockenspiel's back panel. From the left, MIDI Input, USB I/O, IEC socket with fuse and illuminated switch, 1/4" audio output, and volume knob.

A standard IEC socket is utilised for providing the instrument with mains power, and it is specified for the 230 V / 50 Hz power standard used in New Zealand and many other countries. The housing of the IEC socket also contains an illuminating power switch and a 3 A fuse to protect the unit in short circuit or other over-current situations.

The left side of the panel is concerned with communications, housing a 5-pin DIN socket for MIDI input and a full-sized USB-B socket which is used for USB-MIDI input and output. The USB socket is mounted internally to a breakout board which is populated with an Atmel Atmega16u2 USB-capable microcontroller. This microcontroller is loaded with the HIDuino firmware [21], which converts USB-MIDI messages to standard serial messages inside the instrument. The firmware has been recompiled so as to allow the instrument to be recognised as a Robotic Glockenspiel by a USB host device.

The right hand side of the panel houses a standard unbalanced 1/4" socket

that provides the audio output from the system. A volume knob is also provided to attenuate the level of the analogue signal, and is inserted into the audio path before the final gain stage.

FIGURE 3.4: A diagram of the pickup mechanism (left side) and the striking mechanism (right side) of the instrument.

3.4.2 The Mechanics

Each individual key of the glockenspiel is paired with a tubular linear push-type solenoid, and there are a number of design considerations that must be made in order to ensure that these actuators create a pleasant sound, free of extraneous acoustic noise. Striking the key with the solenoid's bare metal rod produces a harsh sound in which the higher harmonics dominate the output. A wide range of materials including wood, rubber, and felt are commonly used with glockenspiel beaters, and after trialling several different striking materials it was decided that wooden balls provided the most agreeable timbre for this instrument. The solenoids are positioned so as to strike near the centre of the glockenspiel keys, and the keys are close enough to the solenoid that striking them will not allow the solenoid's plunger to reach its limit, avoiding the extraneous clicking sound that would occur as the plunger hits the end of its travel.

Another potential source of unwanted acoustic noise present in the instrument is on the keys themselves. As traditional ways of playing the glockenspiel involve striking the keys from above, these instruments are not designed to cope with the keys being struck from below. As a result, when solenoids strike the keys from below, the keys tend to travel upwards until they come into contact with the heads of the screws that are holding them in place. Even though the screw-heads are protected by metal spacers and thin felt rings, this can create an audible clicking noise between the spacers and the screw heads which is most pronounced at high velocities. To decrease the significance of this source of noise, 42 1 mm thick rubber o-rings were fitted, one on each screw, between the screw head and the metal spacer beneath. This creates a buffer which does not let the two pieces of metal make contact, decreasing the extraneous noise significantly.

Though some similar systems such as Raes' <Vibi> [133] utilise the combined forces of gravity and that of the solenoid's plunger bouncing off the key to propel it back to its resting position after a strike, in the Closed-Loop Robotic Glockenspiel a spring is attached to the plunger to expedite this action. This provides an increase to the strike rate. To avoid the extraneous noise associated with the solenoid reaching its extreme on the return stroke, strips of acoustic foam are mounted inside the enclosure underneath the solenoid's plungers. This foam prevents the plunger from reaching the bottom end of its travel and avoids any clicking sound that would occur should the plunger reach its limit.

As shown in Figure 3.4, the neodymium magnets are attached to the keys near their central extremes using Araldite Super Strength glue. As the nodal points of the keys are where they are attached to the instrument's frame, the extremes of the keys are where the highest level of vibration occurs. Beneath each of these magnets, an electromagnetic coil is positioned close enough to pick up the vibrations of the magnet, but not so close as to inhibit the key's free movement with its magnetic attraction. Since there is no wiring attached to the keys directly, they are able to move freely, and the neodymium magnets have a low enough mass that any potential detuning of the keys is imperceptible.

3.4.3 The Electronics

The electronics of the instrument comprise three main sections: the power supply, the digital circuitry which includes microcontrollers and solenoid-controlling Metal Oxide Semiconductor Field Effect Transistors (MOSFETs), and the analogue audio circuitry. The design and construction of these three sections are discussed in turn below.

Power

The instrument receives its mains power through the IEC socket, and is switched on by toggling the illuminating power switch. The mains power is routed through an internal fuse and is passed on to the primary winding of a power transformer inside the unit. A toroidal transformer was chosen rather than a typical EI model due to their high power efficiency, low profile, the required low levels of flux-leakage noise due to being mounted in close proximity to analogue audio circuits [100], and low acoustic noise [163]. In order to prevent anything from coming into physical contact with the mains and potentially shorting it, the mains voltage section of the unit is physically separated from the rest of the internals with a 3D printed barrier that entirely encompasses the internal side of the IEC fixture, leaving very small gaps for the primary wires of the transformer to enter.

The toroidal transformer provides dual secondary voltages of 60 V AC, which is used to drive the the solenoid supplies, and 9 V AC, which is used to drive both the analogue audio and digital circuitry. Similar to Raes' voltage selection explained in [126], 60 V was chosen as a suitable voltage level for the solenoids. Though the solenoids are rated for 24 V DC, this higher-than-specified voltage is permissible due to the very short on times used, that result in relatively little heat generated. These short pulses of high voltage allow the solenoids to strike the keys at a higher velocity without pressing on the keys and dampening their sound.

In order for these alternating current supplies to be of use in the instrument's circuitry, they must be converted to direct current. The single-layer custom circuit board shown in Figure 3.5 was created to carry out this function on both secondary outputs from the transformer. For the solenoid supply shown in Figure 3.6, the board takes in the 60 V AC line, passes it through four high current 1N5404 diodes arranged in a bridge rectifier configuration, and then

FIGURE 3.5: A rendering of the power supply circuit board of the Closed-Loop Robotic Glockenspiel.

utilises a pair of large electrolytic capacitors to provide the bulk capacitance necessary to smooth the rectified power.

FIGURE 3.6: The solenoid power supply circuit.

The circuitry used for the 9 V AC line is shown in Figure 3.7. A pair of 1N4007 diodes of opposite polarities are used to separate the positive and negative sides of the alternating current and the voltages are smoothed with bulk capacitance before being passed to the 7809 and 7909 linear positive and negative voltage regulators respectively. These regulators provide very stable bipolar ± 9 V supply rails for use with the op-amp based audio circuitry, and extra 1N4007 diodes protect the outputs from any accidental reverse polarity connections. The +9 V line also feeds a 7805 linear voltage regulator which outputs the stable 5 V required to drive the digital electronics. The board is also equipped with a power indicator LED, which is attached to the negative supply rail, as that rail is loaded to a lesser extent than the positive one.

Digital Control

As discussed in Section 3.3.3, both the control logic and DSP algorithms of the prototype are implemented on a pair of Arduino Due microcontroller boards. This means that the processing and control functions are spread between the two boards, with one Due responsible for ten keys, and the other for eleven. To control both boards simultaneously, the MIDI and USB-MIDI input connections are forked to UART ports on both boards. In order to send data back to a host computer or other device, one Due has a direct connection to the panel, while the other Due sends its information to the first, to be passed through in a daisy-chain configuration.

FIGURE 3.7: The circuit which supplies power to the analogue audio and digital control circuitry.

Programming the Dues while the instrument is assembled is achieved by removing the small front panel of the instrument, which leaves the micro-USB ports of both boards accessible. In order to simplify the programming process, the firmware is written as a single project that encompasses the functionality of both boards together, with a single software flag used to indicate which board is currently being programmed.

Control signals consisting of short pulses to activate the solenoids are transmitted from the microcontroller boards to the solenoid control board. A rendering of the board is presented in Figure 3.8 without cabling so that its componentry can be seen clearly. This simple board implements the same MOSFET based solenoid control circuit outlined in [126] for each of the 21 keys of the instrument.

Audio Circuitry

A rendering of the analogue audio processing board of the instrument is presented in Figure 3.9 without cabling in order to provide a clear view of the componentry. The board receives signals from each of the 21 pickup coils positioned under the glockenspiel keys and prepares the signals for the digitisation stage of the process. The form factor of the board was designed so that it could be positioned vertically within the instrument's case, attached to the floor of the case via a pair of aluminium L-brackets. Since this is the first instrument of this type constructed, through-hole componentry was utilised to allow practical experimentation during the development process.

FIGURE 3.8: A rendering of the MOSFET-based solenoid control circuit board.

FIGURE 3.9: A rendering of the analogue audio conditioning and mixing circuit board.

Figure 3.10 shows the schematic of one of the analogue audio processing channels. A coil is connected at the input stage, with its current flowing through R_1 to arrive at the TL074 quad low-noise JFET-input operational amplifier. The amplifier is powered at $\pm 9\text{ V}$ in a bipolar configuration. Negative feedback is achieved with R_3 , and the gain coefficient is dictated by the position of the onboard trimmer R_2 , from a gain of approximately $-10:1$ to $-\infty:1$.

After amplification, the signal is passed through a $10\text{ }\mu\text{F}$ electrolytic capacitor in order to reduce DC offset, AC coupling the signal. Following that, R_4 and C_2 form a low-pass filter with a cut off frequency of approximately 17 kHz in order to attenuate any extraneous high frequency noise that may be present. Finally, the signal is biased to a virtual ground at 1.65 V with R_5 and R_6 prior to being sent to the microcontroller boards for analogue to digital conversion.

The output pins of each of the amplifier circuits are also connected to a separate TL074 based summing amplifier circuit before the biasing stage. The summing amplifier is equipped with an internal trim-pot to set the maximum output voltage of the mix-down. The signal is routed through the panel-mounted volume potentiometer, and is buffered by a final op-amp circuit in a voltage follower configuration before being output to the $1/4"$ panel-mounted TS socket.

FIGURE 3.10: The schematic of one of the 21 channels of audio conditioning circuitry.

In Figure 3.11, an exploded rendering of the Closed-Loop Robotic Glockenspiel is presented with one panel made transparent in order to clearly illustrate the layout of the internal componentry of the instrument. From the right, the main panel with all connections and controls is shown, with the 3D printed mains power protection bracket shown to the left of that in green. The toroidal transformer is placed to the left of the mains connections, with the power supply board positioned to the left of that. The pair of Arduino Dues are shown to the left of the power supply board, with the MOSFET board split into two pieces due to space constraints shown at the leftmost side of the image. The audio processing circuit board is vertically positioned behind the other circuitry, shown in blue. By integrating all of these subsystems in a single enclosure, the complexity of the logistics of transporting, setting up, and performing with the instrument are greatly reduced compared with other musical robots with closed-loop feedback functionality, which often require separate control hardware, cabling, and generally require a personal computer for calibration. This fulfils the portability design goals for the instrument stated in Section 1.3.

FIGURE 3.11: An exploded render showing the internals of the Closed-Loop Robotic Glockenspiel (top), and a photograph of the internals of the instrument (bottom).

3.4.4 The Firmware

A single firmware in the form of an Arduino sketch is written to both of the microcontrollers, with a flag which assigns each of the two boards to one of the two groups of solenoids and pickup coils. The firmware carries out multiple functions:

- It manages communications with the outside world via MIDI In and Out signals.
- It produces the pulses used to control the solenoids which strike the keys of the instrument.
- It implements audio acquisition ADC routines and Digital to Analogue Converter (DAC) signal output routines for debugging.
- It carries out continuous onset, latency, and amplitude detection on the audio received from the coils, and continuously adjusts the instrument's performance based on these measurements and analyses.
- It loads and saves calibrations and configurations to and from the external EEPROMs.

The implementation of each of these functions is described in the following subsections.

MIDI I/O

As shown in Figure 3.12, a pair of 5-pin DIN connectors on the back panel of the instrument provide standard MIDI input and output connections so that the instrument can be used directly with the large variety of MIDI controllers, sequencers, and other equipment built around the protocol. The USB MIDI connection is also provided so that the instrument may be played directly from a USB host such as a personal computer, laptop, tablet, or smartphone, with bidirectional communication provided by a single USB cable without requiring any additional MIDI interfaces. One, two, or all three of these connections may be utilised at once, depending on the performance application.

The input from the MIDI or USB MIDI is utilised predominantly to receive MIDI Note On commands with velocity information from a controller or sequencer to both Due's simultaneously. They are also sensitive to a few specific Program Change numbers for optional configuration. The MIDI and

USB MIDI outputs are used for sending MIDI Note On and Off messages from the instrument as well as potentially sending out information about the performance of the instrument for debugging purposes. These functions are described in greater detail in Section 3.4.5.

FIGURE 3.12: Diagram of the serial communications scheme of the Closed-Loop Robotic Glockenspiel.

The firmware of the instrument handles these MIDI communications via a series of callback functions provided by version four of the Arduino MIDI library¹. The instrument ignores Note Off commands as they are not necessary for a glockenspiel without controllable sustain, and acts on Note On messages that correspond with the list of keys which that Due is responsible for actuating. When a MIDI command arrives that must be acted upon and a note is not currently being struck by that solenoid, the corresponding GPIO for that note is driven high. The channel is then marked as busy, taking record of the time of the strike with the `micros()` function, which returns a number corresponding to the number of microseconds that have elapsed since startup. Referring to the velocity calibration tables, the routine extracts the pulse width that corresponds with the received MIDI Note On velocity value, and the firmware's main loop regularly polls the `micros()` value, bringing the solenoid-controlling GPIOs low when their pulse-width times have elapsed.

At the point of triggering the note, the firmware also checks whether or not that key has recently been played by an instrument or a human, and only triggers audio analysis to be conducted on strikes played by the robot with

¹https://github.com/FortySevenEffects/arduino_midi_library/

the key at rest. This prevents automatic continuous calibration values from being skewed by strikes that have accumulated higher vibrating amplitudes as a result of being struck in quick succession. The set time constant that establishes that a note is at rest is equal to the amount of time the sound of a strongly struck key takes to decay to silence. A value of approximately three seconds was arrived at experimentally. The analysis triggering routine also cancels and discards any ongoing analyses that are being undertaken on notes that have been interrupted by another strike by the robot or a human.

Analogue I/O

Each of the Arduino Due boards are responsible for taking in ten or eleven channels of audio for analysis, and outputting one or two channels of audio internally for development and debugging purposes. The Due's onboard SAM3X8E microcontroller is equipped with an ADC peripheral that is capable of sampling at a resolution of 12 bits, at a sample rate of up to 1 MHz. It also has an integrated 16 channel multiplexer, twelve channels of which are accessible on the Due board.

The built-in `analogRead()` function in the Arduino language provides a convenient method of sampling these inputs, however, its implementation is such that it blocks the processor while it is being conducted, and it takes approximately 40 μ s to carry out a single read of one channel. The target sample rate of 44.1 kHz translates to 22.68 μ s between samples, and all ten or eleven channels must be sampled within this time frame. This means that in order to achieve the required performance with the board, it is necessary to bypass Arduino's built-in `analogRead()` function and configure the ADC manually.

Upon startup of the board, a routine is run that configures the SAM3X8E's ADC peripheral, before enabling it, to start signal acquisition. The routine begins by writing to the ADC's Mode Register to set the unit to 12 bit resolution mode, ensure that the sleep mode is off, and disable free-running mode, as the sample rate is specified. The ADC peripheral clock is then set to 21 MHz, the highest compatible value given the master clock of 84 MHz and a maximum ADC clock of 22 MHz, and the startup, settling, tracking, and transfer time periods are set to their maximum values for increased stability of input. Following this, the ADC's user sequence mode is activated, and the channel sequence registers are loaded with the specific channels of

the ADC multiplexer that are to be read. Those channels are then enabled, with all gains set to unity, with all interrupts disabled, and all inputs set to single-ended configuration.

Following the ADC configuration, the Peripheral DMA Controller is configured for ten or eleven transfers per call, and a pointer is set up to automatically populate defined tables in Static Random Access Memory (SRAM) with the input samples from the ADC. A 32 bit timer is configured to trigger an interrupt routine 44100 times per second to undertake all of the audio-rate processing of the system. At the beginning of each activation of this interrupt, a single command triggers the ADC sequence to sample each of the audio channels in turn, with the DMA controller populating the following sample's values automatically. This occurs independently of the main processor, enabling it to carry out its MIR analyses and other functions in parallel.

While not required for the operation of the instrument, the on-board two channel DAC peripheral provided by the SAM3X8E is useful for development and debugging purposes, and is also configured on startup. The DAC is set to free running mode, with an output requiring a write to the DAC's Mode Register to specify the channel, and the Conversion Data Register populated with the desired value. By passing audio channels from the ADC through to the DAC it is possible to monitor specific signals, and observe the results of analyses via an oscilloscope.

Onset Detection

Onset detection is a technique utilised in order to determine the point in time when the start of a musical event occurs. It is commonly undertaken in musical systems for triggering musical events, score following, conducting analyses on the timing of musical events, and as a prerequisite for many other processes such as beat, rhythm, tempo, and melody detection among others.

Before conducting amplitude and latency analyses in the Closed-Loop Robotic Glockenspiel, it is necessary to determine when the strike took place as precisely as possible using onset detection. There are a number of different algorithms that are used to detect onset summarised in [22], and their effectiveness varies greatly depending on the source material. Because of the unique nature of the Closed-Loop Robotic Glockenspiel, a custom onset detection scheme was devised, described henceforth.

The process of onset detection in this instrument is significantly simplified in comparison to general purpose applications due to the fact that several pieces of information that are generally unknown in a human performance are pre-determined in a robotic one. Firstly, the approximate point in time that an onset created by the robot is expected to occur is established. This enables the continuous calibration algorithms to ignore any incidental onsets that could be detected outside of a small window of time. Secondly, since the glockenspiel is a percussion instrument, the envelope of each strike consists of a short, sharp attack, and a slow and predictable release. This means that the algorithm need not account for sounds with slow attack times or unpredictable envelopes.

Since the audio transduction method makes use of magnetism rather than sound pressure waves, it is not necessary to account for the possibility of unintended acoustic sounds within the proximity of the instrument interfering with the audio feed. Also, because a coil is assigned individually to each key of the instrument, the fundamental frequency, and therefore, the period of each of the expected notes is also a known constant.

FIGURE 3.13: Onset and amplitude detection.

With these assumptions made and referring to Figure 3.13, the onset detection works as follows. After a MIDI Note On message is received and a pulse is sent to the MOSFET controlling the referenced solenoid, a short impulse of interference is present in the audio signal of the note's pickup coil, as shown on the left of Figure 3.13 and noted in [80]. In order to prevent this impulse

from interfering with the onset detection, the algorithm does not begin listening for onsets for an amount of time equal to the width of the pulse applied, plus a 1.5 ms buffer.

When the machine begins listening for an onset, a circular buffer maintains an averaged signal that acts as a software virtual ground from which to make amplitude comparisons. With reference to this average, the algorithm marks negative-going zero-crossings in the audio stream. When three subsequent zero-crossings have been observed that are each separated by the expected period of the played note within a hard-coded margin of error, an onset is recorded as having taken place two periods earlier. The latency of the strike of that particular key, and at that particular velocity is then recorded as the time from the note being triggered to this onset time. If an onset has not been detected within 50 ms of the note being triggered, a time-out occurs and it is established that the solenoid did not make contact with the key. This frequency dependent approach to robotic pitched percussion onset detection allows even very soft strikes just above the noise floor to be detected, and has relatively low CPU overhead.

The averaging and onset detection algorithms run in the high priority audio routine that is carried out for each sample in real time at 44.1 kHz. The flow diagram presented in Figure 3.14 outlines the processes that are carried out in this routine, starting with retrieving the new audio sample, maintaining averaging buffers, listening for and differentiating between human and machine actuated strikes, and implementing the zero-crossing-based onset detector. When machine-activated onsets are detected, the amplitude and latency measurement routines which run at a lower priority and are described in the following subsections are activated.

Amplitude Detection

Amplitude detection is the process of determining the voltage swing of a musical input at a certain point in time. It can be used to determine the comparative average loudness of an actuation of an instrument, or to generate a curve that illustrates the amplitude envelope of a sound throughout its attack, sustain, and release stages. This technique is often utilised in both analogue and digital domains for creating envelope followers. These are often used to apply one sound's amplitude envelope to a characteristic of another

FIGURE 3.14: Flow diagram of the high priority audio rate audio analysis routine.

sound, to determine the strength of the actuation of an instrument for providing velocity output messages, and for a range of other analyses.

Because of the predictable envelope of a glockenspiel strike, the most important amplitude value to record is that of the wave around the strike's onset, as it is generally possible to infer the curve of the note's decay from this value. To retrieve this information, each time the onset detector algorithm marks the first in a potential series of negative zero-crossings, the amplitude detector begins recording the input audio into a dedicated buffer. When three correctly spaced zero-crossings have occurred and the onset has been successfully detected, the amplitude detector scans its buffer of the previous two wavelengths, and records the minimum and maximum values in the buffer. The difference between these two points is recorded as the signal's amplitude.

Continuous Calibration

At startup, the instrument initialises the velocity-to-pulse-width and velocity-to-latency tables with naive data that may not correspond with the current ideal amplitude values of the instrument. MIDI Program Change commands are used to set the calibration settings of the instrument. Different values of Program Change instruct the instrument to turn automatic continuous calibration on or off, load the previously saved calibration configuration from the external I²C EEPROMs attached to the Arduino Due boards, and save the current calibration to the EEPROMs.

When continuous calibration mode is on, after each key is triggered and the MIR algorithms have recorded latency and amplitude values for the key at a certain pulse width, the continuous calibration routine is run in the background. As shown in Figure 3.15, this routine compares the measured amplitude value with a table of ideal amplitude values vs MIDI velocities, and writes the pulse width used to its ideal position. It also adjusts the tested velocity towards its ideal value. It then conducts an interpolation between the new value and the previous recorded value, and between the new value and the adjusted value. This brings the current array of amplitudes more in line with the desired response graph which represents 126 equal steps in rising amplitude.

The algorithm also creates a comprehensive record of the latency of each key, at each velocity. In the case that a time-out has occurred and the no-contact

Dynamics Calibration

FIGURE 3.15: Flow diagram of the lower priority dynamics and latency calibration routines.

flag has been set, the algorithm simply adds a small length to the pulse width at that velocity value and interpolates from that point. A value of 0.5 ms for this extra length was arrived at experimentally and is designed to give a significant boost to the velocity of the note, reflecting the fact that it is more desirable for a note to sound slightly too loud than for it to be missed altogether.

The fact that lower velocity strikes take longer to reach the instrument's keys than higher velocity strikes often creates timing discrepancies with robotic musical instruments. While this issue cannot be avoided in live performance situations where the arrival of each individual note cannot be predicted, for performances that involve sequenced material, a latency compensation system is implemented. To activate the instrument's latency compensation, the timing of the sequencer's MIDI track is brought forward and the instrument's latency compensation feature is enabled with a Program Change command.

A value of 50 ms is chosen for the latency compensation as it is equal to the slowest measured strike plus a safety margin. The instrument then uses the information it has gathered about the latency of each velocity of each note to provide a uniform level of latency to every sequenced strike. It achieves this by playing softer hits with higher latency earlier and harder hits with lower latency later. As a result, it is possible to make improvements regarding the tightness of the instrument's timing, even in compositions with great dynamic variety. Since the instrument is constantly updating these latency values, the latency compensation also adjusts by itself dynamically.

This continuous calibration enables users to conduct performances with the instrument without lengthy manual calibration and without worrying that changing conditions during play will have a negative effect on the instrument's performance. If the instrument has just been installed in a new environment, a play-through of the intended composition during the sound-check will in most cases provide adequate opportunity for the instrument to automatically calibrate itself, but a more thorough calibration routine can also be conducted. The effectiveness of the described calibration routines are evaluated in Section 3.5.

3.4.5 Other Features

As a result of the closed-loop upgrades to the robotic glockenspiel, in addition to the improvements to the reliability, ease of use, and dynamic range, several extra features are made possible by the additional hardware. The audio output from the instrument allows it to be amplified in performance contexts without the extra equipment, setup time, acoustic bleed, and feedback problems inherent in using microphones. This also means that the audio feed can be routed through standard audio effects, or to a PC for creative digital signal processing. The extra performance information available to musical robots means that with creative programming, parameters of audio effects applied to the output signal can be perfectly synchronised with the sequence that is triggering the actuation of the keys. This level of synchronisation can provide many new creative possibilities and avenues for musical expression as utilised in the composition described in Section 3.7.1.

Another feature that this system adds to the instrument is the capability for the instrument itself to be used as a MIDI controller. By detecting onsets that were not caused by the solenoids, the instrument can extract hit and velocity information and output MIDI Note On commands followed by Note Off commands sent after a fixed delay. By processing the output MIDI signals externally, it is possible to apply MIDI effects and processing and send them back to the instrument. This creates an interactive robotic instrument, also defined as a ‘cooperative musical machine’ by Barton [6], and opens up performance possibilities in a number of ways. This feature was utilised in the installation piece described in Section 3.7.2.

3.5 Evaluation of Functionality

In order to establish whether the closed-loop audio-based continuous calibration systems implemented in this upgrade have been successful in improving the specifications of this instrument, it is necessary to carry out tests of the calibrated parameters. The two primary characteristics that are tested are dynamics and latency. Fortunately, as the upgrades have left the instrument’s mechanics unchanged from its earlier open-loop state, it is possible to evaluate both open and closed-loop versions, in order to make comparisons between their characteristics.

Throughout these evaluations, RMS is utilised as an objective measure for comparison between the results of tests done on the same instrument, but is not useful for comparing loudnesses of different instruments, as the measurement depends on a fixed laboratory configuration, and different instruments naturally have differing dynamic capabilities. An imaginary ideal dynamics graph for an instrument would consist of a straight line that starts from zero at MIDI value zero, and rises in equal increments for each step, arriving at the maximum possible loudness of the instrument at value 127. By comparing the presented graphs with this, a sense of the effectiveness of the results in relation to this ideal can be attained.

3.5.1 Dynamics

The dynamic response of the keys of the Robotic Glockenspiel are tested by positioning an SM57 dynamic cardioid microphone, oriented 90 degrees perpendicular to the key under test, pointing at the centre of the key with its diaphragm at a distance of 100 mm. A high, middle, and low key are selected for measurement to achieve a more comprehensive perspective on the range of responses, and each measurement is taken ten times to provide standard deviation values.

The recordings are made at a fixed level of gain by a MOTU 828 mk3 Hybrid audio interface at a 44.1 kHz sample rate and 16 bit resolution, and triggered by the same device's MIDI interface. 2-second-long samples of each strike are rendered via a digital audio workstation (DAW), and are analysed by a custom script written in the ChuckK musical programming language. This script reads through each sample, taking spectral RMS measurements at an FFT size of 1024, adding each RMS value together to provide a total RMS value for the sample.

Open-loop Response

The glockenspiel is commanded to perform strikes powered by 20 different pulse widths spanning from not touching the key at all to striking it excessively hard, in 50 μ s increments. The resulting graphs of RMS loudnesses measured at these pulse widths for the top, middle, and bottom notes are shown in Figure 3.16.

FIGURE 3.16: Loudness vs pulse widths of three notes of the Robotic Glockenspiel without automatic calibration (first trial).

Each graph can be separated into three main sections: low pulse widths, ideal range, and excessively hard. The low pulse width section of each of the three graphs consists of values of very low RMS where the beater of the solenoid was not provided with enough power to make contact with the instrument's key. The low pulse width section of the top graph of Figure 3.16 is the range from 1200 to 1350 μs . At the top of this range is a higher value that often also has a relatively high standard deviation, shown by the error bars. This occurs when the tested value is only making contact with the key some of the time, partially the result of inaccuracies in the pulse widths sent by the Arduino Due.

The ideal range section in the graph is the section above the low pulse widths section, consisting of a set of values that make a relatively straight line of increasing loudness values for increasing pulse widths. The ideal range of Figure 3.16's top graph is from 1450 to 1750 μs . This range also tends to have the lowest standard deviation values of the graph. It is labeled the 'ideal range', because this is the range of values that the automatic calibration routine must identify in order to implement a striking calibration with reliable dynamics.

The excessively hard range is the top section of the graphs where the trend of ideal values is broken and the dynamics of the strikes no longer correspond predictably to increases in pulse width. In Figure 3.16, the excessively hard sections of the top and middle notes start at 1800 μs and show very high standard deviations. This is a result of some of the strikes being quickly dampened by being pressed too hard into the stoppers above the keys, or by the beaters touching the keys for too long or more than once before they fall back to their resting positions. The waveform of one of these dampened strikes is shown on the bottom of Figure 3.17, and compared with a normal strike shown on the top. This kind of unpredictability in strikes is very undesirable in performance, so automatic calibration should exclude these values. The bottom note in Figure 3.16 was not affected by this quick dampening issue, so the standard deviations did not rise significantly in its excessively hard section. Despite this, these values remain highly non-linear and further investigation into the precise causes of this would require a high frame rate motion capture study that is beyond the scope of this thesis.

Section 3.2 mentions that one of the issues experienced with the open-loop Robotic Glockenspiel is that even with precise manual calibration, the device would fall out of calibration with use. In order to demonstrate this issue, the

FIGURE 3.17: Waveform view of example uninterrupted (top) and interrupted (bottom) strikes, over a period of approximately 20 ms.

above trial is carried out two additional times. Before the second trial, the device was subjected to conditions intended to simulate packing up, transporting, and setting up the instrument again. Measurements of the unit's dynamic curves were taken on the same three notes, and the results are shown in Figure 3.18.

The graphs of the second trial resemble those of the first trial relatively closely. The most noticeable differences are the “excessively hard” sections of the graphs, where the shapes of the lines are quite dissimilar. This emphasizes the unpredictable nature of the response of the instrument at long pulse widths. Before the third trial, the unit was subjected to conditions intended to simulate maintenance by removing the top panel and structural screws and reassembling the unit. The result is that screws and parts of the unit likely have slightly different tensions and positions to the first and second trials. The results of the third trial are presented in Figure 3.19.

The “excessively hard” sections of the third trial's graphs exhibit quite dissimilar shapes to the first and second trials. In addition, it is observed that the low pulse width sections of each of the three notes have grown larger by varying amounts, pushing the ideal range into higher pulse-width values. This demonstrates the possibility of calibration drifting in the instrument, especially when the hardware is adjusted or maintained, and underscores the necessity of a closed-loop automatic calibration system for keeping the instrument within its ideal range.

Closed-loop Response

Following the three trials of the instrument's dynamic response without calibration, an evaluation of the automatic audio-based closed-loop dynamics calibration routines described in Section 3.4.4 is undertaken. The instrument is powered on with the default pulse width to MIDI velocity values that linearly span the ranges shown in Figures 3.16, 3.18, and 3.19. That is, a straight line of values between 1200 μs and 2150 μs for the high and middle notes, and between 1450 μs and 2400 μs for the bottom note.

The instrument is then sent sequences of MIDI note commands that simulate a typical initialisation routine or performance sound check. These sequences include five iterations of each key's lowest and highest velocity values, and a strike at each of the graphed velocity points. Following this preparation, the

FIGURE 3.18: Loudness vs pulse widths of three notes of the Robotic Glockenspiel without automatic calibration (second trial).

FIGURE 3.19: Loudness vs pulse widths of three notes of the Robotic Glockenspiel without automatic calibration (third trial).

FIGURE 3.20: Graphs of Loudness (RMS) vs MIDI velocity for the Closed Loop Robotic Glockenspiel after automatic calibration.

instrument's automatic dynamic and latency calibration functions are deactivated remotely via MIDI Program Change commands. This ensures that the evaluations capture a snapshot of the instrument's calibration at that point in time and that it does not continue to be automatically adjusted while the measurements are being taken.

Twelve points in the MIDI velocity scale of each of the three notes are measured, with ten samples taken of each measurement to provide standard deviations. The results of these measurements are presented in Figure 3.20. It can be seen that the automatic calibration has eliminated the low pulse width and excessively hard ranges from the dynamic graphs, leaving the relatively straight lines from the ideal ranges of each of the keys. The standard deviations of the lowest dynamic points in the top and middle notes are somewhat higher than the rest of the points in the graphs, similar to the edge between the low pulse width and ideal ranges in the open-loop graphs. As the automatic calibration cannot improve the accuracy of the pulses themselves, the overall standard deviations are also very similar to those of the open-loop graphs. This situation is improved upon with the development of a Field-Programmable Gate Array (FPGA)-based controller board in Chapter 5.

3.5.2 Latency

In order to retrieve information about the latency of strikes before and after automatic calibration, it is necessary to take measurements of the time difference between sending a MIDI command for a strike and the solenoid beater making contact with a key. These measurements are carried out on the audio recorded during the dynamics testing, and each value is obtained by observing the difference in time on the DAW timeline between the MIDI Note On command and the onset of movement in the microphone's audio channel.

Open-loop Response

Figures 3.21, 3.22, and 3.23 present the latency analysis of the three trials outlined in Section 3.5.1. They each plot the latency values in milliseconds observed at a range of different pulse widths, for the three glockenspiel keys tested. The lower end of the pulse width ranges do not correspond with those of the dynamic response graphs because latency values cannot be measured for strikes that do not reach the instrument's key.

FIGURE 3.21: Latency values of pulse widths of three notes of the Robotic Glockenspiel without automatic calibration (first trial).

FIGURE 3.22: Latency values of pulse widths of three notes of the Robotic Glockenspiel without automatic calibration (second trial).

FIGURE 3.23: Latency values of pulse widths of three notes of the Robotic Glockenspiel without automatic calibration (third trial).

All of the graphs show a relatively linear relationship between the width of the pulses of current sent to the solenoids and the amount of time it takes them to strike their corresponding key. The left side of the graphs show the strikers' softest hits taking just under 20 ms to make contact, and the stronger strikes on the right sides taking approximately 13 ms. This approximately 7 ms difference in latency between soft and loud notes means that some musical passages such as those with a steady pulse that vary in dynamics may be audibly out of time when realised on an instrument such as this. A study done by Friberg and Sundberg measured a 'just noticeable difference' at 6 ms of perturbation for participants listening to isochronous sequences of sounds [33]. In the case of several notes being performed simultaneously in a chord, this value may be even lower.

Closed-loop Response

As the instrument is capable of automatically calibrating both dynamics and latency simultaneously, measurements of the instrument's latency calibration performance is done on the same audio used for the dynamic calibration described in Section 3.5.1. When latency calibration is active, the instrument maintains a table of latency values of the key throughout its velocity range, and adds an extra delay to each strike so as to create an even latency over the dynamic range, as described in Section 3.4.4.

Figure 3.24 shows the latency of strikes of the three tested notes while latency equalisation is activated on the instrument, and latency compensation is activated on the DAW. All three of the graphs show values over their range that are near the ideal latency of zero, with only slight standard deviations of 0.69 ms, 0.79 ms, and 0.24 ms for the top, middle, and bottom notes respectively. This demonstrates that the closed-loop audio based automatic latency calibration and compensation works as intended, providing significant reductions of perceived latency in prepared musical sequences.

3.6 Practical Issues Encountered

Since this iteration of the Closed-Loop Robotic Glockenspiel is the first example of such an instrument, it was created using prototyping microcontroller boards and changeable through-hole circuit boards, a workflow that has been

FIGURE 3.24: Graphs of Latency (ms) vs MIDI velocity for the Closed-Loop Robotic Glockenspiel after automatic calibration and DAW-based latency compensation.

useful over the course of developing the instrument. However, throughout the course of evaluating, composing, and performing with the instrument, a number of issues were observed that should be addressed in future iterations of this instrument or when designing any new closed-loop robotic musical instrument with similar features. Many of the issues identified below are improved upon in the robotic unpitched percussion system described in Chapters 4 and 5.

3.6.1 Noise

Though the Closed-Loop Robotic Glockenspiel makes use of a toroidal transformer in order to reduce the extent to which power supply noise interferes with the audio circuits, due to the size constraints inside the instrument, the audio circuits still pick up a small amount of ground hum interference. Because of the fact that the audio preparation and mixing circuit-board mixes all 21 channels at full volume and the ground hum is similarly present at each input, the resulting noise-floor of the output is high enough to be noticeable to an audience, especially in quiet concert settings.

In order to remedy this situation in concerts, external noise-cancelling processes were applied to the audio output of the glockenspiel as mentioned in Section 3.7.1, providing a much cleaner signal. A more integrated solution for the future would be to create a system that provides noise-gating functionality for each channel, so that channels that are not in use do not contribute to the noise-floor of the instrument.

The Arduino Due boards themselves also often create acoustic noise. Upon power-up, the boards frequently emit a high pitched ‘whine’ which centres around a frequency of 4 kHz. The audible noise is created by the power-conditioning circuitry on the board, and is also present as an interference on the power-rails. A simple power cycle of the instrument often remedies the issue, but a custom board that does not have such a problem should be created.

3.6.2 Accuracy

The error bars shown in the measurements taken in Section 3.5 demonstrate that the accuracy of repeated actuations of the instrument can be improved.

In order to pinpoint the cause of these inconsistencies, the actual pulses output by the SAM3X8Es to the solenoids were measured on an oscilloscope. This investigation showed that pulses that were intended to be of identical length often varied by up to 100 μs . As the Robotic Glockenspiel provides a range of velocities by applying very short pulses of voltage at several times the voltages the solenoids are rated for, this margin of error has observable effects on the performance of the instrument.

The Arduino Due boards in the Closed-Loop Robotic Glockenspiel are responsible for both the control of the solenoids and the embedded closed-loop DSP of the instrument. This multi-tasking creates a compromise where one of these functions takes priority in the processor, with the audio processing interrupt routine taking priority in this case over the communications and actuator management routines. The result is that as the MIR functions take up more of the CPU's overhead, less space is left for the processor to poll the state of the solenoid triggering outputs. This sometimes causes delays in bringing the solenoids' pulses down, often in multiples of 22.68 μs due to that being the audio processing sample period. In order to make further improvements in the accuracy of the instrument, a system of parallel processing where the solenoid control functionality is handled separately to the audio processing would likely be required.

3.6.3 Programming

One of the cited reasons for selecting the Arduino Due as the platform on which to implement the Closed-Loop Robotic Glockenspiel is that Arduino provides a comprehensive array of libraries and functions that expedite the processes of writing and testing firmware. However, throughout the course of creating the firmware for the instrument, it was found that the convenient functions and libraries provided by Arduino included significant compatibility overhead and suboptimal peripheral configuration, decreasing the performance of the routines to the point that they were impossible to implement. As a result, in order to carry out functions like sampling eleven channels of audio at 44.1 kHz, communicating with the EEPROM via I²C, controlling GPIO, utilising DMA, and other functions, it is necessary to create custom libraries and routines. This negates many of the previously perceived benefits

of utilising Arduino boards in the instrument, making the platform a less desirable option for this application. A more powerful processor may therefore be used.

3.7 Creative Outputs

In order to test the utility of the upgrades to the Closed-Loop Robotic Glockenspiel in a practical, real-world setting, two new works were created for, and exhibited with it: a composition, and an interactive installation. The following two subsections describe the conceptual aspects, technical configurations, and performance considerations of these works.

3.7.1 Composition: Broken Mirror

The first original piece written specifically for the Closed-Loop Robotic Glockenspiel is entitled *Broken Mirror*. In addition to taking advantage of the automatic continuous calibration and the higher level of integration of the instrument as a matter of course, the composition of *Broken Mirror* aimed to utilise the improved dynamics of the instrument, making use of lower velocity notes with more temporal space than would have been safe to use with the previous incarnation of the instrument.

The piece also makes use of the instrument's bleedless direct audio output by feeding the audio back to the digital sequencer controlling the instrument, and applying composed, precisely automated digital signal processing to the audio in tight synchronisation with the MIDI control stream. The sequencer then spatialises this audio around the performance area via a quadraphonic array of loudspeakers. The aim of this configuration is that the audience will hear both the acoustic sound from the actual instrument and the processed sound from the speaker array blended together in such a way that it is at times unclear where the instrument finishes and the speakers continue. Written in the form of a duet between the robotic musical instrument and real-time electronic processing, rather than simply instrument and effects, the speakers extend and 'mirror' the acoustic sound of the instrument, though that reflection is transformed, distorted, or 'broken' at times, leading the audience to question their perceptions and reflect on the sound of the instrument, the characteristics of the speakers, and the acoustic environment.

In addition to applying the technical advances made with the construction of the Closed-Loop Robotic Glockenspiel, *Broken Mirror* also explores the new artistic avenues offered by this novel configuration. One of these areas of exploration is the existing and potential relationships between musical robotic performance and the concept of ‘liveness’ as described by John Croft in his *Theses on Liveness* [20]. The piece raises questions such as: “When the performer is a musical robot, is it a ‘live performance’?”, and “When real-time processing in a piece is conducted by automation, is it a piece for ‘live electronics’?”. While this configuration is certainly *procedurally* live despite not involving humans in real time, the extent to which a piece such as this can be *aesthetically* live is explored through balancing consistent and varying spectromorphological processes.

Croft states that “if many perceptibly different inputs generate outputs with no pertinent differences, ... then the liveness is merely procedural and not aesthetic... At the other extreme, if the mapping is too explicit, too transparently one-to-one, the result is ... tedious”. The composition of *Broken Mirror* attempts to occupy some of the “surprisingly narrow” space between these two cases by utilising processing that preserves the detail of the live sound, with any peculiarities or imperfections mirrored in the electroacoustic extensions, while applying effectors like pitch shifting, bending, reverberation, echo, and reversing that maintain a rigid relation to the source material while providing a wide range of expressive possibilities.

It is clear that the electronic processings of *Broken Mirror* do not fit Croft’s *Backdrop*, *Accompanimental*, and *Environmental* paradigms of performance for live electronics. The closest assignable category seems to be *Resonsonal/proliferating*, which is when the electroacoustic sound has an antiphonal relationship to the live sound, responding to or extending it after it has occurred. However, the synchronised control of the instrument’s actions with the ‘live’ electronics allow gestural possibilities where the processed sound seems to lead the cause and effect relationship in the duet. This reversal of the cause and effect relationship between the acoustic and electroacoustic parts is difficult to achieve in a traditional live electronics context, as human beings lack the ability to precisely ‘look ahead’ into the future.

Another aspect of performance that *Broken Mirror* explores is the concept of gestural surrogacy in musical robotic contexts. Being a physical and acoustically audible musical instrument, the mimetic gestures performed by the Closed-Loop Robotic Glockenspiel in this composition inherently depict the

first level of gestural surrogacy as defined by Smalley in [25] despite the instrument's inability to display 'effort' and evoke proprioceptive responses. However, as the electroacoustic elements of the piece are explicitly presented as derivative of the sound emanating from the glockenspiel itself, the content is entirely source-bonded, though the electroacoustic gestures are not. This means that despite the lack of human performers, the loudspeaker-projected elements of the performance are only quasi-acousmatic, and comprehension of the cause and effect and structural relationships within the piece will depend on a listener's vocabulary of processes.

The composition of *Broken Mirror* also investigates the expressive possibilities of creating and applying various types of space throughout the piece. While enabling the manipulative aspects of the work, the quadraphonic loudspeaker parts also construct a number of different perspectives of space that would not be possible for the glockenspiel alone to create. Reinforcement of the acoustic sound creates a mediated intimacy that brings gestures from an intimate space to an ensemble space, and while Smalley states that "source-bonded sounds in particular... are space-bearers", the bleedless coil-based pickup system utilised in the instrument has the advantage that it does not interact with space in the way a microphone would, allowing the processing to discard source-bonded space, and apply its own transcendental space to the performance area. Performances of the piece in art gallery settings have also applied peripatetic space which allows audience members to manipulate their own vantage points while observing the piece.

While it is possible to examine many of the concepts outlined above without this specific instrument, the creation of the Closed-Loop Robotic Glockenspiel has made the exploration of these new performance paradigms more apparent and feasible than with other configurations of instruments. The architecture of the instrument pushes this young paradigm of contemporary music into new areas of accessibility and expressiveness, and makes the questions raised early in this section more pertinent than ever. It is hoped that its creation will aid in the development of a new performance practice and theoretical framework for contemporary music composition with musical robots and 'live' electronics. A description of the formation of the piece follows with reference to Appendix A, which contains a top level view of the entire eight minute sequence of *Broken Mirror* complete with descriptions of the functionality of each of the tracks.

FIGURE 3.25: An extract from the piano roll of the first track of *Broken Mirror*. Sequences of exponentially expanding and contracting rhythms of cluster chords with precisely maintained timing relationships (top), and corresponding velocity curves (bottom).

Broken Mirror explores a range of performance techniques that would be impractical for human performers to realise. This desire to create pieces that are impossible for a human to play has been a commonly occurring thread and frequent motivation for creating musical automata and robots throughout history as discussed in Section 2.2.2. The passage presented in Figure 3.25 includes sections with large numbers of notes being played simultaneously or staggered in complex timing relationships, with very fast playing at times, and a large dynamic range, evocative of much of Nancarrow's work for player piano. The closed-loop upgrades that have been implemented in the instrument help ensure that these dynamics are respected, that very quiet notes can be trusted to sound, and that precise timing relationships between high and low velocity notes are maintained. These improvements, enabled by the novel audio-based closed-loop architecture, have provided the instrument the means to act as a successful tool for exploring mechanically-realised expressiveness and nuanced dialogue between the physical and the virtual.

Given the instrument's ability to accurately realise quiet performances, the National Library of New Zealand was chosen as an appropriate venue for the debut performance of *Broken Mirror*. Subsequent performances have also taken place at Wellington's Adam Art Gallery, Lower Hutt's The Dowse Art Museum, and as part of Victoria University of Wellington's Creativity Week in the Hub. The piece was also performed at the 2017 Lilburn Trust NZSM Composers Competition, winning first prize in that event. An audio-visual recording of that performance is available at <https://vimeo.com/264917454>. Given the varying acoustic properties of each of these performance spaces, it is the composer's opinion that a small venue that seats an

audience of 50 to 100 in a room with a moderately reverberant quality is the ideal venue for this piece.

3.7.2 Installation: AutoReich

There are several paradigms in existence related to creating interactive musical robots. Firstly, a number of robotic musical instruments are constructed in such a way as to allow humans to play the instrument at the same time as mechanical actuations are taking place. Examples include the Robotic Taishogoto [85], the Yamaha Disklavier, Jason Jahnke's Super Cymbalum, and Scott Barton et al.'s Cyther [6]. Other robotic instruments such as Weinberg's Haile [168] and Krzyzaniak's Kiki [70] are examples of musical robots that interact with human musicians by parsing the sound of the human performers' playing, and generating complementary or responding musical passages. These can be seen as attempts to imbue robots with their own sense of musicianship so that they can perform together with humans as other players would.

Another example of an interactive musical robot is Bretan et al.'s Robotic Prosthesis for an Amputee Drummer [14]. This wearable robot arm allows a human drummer who is without a wrist to carry out both individual strikes and drum rolls. Other robotic instruments such as Gurevich's Stringtrees [42] and Sheffield and Gurevich's robotic percussion array [141], were built to distribute the sounding-object activation responsibilities between a number of electromechanical actuators and human musicians.

The Closed-Loop Robotic Glockenspiel shares characteristics with several of the aforementioned instruments in that the top surface of the glockenspiel is left clear for manual human playing, and an audio acquisition and analysis system is built in that is capable of detecting onsets. Though the circuitry inside the instrument is not powerful enough to carry out machine learning and other such algorithms, the manual onsets that it can detect may be output in the form of MIDI notes for external processing.

In order to initiate explorations of the instrument's capabilities as an interactive, cooperative robotic musical instrument, a short-term installation was presented during Victoria University of Wellington's Creativity Week that demonstrates some of this functionality. The installation utilises the MIDI Out capabilities of the instrument, inviting participants to use a wooden

beater to manually strike keys of the instrument. Upon receiving a strike, the Closed-Loop Robotic Glockenspiel sends a MIDI Note On to an attached laptop running the Max patch shown in Figure 3.26. The patch then responds to the notes played by sending related notes back to the instrument to be played robotically.

FIGURE 3.26: The software backend of AutoReich - a Max patch.

In contrast with many of the cooperative robots outlined above, conceptually, AutoReich seeks to place complete control of the glockenspiel and its actuations in the hands of the participant, creating a concrete sense of cause and effect that responds to stimuli in a predictable and controllable manner. By restricting its capabilities in this way, this encourages the participants to regard the device as a playable musical instrument in and of itself, with which performance skill can be acquired via practice and experience. This has the effect of maintaining the strengths of human performance, such as expressiveness, while robotically extending attributes such as dexterity and rhythm as described by Kemper [65].

The way the instrument responds to the participant corresponds with which of the modes are selected. Three MIDI effects that are commonly found in synthesisers and other electronic instruments were created to aid in participants' familiarity and comprehension of the interaction: Echo, Arpeggiator, and Step Sequencer. The simple user interface for selecting interaction modes and parameters shown in Figure 3.27 is made available to the participants on the attached laptop.

FIGURE 3.27: The optional software interface for AutoReich.

The Echo effect repeatedly replays the same note that the participants struck with an adjustable delay between each subsequent note and decreasing velocity values. The implementation of the echo by physical re-strikes of the glockenspiel key creates a natural acoustic sound that contrasts with the digital sampling methods that are ubiquitous in modern music. The Arpeggiator is similar to the echo effect except that the echoed notes play a broken chord starting from the note that was struck by the participant. The speed of the arpeggio and the individual intervals from the root note are adjustable.

The Step Sequencer is a monophonic eight step looping device similar in effect to analogue modular sequencers like the Roland System 100-m 182. Starting with a blank canvas, it loops at an adjustable speed through eight note spaces, and when participants strike a note, the currently active space is populated with that note. This allows the participant to quickly and intuitively create and alter note patterns with the instrument simply by hitting it, and the physicality of the playback encourages natural interaction with keys such as dampening them with fingers, and using other objects to alter the sound.

When utilising the Step Sequencer mode, an interactive sonic trajectory is formed that flows from silence, through a repeating sequence that begins as sparse individual notes and results in a full sequence of slowly morphing note loops as the participant carries out subsequent strikes on the instrument. This structure and hypnotic looping quality is reminiscent of works such as Steve Reich's *Six Marimbas*, inspiring the installation's name.

3.8 Conclusions

This chapter has presented the Closed-Loop Robotic Glockenspiel, a new robotic pitched-percussion instrument that brings musical robots a step closer towards achieving the repeatability, reliability, and usability of commercial electronic musical instruments. The embedded MIR system offers

continuous calibration of the instrument that is invisible to the end user, and playing it is no more complicated than plugging in power and MIDI. The closed-loop system also adds new features to the instrument such as audio output via a standard 1/4" jack, and MIDI Out functionality for interactive and computer synchronised performances.

The evaluations of the dynamics and latency of multiple notes and velocities of the instrument in its original state demonstrate the issues that the audio based closed-loop continuous calibration system is intended to address: incorrect and changeable dynamic scaling, and variable latencies for different velocities. The tests show that the system is capable of automatically scaling the relatively linear ideal range of pulse widths to each note's MIDI velocity, significantly reducing instances of missed notes and dampened notes, and maintaining this calibration over time. The tests also demonstrate that the automatic latency calibration and compensation algorithm coupled with MIDI delay compensation is capable of significantly reducing the standard deviation of latency in dynamic sequences.

The process of creating an original composition and installation piece for the Closed-Loop Robotic Glockenspiel has made it clear that the instrument enables many new musical possibilities in several different areas, and these works barely scratch the surface of these. The instrument enables the accurate implementation of compositions of super-human speed, complexity, and synchronisation, providing a platform for new types of contemporary music to be explored. It also joins the relatively new field of robotic cooperative instruments with which novel modes of interactivity and experimental musical expression can be achieved. As the first existing instrument of its kind, it provides a valuable and highly educational starting point from which to explore the possibilities of closed-loop audio feedback based control systems with polyphonic pitched percussion musical robots. It demonstrates encouraging results as to the technique's effectiveness and provides valuable information that informs the direction of future development of the area.

The research presented in this chapter is featured in the following publication:

J. Long, A. Kapur, and D.A. Carnegie. "The Closed-Loop Robotic Glockenspiel: Improving Musical Robots Using Embedded Musical Information Retrieval". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Brisbane, Australia: Griffith University, 2016

Chapter 4

A New Robotic Drum Kit: Mechanics

4.1 Overview

Unpitched percussion arrays are popular candidates for automation by musical roboticists. One reason for this is that they are capable of creating a large variety of different sounds using a lower number of actuators than related areas such as that of polyphonic pitched percussion instruments. These arrays are predominantly modular in nature, utilising a central interface with a number of channels to control a range of different instruments from a single sequencer. This allows the composer or installation artist to use a single set of control electronics for a large variety of different musical works or performance contexts simply by adding, subtracting, or substituting instruments connected to the system.

There are many existing examples of robotic unpitched percussion arrays including many of Trimpin's installations such as *Circumference* [31], a number of Meiwa Denki's works such as the *Ongen* series of arrays, Raes' <HAT> (Hit Any Thing) robot [127], Eric Singer's Modbots [144], and Kapur's MahaDevi-Bot [61], among many others. There are also several examples of academics using unpitched percussion arrays as points of entry to introduce students and newcomers to the musical robotics field [45, 58, 115, 141].

A specific unpitched percussion array that is frequently used in conjunction with these modular robotic control systems is the drum kit, with a standard kit containing a kick drum, snare drum, 10", 12", and 16" tom-toms, and hi-hat, crash, and ride cymbals. Examples of robotic drum kit players include

Auto-PRK [7], the Karmetik Breakbot [60], and Craig Newswanger's Percusibot [113]. The robotic drummers of the Compressorhead, Z-machines [147], and Captured! By Robots [162] bands are examples of partially anthropomorphic robots playing drum kits.

Despite their popularity, utilising arrays of varieties of percussion instruments presents particular challenges to composers that desire to create reliable, reproducible performances with these instruments without lengthy, difficult, and error-prone calibration procedures. Frank Barnes et al. of Compressorhead have written that "Programming and teaching the robots to perform the new songs takes several months. Fine-tuning and calibration is an ongoing process. Especially after robots have been serviced and parts have been exchanged, programming has to be adjusted." [4]. Ness et al. have also stated that "It is a time consuming and challenging process to setup robotic instruments" and that "[The calibration] process is frequently tedious and typically requires extensive manual intervention" [111].

These experiences of difficulty with setup and calibration of robotic unpitched percussion arrays are echoed by the author's own experiences with the Flea Market Ensemble described in [74]. Issues such as lengthy manual calibration, calibration drifting, drums and actuators shifting throughout performances, and lack of scalability were observed while working with the ensemble. In order to make improvements upon this situation, a new robotic drum kit with a closed-loop audio feedback based control system is developed and described in this and the following chapter.

This chapter begins with a study of the various types of robotic percussion actuation in use, recreating several existing designs and evaluating their performance in order to inform the design of the new system. The results of these evaluations are then discussed and the design and construction of a new, improved set of mechanical drum strikers is presented. Variations necessary for individual drums are discussed and the creation of a novel hi-hat movement system is then outlined. Chapter 5 then presents the electronic hardware that implements the embedded closed-loop audio feedback based system that controls the mechanical hardware described in this chapter along with descriptions of the firmware and the system's use in musical performances.

4.2 Evaluating Existing Percussion Mechanisms

In order to create a next generation robotic drumming system that is flexible, integrated, accurate, reliable, and expressive, it is necessary to carry out an investigation of the range of robotic percussion solutions currently in use before designing the hardware. There have been some earlier studies that evaluated a small range of mainly solenoid-based solutions for robotic drumming. These studies compared relatively few mechanisms and actuator types in relation to the breadth of solutions that are currently in use among various projects [59, 110]. These earlier studies also measured relatively few of the characteristics necessary to make an informed selection of actuator and mechanism for a range of purposes, and their methodology is not clearly specified.

This section seeks to rectify this with a study of a larger range of mechanisms and actuators, and with an evaluation methodology that more thoroughly and fairly compares the various options with regards to the parameters relevant to robotic drumming. It achieves this by initially outlining the types of actuators to be compared, and then describing the representative mechanisms that were built for carrying out these investigations. Following that, the methodology devised for undertaking the evaluations is presented, the tests of the various characteristics are outlined and the results of the evaluations are discussed.

4.2.1 Percussion Mechanism Types

Though most robotic percussion designs are project-specific, there are several categories into which they may be grouped. This grouping may conveniently be based upon actuator type, discussed below.

Solenoids

The most common type of actuator used in robotic percussion mechanisms is the solenoid. Solenoids are durable, relatively low cost, simple to drive, and generate relatively low levels of extraneous acoustic noise. The remainder of this subsection details linear and rotary solenoids.

When using linear solenoids for robotic percussion, the target instrument is either struck directly by a push-type solenoid's plunger (as is the case with the Closed-Loop Robotic Glockenspiel described in Chapter 3), or by an attached beater which rotates about a pivot (as in [144]). For acoustically quiet operation, linear solenoids should be mounted with suitable dampening.

In applications that utilise rotary solenoids (like the one shown on the right of Figure 4.1), beaters may be attached to the rotary solenoid's external rod, or to its rotating plate.

FIGURE 4.1: Left to right - pull-type tubular linear solenoid, push-type tubular linear solenoid, pull-type D-Frame linear solenoid, rotary solenoid.

While rotary solenoids are often more expensive than linear solenoids, they afford designers the ability to create mechanisms with few additional moving parts, reducing parts count and complexity. Rotary solenoids are notably used in Murphy's Nudge [103] and the KarmetiK NotomotoN [63] among others.

Pneumatics

While solenoids are the most common actuator for robotic percussion systems, pneumatic actuators are used by some notable projects, including Ken Caulkins' musical installations [17], Compressorhead, and the Zima Z-Machines [147].

The benefits of pneumatics are that they are powerful and durable, though they are also relatively expensive. There is a large variety of pneumatic actuators from which to choose, some of which provide longer travel than most

solenoids. A drawback of pneumatic systems is that the required air compressors and exhaust lines often produce high levels of extraneous noise, occupy excess space, and further add to the cost of the system.

Servos

Servos are electronic motor assemblies which are characterised by their ability to rotate to specific positions on command. They are used in a range of musical percussion projects, including Talmudi et al.'s 'Drums v0.1', Joakim Sand and Espen Stromme's 'Drum Machine', and Gadd's Beatbot [36]. Because of their ability to be positioned explicitly in software, they are able to be multifunctional, and can strike from various heights without the manual adjustments needed with other actuator types. However, servos generate a relatively high amount of acoustic noise, and are generally less durable than pneumatic devices and solenoids. In addition, since servos' travels are explicitly specified in both directions, the risk of the servo applying pressure to a drum head rather than bouncing from it freely is higher than with solenoids.

4.2.2 Test Mechanism Construction

To compare the different types of mechanisms and demonstrate the evaluation methodology, eight mechanisms that are representative of popular musical robotic percussion systems, are presented below. Two variations of each of the actuator types outlined above are created in order to more thoroughly explore different varieties within those actuator groups and configurations.

Rotary Solenoid Strikers

Two rotary solenoid strikers are shown in Figure 4.2. Both make use of aluminum mounting brackets to affix them to microphone stands and to mount input sockets. The left device also uses an aluminium plate to affix a hollow 6 mm aluminum striking rod to the rotary solenoid's plate, whereas the right mechanism's 3 mm steel striking rod is welded to a shaft collar which is attached to the solenoid's rod.

The left device uses a smaller 39 mm Ushio SP-B model, and the right device uses a larger 48 mm diameter Shindengen model. Both are rated for 24 V at

a duty cycle of 100%. The devices are chosen as representative of those often used in robotic percussion applications. The SP-B is capable of 25 degrees of rotation, so a dampener is not utilised to limit its travel, and the Shindengen model has a 90 degree rotation, limited by a rubber stopper. Both of their internal springs are set to maximum tension.

FIGURE 4.2: The rotary solenoid-based percussion mechanisms. Left - Ushio Small Rotary Striker, right - Shindengen Large Rotary Striker.

Linear Solenoid-based Strikers

A linear solenoid with a pivot is shown on the right of Figure 4.3 and is inspired by the Trimpin Hammer [59] and the Kaltron [63]. It is built around a tubular, linear, pull-type solenoid with a plunger that is tapped to allow a male ball-joint to be affixed. A female ball-joint is also attached to the main bracket at the at-rest height of the male one. A 6 mm hollow aluminium rod is then inserted through both of the bearings, and terminated with a wooden ball to act as the beater. To hold the rod in place, two shaft collars are affixed to the rod on either side of the bearings. The solenoid used is an HCNE4-25/50TL, which has been chosen for its rating of 24 volts, and a duty cycle of 25%. These specifications match the requirements of this application in which the solenoid is delivered short pulses of higher voltage. This solenoid is a mid-range device which can be obtained at a relatively low cost and can be considered representative of devices used in these applications.

The bracket, which is a single piece of 3 mm laser-cut and heat-folded acrylic, provides both a mounting plate for the mechanism, and a stopper for the top end of the striker's travel. Dampening material is attached to prevent the mechanism from bouncing upon return and to reduce extraneous acoustic noise.

The direct linear solenoid striker is shown on the left of Figure 4.3 and is inspired by mechanisms such as the Kapur Finger [59], and <Snar> [131]. It

utilises a linear, tubular, push-type solenoid (model HCNE4-25/50L), rated for 24 V and 25% duty cycle. This is the exact same model as the solenoid used for the pivot mechanism described above, but in a push-type configuration. The laser-cut, heat-folded 3 mm acrylic bracket provides mounting capability via a standard microphone stand connector, and dampening of the return action of the solenoid to prevent bouncing upon return.

FIGURE 4.3: The linear solenoid-based percussion mechanisms. Left - Direct Linear Solenoid Striker, right - Linear Solenoid with Pivot.

Servo Strikers

An RC-Servo-based drum striker is shown on the left of Figure 4.4, and is built to closely resemble the rotary solenoid mechanisms described above, providing the same simplicity of integration. It consists of a single piece of laser-cut acrylic, heat-folded at a 90 degree angle to create a mounting bracket for the RC Servo. A hollow 6 mm aluminium rod is connected directly to the servo's shaft-mounted horn. The servo used is a Blue Bird 617DMG+HS, which is a popular, standard form-factor, low-cost, metal-gear RC hobby servo. The jarring motion that robotic striking mechanisms must produce make metal-gear servos such as these preferable to their nylon-gear equivalents, whose more fragile gears may be stripped by repeated striking events.

While analogue servos are controlled by regular pulses of voltage whose widths dictate their position, digital servos are controlled by serial messages. This allows a number of different messages to control a range of aspects of digital servos' performance, and when bidirectional communication is implemented, the servo can send position and other information back to the controller. A digital servo-based drum striker was created, shown on the right of Figure 4.4. This allows a comparison of the suitability of analogue and

digital servos for robotic percussion applications. This mechanism utilises a Dynamixel AX-12 running at 1 Mbps baud rate and maximum speed. The mounting bracket and beaters were made to resemble those of the analogue servo striker as much as possible, with the same laser-cut and heat-folded acrylic, and identical 6 mm rods and 24 mm diameter wooden balls.

FIGURE 4.4: The servo-based percussion mechanisms. Left - Analog Servo Striker, right - Digital Servo Striker.

Pneumatic Strikers

The pneumatic striking mechanism shown on the left of Figure 4.5 has been built to closely resemble the linear solenoid with pivot mechanism outlined above. It is built around a Festo DSNU-16-MP-A double-acting pneumatic cylinder, which was chosen due to its size and travel closely matching that of the linear solenoids used in this evaluation. A female ball-joint is affixed to the cylinder's plunger and a second female ball-joint is attached to the main bracket behind the first. A 6 mm hollow aluminium rod is inserted through both of the bearings, and terminated with a wooden ball to act as the beater. To hold the rod in place, two shaft collars are affixed to the rod on either side of the bearings.

The bracket, which is a single piece of 3 mm laser cut acrylic, provides a mounting plate for the mechanism, with a standard microphone stand socket attached. A stopper was not used for this device, as the cylinder is double-acting, meaning both directions are individually pneumatically powered.

The novel pneumatic striking mechanism shown on the right of Figure 4.5 was designed to test the utility of magnetically coupled rod-less cylinders in robotic percussion applications. This single-acting with spring arrangement was created to provide a point for comparison with the pneumatic double-acting striker. The device has been built to closely resemble the linear solenoid with pivot mechanism utilised with the other striker with slight modifications to accommodate the design of the actuator. It is built around

an SMC CY1B6H-40 magnetically coupled rod-less cylinder and, in order to achieve a similar action to the other units for comparison, the carriage is fixed to the acrylic bracket and the cylinder barrel creates the actuation. Similar to many of the strikers in this evaluation, this design also features a 6 mm aluminum beater manipulated by two ball-joints and held in place by two shaft collars. Stoppers were not used for this device, as the actuator naturally limits its movement.

In testing, the internal cushioning feature of this actuator was found to greatly restrict its speed, causing it to exhibit unacceptably high latency and very low striking power. The result of this experiment shows that this actuator should not be recommended for percussion applications, and it is excluded from the more detailed analyses presented in the following sections.

With eight mechanisms built, and seven of those mechanisms deemed suitable for comparative evaluation, the experiments can begin.

FIGURE 4.5: The pneumatic percussion mechanisms. Left - Double-Acting Striker, right - Magnetically Coupled Rod-less Striker.

4.2.3 The Evaluation Methodology

This section details the specifics of the testing setup and provides information about the equipment used and how the measurements are made.

A Remo Fiberskyn frame-drum was used as the target for all of the tests, as it is a portable, single head drum that allows easy access for the attached hardware. In order to measure the movement of the drum's membrane, a 35 mm diameter piezoelectric disc is affixed to the underside centre of the drum's skin. The audio interface used is set to a sample rate of 44.1 kHz and

a resolution of 16 bits. All strikes occur 55 mm from the piezo disc. A sound-level meter is positioned 300 mm above the drum and is pointed at the centre of the drum at a 45 degree angle. The meter is set to dBC mode.

To gain insight into the root causes of device latency, an electromagnetic pickup mounted near the solenoids and DC-motors is used to sense the on-set of coil power-up. By viewing the pickup's output, the latency contributed by the electronics can be compared to the latency contributed by the actuators. This process is illustrated in Figure 4.6 and the latency components are discussed in Section 4.2.4.

FIGURE 4.6: Measuring the various stages of the system's latency.

A 48 V driver box using the IRL640 MOSFET control circuit outlined in [126] is used to actuate the solenoid-based mechanisms. The analog servos receive 6 V and the digital servos, 9.6 V. The pneumatic devices receive 6 bar regulated air pressure.

4.2.4 The Evaluations

The percussion mechanisms' latency, maximum loudness, dynamic consistency, and maximum repetition rates were evaluated, as collectively, these characteristics describe the general behaviour of a robotic drumming mechanism. The methodology for carrying out these individual tests and the results are shown below.

Latency

The time it takes for a mechanism to strike a drum after being commanded to is considered its latency. This is an important attribute when evaluating musical robotic systems, as instruments with lower latency will respond to live instructions quicker, and help to maintain an adequate cause and effect, instrumental relationship with the human issuing the commands. While given the physical nature of mechanical movement, latency can not be eliminated entirely, in contexts that involve pre-generated musical sequences, latency can be compensated for in software as long as that latency is consistent.

As shown in Figure 4.6, there are several components to the latency exhibited by robotic percussion mechanisms, and it can be useful to observe these when attempts are made to reduce overall latency. The first value is the time between MIDI note transmission and the microcontroller receipt. In this testing setup the first latency value was an average of 1.69 ms with a standard deviation of 0.39 ms.

The second component of latency is the time between microcontroller transmission and actuator receipt, as shown by an electromagnetic pulse observed by the pickup. For the solenoid mechanisms, this value was an average of 1.89 ms with standard deviation 0.58 ms. These small latencies are inherent to any MIDI based electronic instrument due to the serial protocol and standard baud rate of 31250, and are relatively insignificant compared with the mechanical latencies. The standard deviation of these values can be explained by the polling-based UART reading and GPIO writing Arduino routines utilised, and are small enough not to affect the latency comparisons between mechanisms. For the analogue servo, this value was an average of 15.5 ms with standard deviation 1.8 ms. This much larger value can be attributed to the electronics built into the servo needing to process the signal before moving the motor, and the greater standard deviation is a result of variance in the phase of the PWM signal used to control it. This contributes a significant amount to the servo's overall latency, and while it can be compensated for in contexts with predetermined musical sequences, this servo's performance in real-time interactive contexts is unavoidably worse than the solenoid's.

The final component of the latency is the time between actuator receipt and the beater making contact with the drum, and is related to the speed and height of the striker. The total latency is the time between the original MIDI

FIGURE 4.7: The latencies of the evaluated mechanisms.

note being sent and the beater making contact with the drum. To obtain an average value, these total latency measurements were performed in sets of ten strikes.

Each mechanism was tested at its low, medium, and high heights above the drum head, allowing for a more comprehensive view of its overall latency characteristics. Figure 4.7 shows the results of the latency tests of the evaluated mechanisms. The Ushio Rotary Solenoid's middle position value is higher than its high position value because it was necessary to program a lower velocity in the middle position to prevent bouncing in the test setup. The graph illustrates that the linear solenoids provide the lowest latency characteristics, with the rotary solenoids and pneumatic system providing moderate performance, especially in lower positions. Both of the servos exhibited high levels of latency, with the digital servo the highest by a significant margin.

Maximum Loudness

How loud an instrument can play may dictate the performance environments that are most suitable for its application. In some instances it is necessary to choose the mechanism that can hit the loudest, regardless of other variables such as latency or speed. As the height of maximum loudness varies across devices, the devices under test were positioned in order to maximise their loudness and the absolute maximum values are shown in Figure 4.8, together with the heights at which those loudnesses were achieved. The graph shows that the pneumatic device is capable of actuating the loudest strikes, the rotary solenoids and linear solenoids with pivot are middle of the range, the servos are quieter, and the direct striking solenoid had the lowest maximum loudness of the group.

Dynamic Consistency

When writing and performing music with robotic percussion instruments, it is important that strikes that are programmed to be identical have consistent dynamics. To assess the extent to which the devices are able to strike at a consistent loudness, sets of ten 1000 ms strike recordings were analysed. RMS values were taken, and the standard deviation of these at varying heights

FIGURE 4.8: Maximum loudness values (dBC) of each striker with heights included.

are presented in Table 4.1. Lower numbers in this table mean more consistent strike volumes. As the pneumatic mechanism tested has a fixed travel in both directions, it can only be tested at a single height.

	Low	Medium	High
Linear Solenoid (Pivot)	3.50 (10 mm)	3.53 (20 mm)	4.25 (30 mm)
Linear Solenoid (Dir. Strike)	5.05 (4 mm)	7.12 (6 mm)	9.52 (8 mm)
Ushio Rotary Solenoid	4.34 (10 mm)	2.36 (55 mm)	1.44 (100 mm)
Shindengen Rotary Solenoid	6.99 (10 mm)	7.40 (100 mm)	3.22 (190 mm)
Analogue Servo	4.34 (10 mm)	5.21 (65 mm)	6.08 (120 mm)
Digital Servo	2.93 (10 mm)	3.71 (65 mm)	3.16 (120 mm)
Pneumatic Striker	N/A	10.83 (50 mm)	N/A

TABLE 4.1: Percentage standard deviations of the RMS Values of ten strikes at various heights.

Maximum Repetition Rate

Every robotic striker has a limit to how many strikes per second it is able to achieve, and this limit dictates the speed of the musical passages that can be written for it. In some performance situations, extremely fast rolls are called for, and a mechanism that is capable of those speeds may be chosen. A maximum repetition rate for each mechanism was achieved by positioning

the strikers near the drum and adjusting the MIDI control signals. During these tests, the strikers did not necessarily return to an at-rest position. To prevent damage to the drum from the pneumatic beater, a height of 50 mm was used for its test. The numbers of strikes per second were observed; the results are shown in Figure 4.9. The graph shows that the linear solenoid with pivot and Ushio solenoids provided the highest maximum repetition rates, the analogue servo, pneumatic striker and the larger Shindengen solenoid occupied the middle of the range, and the digital servo and the direct striking solenoid were the slowest of the group.

FIGURE 4.9: Maximum repetition rates achieved by each striker in strikes per second with heights above drumhead specified for each striker.

4.2.5 Results

The results of these evaluations allow several observations to be made about the range of robotic percussion devices' suitability for various musical contexts, and are described below.

1. The servos both received low scores for loudness and repetition rates and exhibited the highest tested latencies. While dynamically consistent, the digital servo exhibited especially high levels of latency. The analog servo exhibited lower latency but with lower consistency. Both

produce high levels of extraneous acoustic noise. These findings suggest that the other types of devices are more suitable for robotic drumming contexts than these servos.

2. The double-acting pneumatic device showed high loudness and latency consistency, but poor dynamic consistency. It is incapable of velocity control, produces high levels of extraneous acoustic noise, and is expensive. Though further research is needed with a variety of pneumatic actuators, preliminary testing of a new pneumatic striker indicates that magnetically coupled cylinders with cushioning are unsuitable for percussion applications.
3. The smaller rotary solenoid exhibited good dynamic consistency and high repetition rates. The larger model showed similar latency, but worse dynamic consistency and repetition rates. It was also found to hit harder than the smaller model.
4. The direct striking linear solenoid scored the best in latency, but did not do well in other tests. The linear solenoid with pivot was shown to be a good general purpose unit with acceptable results in every characteristic measured.

While the selection of actuators utilised in these evaluations cannot be considered exhaustive, the results of these tests suggest that pneumatic percussion mechanisms can be particularly useful in situations where space, extraneous noise and high cost are not prohibiting factors. They also indicate that the tested models of servos can be considered inferior choices for robotic percussion compared with the other options. For cases that extraneous noise, high cost, or spacial constraints are concerns, the evaluations show solenoids to be the most effective of the actuators tested for the task. In addition, the rotary solenoid and pull-type solenoid with pivot mechanisms provided all-round good performance.

For the purposes of creating a new robotic drum kit, the choice between rotary solenoids and linear types becomes one of availability, cost, and mounting options. While it is possible to acquire second hand solenoids for affordable prices, sourcing them in this way is not stable and reliable, meaning that replacement parts and future upgrades may be difficult or costly. In the case of parts acquisition for this project, the cost of new rotary solenoids is significantly higher than the cost of equivalent new linear solenoids, even with the extra mechanical parts required to implement the pivot. It is for this reason

that the linear solenoid with pivot design was selected for this robotic drum kit.

4.3 Striker Construction

With the linear solenoid with pivot design decided upon, a set of robotic drum striking mechanisms are created. The key goals of the new striker design are that they are:

1. Fixed to the instruments they are playing so as to minimise shifting of the strikers in relation to the drums during performance. This also eliminates the use of microphone stands for positioning strikers, simplifying setup and transportation logistics.
2. As modular as possible, allowing parts to be quickly and easily replaced in the system with backups. The parts should also be generalisable and flexible enough to be used in conjunction with new percussion instruments acquired in the future.
3. Strong enough to endure the rough environment of touring, capable of surviving being dropped.
4. High performance: matching or exceeding the measurements taken of the example strikers in Section 4.2.
5. Cost effective, so that it is not prohibitive to make a higher number of them, keep backups, and replace them if necessary.

The following subsection describes the design and construction of the new robotic striking mechanisms and addresses the particular mounting requirements for the individual drums and cymbals.

4.3.1 Modular Drum Strikers

Figure 4.10 illustrates the design of the new robotic percussion striking mechanism in its default configuration. The design maintains several elements of the linear solenoid with pivot, and the Timpin Hammer [59] and Kaltron [63] mechanisms, in that it utilises a pair of male and female ball joints with a 6 mm rod inserted through them, held in place by shaft collars. A spring is fitted around the solenoid plunger to assist with the beater's quick return,

and dampening foam is mounted above the beater. This dampener ensures that the beater begins its stroke from a set point, aiding in improving the consistency of the velocity and latency of strikes. A 2.5 mm DC barrel socket is also included in the mechanism for connection to the control system.

FIGURE 4.10: A rendering of the new robotic striker design.

The mounting assembly is based around two laser-cut pieces of 2 mm thick steel held together in three places by two M4 machine screws and hex nuts, and the solenoid coil and hex nut. Utilising steel instead of the previously used acrylic and aluminium constructions increases the rigidity of the mechanism, allowing it to maintain its shape, and therefore, accuracy even after sustaining multiple impacts. By designing the brackets to be laser cut, with just three 90° bends, many sets of them can be created efficiently and relatively inexpensively using modern Computer Numeric Control (CNC) construction techniques.

The bottom mounting bracket of the mechanism is fitted in order to avoid using microphone stands for positioning, and to enable flexible attachment to a range of mounting apparatuses. On the bracket, there are two columns of M4 holes spaced 40 mm horizontally and 6 mm vertically. These holes allow for vertical displacement in order to fine tune the position of the beater, while still maintaining the ability to lock the mechanism in place. By adjusting the

height using these holes in conjunction with adjusting the height of the spacers positioned under the female ball joint and the height of the dampening foam used, a large variety of heights and angles of strikes can be achieved. In this way, the user can customise the configuration of the unit to suit the instrument being struck. The configuration and mounting strategies used for the various instruments of the standard drum kit are outlined below.

Snare

A pair of robotic striking modules are shown mounted to a 13" snare drum in Figure 4.11. The strikers shown are mounted with standard maple 5A sized drum sticks that have been cut to an appropriate length, and have had 5.9 mm holes drilled in the centre of the newly created back plane of the stick. The 6 mm metal rod (in this case silver steel) is then pressure fit into the drum stick for a tight, secure connection. The position of the stick is still adjustable by loosening the metal shaft's two collars, adjusting the position, and locking it into place. Drum sticks can also be easily substituted by removing the rear shaft collar and pulling the stick out via the front of the device.

FIGURE 4.11: Photograph of two robotic strikers attached to a 13" snare drum via a custom mount.

The speed of the stick's rebound can be coarsely adjusted by adding a second spring, as illustrated in Figure 4.11. Up to three springs can be layered in this

fashion. As the dampening material is soft, it cannot be affixed reliably in place by simply tightening a single hex nut. Instead, one hex nut is fitted above the material to firmly attach the M3 machine screw to the steel striker mount, and two hex nuts compress a washer beneath the material to fix it in place without applying pressure directly to the foam.

A custom aluminium mount compatible with the hole spacings mentioned above is created to enable the robotic strikers to firmly attach to the snare drum. The mount is created from a 3 mm thick L-bracket, cut to 150 mm in length, with sides each protruding 30 mm. A semi-circle is also ground into the drum-facing edge of the aluminium bracket to follow the shape of the snare drum. The bracket attaches to a pair of the snare drum's adjacent UNC 12-24 tension rods, with a hex nut on each side securing the bracket in place. The bracket should be secured after tuning of the drum head takes place, to ensure that tuning using the rods does not affect the mounting strength. Tension rod locks may also be used to ensure that securing the mounting bracket does not affect the tuning of the drum.

Four columns of M4 holes are drilled into the striker-facing side of the mount, with 6 mm vertical spacing and 40 mm horizontal spacing between the holes. This pattern of holes allows several different striker configurations. A single striker can be mounted in any of five positions along the bracket, with the middle three being the most secure with two columns of connectivity. Figure 4.11 shows strikers mounted in positions two and four. Since the striker's lower brackets are wider than their upper ones, a third striker can be mounted in position three between them. The fact that each column contains multiple vertical holes means that an extra two strikers can also be mounted in positions one and five, to the outside of the two strikers shown. The downside to this is that these positions will be less secure than the central positions. If more strikers are desired at different positions of the drum for extra techniques such as rim shots and rim clicks, or different beaters such as brushes or nylon tips, an identical mount can be secured on each side of the drum to accommodate up to 20 strikers at once.

Rack Toms

When selecting rack tom configurations in a drum kit, there are two predominant types: those mounted to the kick drum, and those mounted to the cymbal stands. The latter type was chosen for this robotic drum project, as

the toms with isolated cymbal stand mounts affixed provide convenient attachment points for robotic striking mechanisms, and the empty insides of the drums leave space for internal mic'ing. The drum kit chosen is a Tama Superstar Hyperdrive model. Figure 4.12 shows the front and back of the 10" tom with a pair of robotic strikers mounted.

FIGURE 4.12: Photographs of two robotic strikers attached to a 10" tom drum.

As shown in the picture on the right side of Figure 4.12, the drums come equipped with a stand mount that is attached to a black, flat metal bar that attaches to the drum in several places around its perimeter. This bar is acoustically separated from the drum with rubber isolators at its connection points, which helps to ensure that vibrations from the cymbal stand are not transmitted to the drum head. By drilling a pair of 40 mm spaced M4 holes in this stand mounting bar, it can double as a secure striker mount.

The left picture of Figure 4.12 also shows an Audix D2 hypercardioid dynamic instrument microphone mounted inside the drum. This is attached using a custom aluminium bracket, affixed to the existing internal screws and washers of the drum. The microphone is connected to that bracket, and the microphone's XLR cable is routed out of the drum via the ventilation hole. By mounting the microphone internally, the system is more integrated, requiring less setup time and eliminating the use of extraneous microphone stands or mounting hardware. It also provides an isolated, close-mic'd sound that includes very little bleed from other nearby drums and sound sources. These characteristics aid the implementation of the closed-loop audio feedback system described in Chapter 5. The 12" rack tom is automated and mic'd in an identical manner to the 10" tom.

Floor Tom

As the 16" tom stands freely on the floor, there are no mounting bars attached and an alternative mounting strategy is required. This tom stands on three legs that are equally spaced around the drum, and the height of each is adjustable. The vertical mount shown in Figure 4.13 was developed to attach to the protruding sections of these legs in order to fix the striking mechanisms firmly to the floor tom. The vertical mount is a rounded rectangle of 2 mm steel with a pattern of holes on one end that matches the mounting holes of the striking mechanisms, and a Tama floor tom leg mount attached to the other. Since this mount is designed specifically to attach to these floor tom legs, they fit well. This mounting strategy allows the strikers to be quickly and easily attached to the drum, and adjusting the striker's yaw and height can be done before tightening the mount's single wingnut.

FIGURE 4.13: Photograph of two robotic strikers attached to a 16" floor tom drum with custom vertical mounts.

As the 16" tom has a larger radius than the smaller 10" and 12" toms, and striking the drum head near the centre provides the most desirable sound out of the instrument, this drum requires longer beaters. When a longer beater is used, more weight is applied to the solenoid actuator. In order to counter this, in addition to (or instead of) adding extra springs, a counterweight can be attached to the opposite side of the beater, as shown on the left of Figure 4.13. This solution resembles the one utilised by Raes in his <Snar_2>

snare drum robot [132]. The weight can be fine-tuned by adjusting the position of the counterweight along the beater's shaft. A microphone is also mounted internally using a similar method to the smaller toms, except an Audix D4 microphone is used. The D4 is similar to the D2 except that it has an extended low frequency response, suitable for capturing the fundamental frequency of the larger tom drum.

Kick Drum

In order to maintain a familiar sound and visual paradigm for the kick drum of this robotic drum kit, rather than adapting the modular striking mechanisms to the much larger drum or creating an entirely new device for the instrument, a commercially available Tama Iron Cobra pedal is modified for automation. The modifications (as shown in Figure 4.14) involve removing the pedal of the mechanism and replacing it with a pair of large D-frame linear pull-type solenoids rated at 24 V. These solenoids are held in place by a 5 mm thick aluminium L-bracket, cut to a length of 80 mm, and sits 80 mm high and 50 mm wide. The bracket is attached to the Iron Cobra frame with an M6 machine screw below the solenoids.

The solenoids are connected to the barrel socket shown on the left of Figure 4.14 in a parallel configuration so that both solenoids are activated with equal power together, not controlled separately. An M2 threaded steel rod connects both solenoid plungers with the kick pedal's chain, and the rod and the chain are held in place with M2 hex nuts. Applying power to the solenoids causes them to pull down on the chain in a similar way to a press on the kick pedal, which rotates the axle that in turn rotates the beater toward the kick drum.

When power is withdrawn from the solenoids, the adjustable tension spring returns the beater from the drum. In order to ensure that each stroke begins from the same spot, maintaining velocity consistency, a dampener is added to the mechanism. The dampener is mounted to the top of the Iron Cobra frame via a pair of 8 mm thick steel tubes with M6 machine screws threaded through them. The dampener itself consists of a 3 mm thick aluminium L-bracket cut to 115 mm in length, with both sides extending 25 mm out. Attached to the L-bracket are three layers of different dampening materials ordered most dense to least dense from the bracket to the beater. The least dense foam cushions the beater on its return stroke, reducing audible effects, while the

FIGURE 4.14: Photographs of the front and side of the robotic kick drum striking mechanism.

more dense materials add support to the foam while isolating the impacts from the L-bracket.

An internal mic'ing system is implemented in the kick drum for similar reasons to that of the tom-toms. Since the kick drum is 22" in diameter and has different practical requirements to the tom-toms, an alternative mic-mounting strategy is required. Inspired by the commercial Kelly SHU™ system¹, the microphone is attached to a metal mount near the centre of the drum, which is held in place by an array of bungee cords that act as shock mounts. These cords are each cut to a length of 6", terminated with hooks, and stretched between an array of custom-made aluminium L-brackets attached to the drum shell's screws, and the 'H' shaped microphone mount, as shown in Figure 4.15.

An Audix D6 cardioid dynamic instrument microphone is used as the internal kick drum mic due to it having a frequency response specifically tuned for kick drums, with an emphasis on the 30 Hz to 100 Hz range, and a dip in the 200 Hz to 1 kHz range. The microphone is also built to withstand the high

¹<http://www.kellyshu.com/>

FIGURE 4.15: Photograph of the internal mic'ing system implemented in the kick drum.

sound pressure levels of a kick drum, and can receive at least 144 dB without distorting the signal. Figure 4.15 shows the microphone mounted in the top middle of the microphone mount, however, the mount is created with three positions, enabling multiple microphones to be used facing the beater head, the resonant head, or in other configurations.

A ring of dampening foam was attached around the circumference of the drum, applying slight pressure to the kick's resonant head. This provides some control over the envelope of the kick drum sound. Although the resonant head pictured in Figure 4.15 has a port hole, internal kick drum mic'ing systems such as this one also allow the use of resonant heads without holes, while maintaining the ability to mic the inside of the drum. The D6's XLR cable is routed through the head's vent hole where it is connected to the audio system externally.

Ride Cymbal

Figure 4.16 shows top and side perspectives of the robotic strikers and mounting system utilised to actuate the drum kit's ride cymbal. A 20" Paiste Signature Full Ride was chosen due to its defined stick sound, complex wash, and pronounced bell sound that is partially isolated from the bow. This enables two distinct timbres to be created with the cymbal and the strikers are positioned to take advantage of this. In the left image of Figure 4.16, the left striker is positioned to hit the bow of the cymbal. In this case, a nylon-tipped Oak Tama Traditional series stick of size 5AN was used in order to achieve an increase in stick sound with less sustain. Nylon-tipped sticks also provide the added benefit of longer durability.

For striking the bell of the ride, a less expensive maple stick with wooden tip was utilised, as the tip will not be used for striking. Instead, the shoulder of the stick is used, providing a more pronounced sound with broader harmonic content. As this stick must reach past the mid-point of the cymbal, it must be significantly longer than the bow striker. This length puts extra weight on the solenoid similar to the floor tom described above, so an additional spring is added to the plunger, and supplementary weight is added to the opposite end of the mechanism to balance it.

In order to fix the position of the strikers to the ride cymbal while enabling quick and flexible adjustment, a mounting system was devised that utilises a combination of commercial drum hardware and custom-built mounting

FIGURE 4.16: The top (left) and side (right) of the ride cymbal striking mechanisms and mounts.

equipment. A Pearl ADP-30 three-way adjustable clamp is attached to the cymbal stand, and can be moved vertically and rotated horizontally before being fixed to its position. This clamp allows two other devices to be attached to the stand.

A custom 10 mm diameter steel mounting bar was shaped by cutting the bar to size, applying heat to sections of the bar, and creating a pair of 90° bends. Several of these mounting bars were created of different dimensions in order to accommodate a range of cymbal sizes. As shown on the right image of Figure 4.16, the mounting bar's attachment to the ADP-30 is not concentric with the ride cymbal. This means that with the ADP-30 in a fixed position, rotating the mounting bar allows fine adjustment of the robotic strikers toward or away from the centre of the cymbal.

A vertical mount identical to that used with the floor tom, described above, is used to attach the robotic strikers to the mounting bar. This allows the vertical position and the rotation of the strikers to be adjusted easily by hand. In the case of using more than one striker on a cymbal, a horizontal mount is also utilised. Similar to the vertical mount, the horizontal mount is a 2 mm thick steel rounded rectangle that has an array of holes in it that mirrors the dimensions of the holes of the robotic strikers' mounting brackets. Similar to the snare drum mount described above, the horizontal mount provides three primary positions in the middle, and two extra positions on its side.

Crash Cymbal

For the drum kit's crash cymbal, a 16" Paiste Signature Crystal Thin Crash is chosen. Figure 4.17 shows a robotic striker equipped with a 5A maple stick mounted to the cymbal's stand via an ADP-30, a mounting bar, and a vertical

mount, similar to the ride cymbal configuration described above. To provide a full crash sound and to reduce the chance of creating cracks in the cymbal, the striker is positioned so that the body of the stick makes contact with the cymbal at an angle, slightly rotating the cymbal during each strike.

FIGURE 4.17: Photograph of the crash cymbal striking mechanism and microphone mounting system.

The Mic Holders MHHH commercial microphone mounting system is utilised to pick up the sound of the crash cymbal, shown on the right of Figure 4.17. This mounting system connects to the cymbal stand with a pair of wing nuts that clamps to the stand's neck via a pair of felt strips. These strips help to isolate the microphone mount from audible vibrations that may be transmitted from the stand. The mount is hollow, and utilises an extra adjustment bar with a pair of screws to allow localised rotation and height adjustment of the microphone without shifting the clamp's position on the stand. This bar is terminated with a solid rubber shaft with a microphone connection thread fixed to the top of it. This shaft acts as a second acoustic isolator for preventing unwanted vibrations from reaching the microphone.

The microphone chosen for use with this cymbal is the Audix ADX51 small diaphragm cardioid pre-polarised condensor, due to its high end frequency response and its selectable 150 Hz high pass roll-off feature. This allows the

mic to be positioned near to the cymbal in order to reduce bleed from other instruments of the kit, while reducing the occurrence of low frequency boost due to the proximity effect.

Hi-hat

The configuration utilised for robotically striking the hi-hat of the drum kit is shown in Figure 4.18. A set of Paiste Signature Precision Sound Edge 14" hi-hats were chosen due to their bright, defined stick sound, and sound edge design that allows air to escape uniformly around the hi-hat while performing a 'chick'. This design feature prevents air from being compressed between the cymbals upon closing them together, which has the effect of dulling and adding undesirable low frequency content to the sound.

The robotic strikers for the hi-hat are configured in a way similar to that of the ride cymbal described above, utilising an ADP-30, mounting bar, and vertical and horizontal mounts. In this case however, the strikers are equipped with two Tama nylon-tipped sticks of equal length. The hi-hat is mic'ed using the same MHMH microphone mounting system and Audix ADX51 condensor microphone as utilised with the crash cymbal configuration described above.

4.4 Hi-hat Movement System

Unlike the other instruments of the drum kit, hi-hats utilise a pedal to adjust the distance between the two cymbals during a performance in order to expressively manipulate the envelope of strikes to the cymbals, alter the timbre of strikes, and to create a 'chick' sound by closing them quickly, among other techniques. In order to enable these modes of expression with this instrument, building a movement system that actuates these types of techniques is necessary.

There are several methods of actuating the pedal mechanism of a hi-hat. If anthropomorphism is desired, as is the case with the Compressorhead [4] and Z-Machines [147] robot bands, a mechanical foot can be used to press on the hi-hat's pedal. This allows the hi-hat to remain entirely unaltered and provides the audience with a visually familiar paradigm. However, the force required to overcome the hi-hat's internal spring is high, and both bands use

FIGURE 4.18: Photograph of the hi-hat striking mechanisms and microphone mounting system.

pneumatic systems that create a high level of extraneous noise and very high cost as outlined in Section 4.2.1.

Another method described by Moritz Simon Geist for his MR-808 ensemble utilises a large solenoid in an on-off configuration to actuate the hi-hat position [39]. To minimise the ‘on time’ of the actuator and avoid overheating, Geist reverses the polarity of the hi-hat movement from default high, to default low. This configuration assumes that the most common position for the hi-hat during most performances is low, and still requires over 30 A of current to power the large solenoid used to lift the hi-hat. It also relies on the force of gravity to keep the hi-hat closed, limiting the tightness the instrument can achieve in the closed state.

Other potential candidates for actuation include servo and stepper motors, both of which would be capable of providing accurate control of the height of the pedal with an appropriate mechanical coupling, but both of which also generate a high level of extraneous noise which is undesirable for performances of very dynamic musical material and quiet recording situations. There also exist a number of linear actuators such as voice coil actuators which provide built-in closed-loop positioning and quiet operation, but require an investment of thousands of dollars for the actuators and corresponding control systems.

4.4.1 Hi-hat Movement Mechanics

With all of the considerations noted above, it is decided that the most appropriate actuator for this application is the Ledex 196658-028 5EPM soft-shift solenoid. These are solenoid actuators that are designed to deliver a relatively linear force response proportional to the current delivered to them. This allows the soft-shift actuator to provide a very quiet linear positioning capability, though without built-in position sensing. While it does not cover every conceivable use-case, the chosen actuator delivers a 10 mm range of movement that is capable of positioning the hi-hat from fully closed with pressure applied, to fully open with the two cymbals not making contact with each other.

In this design, the hi-hat’s original pedal is removed, but the pedal’s base frame is left intact and a replacement mechanism is mounted in the pedal’s place. Figure 4.19 shows the two soft-shift solenoids mounted to their control

FIGURE 4.19: Annotated photograph of the hi-hat movement system.

box. All of the electronics that drive the solenoids and carry out the control are housed inside the 6 mm thick black laser-cut enclosure. The structure attaches to the soft-shift solenoids by their mounting screws, providing a stable platform that is affixed to the hi-hat pedal frame. The bottom of the enclosure includes slots that lock it into the correct position on the frame, and an M6 machine screw affixes the enclosure tightly to it. The cymbal-pulling shaft that was connected to the pedal by way of a chain is instead directly mechanically coupled to the soft-shift solenoids via an aluminium bracket. The solenoids' plungers are also tapped to accept M3 screws, firmly attaching them to the aluminium bracket.

The hi-hat's internal spring requires significant power to overcome even when configured for minimum tension, so in this design it is disabled entirely. The soft-shift solenoids are equipped with internal springs that exert upward pressure on their plungers, but the strength of these springs is insufficient for lifting the hi-hat with an adequate response. In order to raise the hi-hat, new external springs are added between the solenoids and the aluminium bracket to augment the response of the internal ones. The extra springs are chosen to have sufficient strength to allow the hi-hat to retain a responsive default-open configuration, while being weak enough for the actuators to overcome them and apply adequate pressure to the hi-hat in the closed position.

4.4.2 Positioning System

Soft-shift solenoids have a much more manageable response than standard push/pull solenoids (as shown on their datasheet²), though the resulting response is not entirely linear. As a result, a closed-loop positioning system is designed in order to increase the accuracy of the unit's height control.

The design of the position feedback system adheres to the principles outlined in Section 2.4.3. To retrieve the position of the solenoid plungers, the design includes a neodymium magnet which is glued to the lower end of one of them. Below the magnet, an Allegro Microsystems A1301 Continuous-Time Ratiometric Linear Hall-Effect Sensor³ IC is mounted. The strong magnetic

²Ledex Soft Shift Solenoid Size 5EP Datasheet:<http://disti-assets.s3.amazonaws.com/testco-inc/files/datasheets/3995.pdf>

³Allegro Microsystems A1301 Datasheet:<http://www.allegromicro.com/~media/Files/Datasheets/A1301-2-Datasheet.ashx>

FIGURE 4.20: Diagram of the mechanism used to control the hi-hat height, with internals made visible.

field created by the magnet induces a voltage in the Hall-effect sensor's analogue output that is proportional to the position of the the solenoid's plunger, thereby providing the electronic control systems with information about the actuator's height. A diagram is presented in Figure 4.20 where one of the sides of the enclosure is made invisible so that the internals can be shown. This illustrates the connections between the hi-hat position rod, aluminium bracket, soft-shift solenoids, magnet, and Hall-effect sensor.

The design concept was that the position information sent to the onboard microcontroller's ADC would be fed into a PID algorithm, allowing the firmware to set a target height, rather than a voltage level. A diagram illustrating the architecture and signal flow of the hi-hat movement circuitry is presented in Figure 4.21. However, during real-world testing, multiple issues were discovered with the design. In addition to the magnetic field induced in the Hall effect sensor by the neodymium magnet, the soft-shift solenoids also introduced interference to the sensor which varied with the level of current supplied to them. A potential solution to this specific issue could be to utilise position sensors that do not rely on magnetic fields.

However, another issue was observed with the concept of the positioning system that is independent of sensor selection. With hi-hats, much of the

musical expression comes from variation in the pressure applied from the top cymbal to the bottom cymbal. When the hi-hats are closed, variations in pressure show little to no variation in height. As a result, any distance-based closed-loop positioning system will not be able to achieve this fine control when the cymbals are pressed together. Following these experiments, it is decided instead to implement an audio-based closed-loop control system as described in Section 5.8.6.

FIGURE 4.21: Diagram explaining the architecture of the hi-hat movement system.

4.4.3 Hi-hat Movement Circuitry

A photograph of the controller circuit board is shown in Figure 4.22. The only interfacing the unit requires is a single connection on either side. The left side has a 3-pin DIN socket for bi-directional RS-232 communication to and from the main ensemble controller, implemented with a MAX3221ECPWR transceiver. The other side utilises a 5-conductor port that provides the system with -24 V, ground, +5 V, and dual +24 V regulated power rails provided by the custom power supply described in Section 4.4.5.

The control of actuators such as soft-shift solenoids is often carried out by applying PWM signals to the actuators via a power transistor or other driver.

In order to reduce the effect of electromagnetic interference (EMI) on the circuitry of the hi-hat controller, a novel system, shown in Figure 4.23, was invented that utilises the LM3886 *Overture*™ 68 Watt High-Performance Audio Power Amplifier integrated circuit (IC) to drive the actuator with an analogue signal directly. The circuitry of that system is described in Section 4.4.4. Though this system significantly reduces the interference received by the circuitry, some interference is still present in the Hall effect sensor's output as mentioned in the previous section.

FIGURE 4.22: Photograph of the inside of the hi-hat movement system.

These amplifiers require the use of sizable heat sinks, so are positioned off the side of the board to accommodate them, as shown in Figure 4.22. In order to ensure proper air flow to the heat sinks, MB40202V2-0000-A99 MagLev Axial Fans are mounted adjacent to them. These fans are rated for 24 V DC, are 20 mm thick and 40 mm wide and tall. They were primarily chosen due to their low noise construction, specified at 21 dBA at full speed. Due to the fact that the negative 24 V rail is only lightly loaded by the amplifiers, the fans are run from that rail.

Despite these low-noise characteristics, it is still desirable to be able to fine tune the speed of the fans, in order to produce the minimum amount of

extraneous noise, while still providing adequate airflow to the heat sinks. To achieve this, an LM337 adjustable output negative voltage regulator was utilised with a ten-turn trimmer potentiometer, allowing the voltage supplied to the fans to be precisely adjusted. By tuning the speed of the fans so that their fundamental frequency lies just below the audible spectrum, the perceived noise levels are greatly reduced.

For this board, the STM32F410RBT6 Arm Cortex M4-based 32-bit microcontroller is selected for several reasons. It is a very low cost and low power device, operates at a clock speed of 84 MHz, and provides sufficient processing power to carry out all control functions at a high refresh rate. It comes equipped with a 12 bit ADC for sampling the Hall effect sensor, and a rail-to-rail 12 bit DAC for writing power levels to the solenoids via the amplifiers. It also has the required UART peripheral for communication with the outside world, and many of its inputs are 5 V tolerant, for glueless connection to 5 V devices.

4.4.4 Signal Conditioning

FIGURE 4.23: Signal conditioning circuits. Left: From Hall effect sensor to ADC, right: From DAC to soft-shift solenoid.

In order to achieve the best practical resolution while interfacing the microcontroller with the pedal mechanism, circuitry to buffer, scale, and bias the signals is included. Figure 4.23 presents the schematics of the two conditioning circuits used in this system, one for mapping the Hall effect sensor's output to the microcontroller's ADC (left), and one for scaling and amplifying the microcontroller's DAC output to control the soft-shift solenoid (right).

The first circuit receives a voltage from the Allegro A1301 Hall effect sensor proportional to the distance between the sensor and the neodymium magnet attached to the pedal mechanism. The sensor is powered by 5 V, has a quiescent voltage of 2.5 V, and the presence of a south-polarity or north-polarity

magnetic field proportionally increases or decreases the output voltage respectively. By utilising a TLV274 opamp with rail-to-rail output swing in an inverting amplifier configuration, in conjunction with trimmer potentiometers controlling the gain and bias of the device, this signal can be scaled to range from ground to 5 V. A voltage divider on the output of the circuit scales this to the ground to 3.3 V range expected by the STM32's ADC.

The right-hand circuit also utilises the TLV274 in a similar configuration to the left one, in order to prepare the signal from the audio DAC for direct DC amplification. The LM3886, supplied with power rails of ± 24 V to correspond with the soft-shift solenoid's 24 V rating, is utilised in a DC-coupled configuration and amplifies the signal by a factor of approximately 20. The resulting higher-power signal is then applied directly to the solenoid's terminals.

4.4.5 The Power Supply

Since the hi-hat movement hardware must be compact enough to fit within the hi-hat pedal frame and it has various power requirements, a separate custom power supply is created, as shown in Figure 4.24. The power supply is housed in a laser-cut 6 mm thick acrylic enclosure with vent holes cut in the top of the unit to allow heat from the circuitry to escape, and a 9 mm thick bottom plane for added strength. It utilises a standard IEC connector with a fuse and illuminating switch to provide the system with power from a mains connection, and a custom 3D printed protection bracket shields the mains connection points from accidental contact for safety.

FIGURE 4.24: Photograph of the exterior (left) and interior (right) of the hi-hat power supply.

The mains connection is equipped with a 3 A fuse, and is routed to the primary connections of a multi-tap transformer. The secondary connections of

the transformer provide 9 V AC and 24 V AC connections which are then routed to the custom power supply circuit board shown on the right of Figure 4.24. A schematic diagram of the circuitry on the power supply circuit board is shown right of the transformer in Figure 4.25.

1N4007 diodes are utilised for rectification of the low current 5 V and -24 V rails, with 1N5404 diodes used for the higher current +24 V lines. Sets of 3300 μ F and 1 μ F electrolytic capacitors are used for bulk storage, and regulation is carried out by 7824, 7924, and 7805 linear voltage regulators. An onboard power indicator LED is included, running off the -24 V rail, as this is the least loaded. All of the supply rails are then connected to the 5-pin output socket, for connection via 5-core cable to the hi-hat movement system.

FIGURE 4.25: Schematic diagram of the hi-hat power supply's circuit.

4.5 Summary

This chapter has described the process of designing the mechanical aspects of a new robotic drum kit. The design and actuator choices were informed

by building and testing a number of exemplary robotic percussion mechanisms and comparing their various characteristics and attributes, selecting the most appropriate options for this project. The design and construction of a new set of modular striking apparatuses that further develop the state of the art in robotic percussion was outlined, and particular methods of mounting and adjusting the instruments with respect to the various instruments of the drum kit were detailed.

A new closed-loop position control system was trialled for implementing fine-grained movement of the hi-hat, complete with mechanical diagrams and schematics of the signal conditioning circuits and custom external power supply. The custom internal mic'ing systems that were integrated inside the toms and kick drum were described along with the affordances for mic'ing the other instruments of the kit.

Chapter 5 proceeds to continue discussion of the new robotic drum kit presented in this chapter, describing the novel electronic hardware built to control the robotic drum kit and other elements of unpitched percussion arrays. This control system also includes embedded DSP hardware for implementing closed-loop MIR-based audio feedback control on these modular percussion arrays. The hardware and software of the system is described and evaluations of the new system are presented.

The research presented in this chapter is featured in the following publications:

J. Long, J. Murphy, A. Kapur, and D.A. Carnegie. "A Methodology for Evaluating Robotic Striking Mechanisms for Musical Contexts". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Baton Rouge, Louisiana: Louisiana State University, 2015

J. Long, J. W. Murphy, A. Kapur, and D.A. Carnegie. "A Comparative Evaluation of Percussion Mechanisms for Musical Robotics Applications". In: *6th International Conference on Automation, Robotics and Applications, ICARA*. IEEE. 2015, pp. 173–178

J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "A Closed-Loop Control System for Robotic Hi-hats". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Copenhagen, Denmark: Aalborg University, 2017

Chapter 5

A New Robotic Drum Kit: Control

5.1 Overview

When creating an array of robotic unpitched percussion instruments, the system utilised for controlling them is of central importance to the performance characteristics and overall capabilities of the devices. Such robotic control systems must make affordances for communicating with the outside world via musical protocols, accurately controlling a variety of actuators, and providing stable power to them.

There exist many examples of such control systems which utilise central controller boxes that allow a range of different instruments to be connected to them. These are generally controlled by unidirectional MIDI streams. Examples include Kapur et al.'s Novalis system, built for use with their MalletOTon and Modulet actuators [62], The Modular-Muse Motor Drivers (MMMD) of the SolidNoise ensemble [46], Raes' control hardware from instruments like <HAT> [127], and the author's earlier percussion controllers described in [74].

Each of these systems must accommodate many different types of instruments in various positions, with a range of actuation speeds. This means that in order to achieve the desired expressive performance from each of these instruments, much manual calibration and maintenance must be undertaken, often using solutions such as those discussed in Section 3.2. This chapter discusses the implementation of a new control system that improves upon the performance of these existing systems, implementing closed-loop audio feedback-based continuous calibration functionality for robotic percussion arrays.

While the previous chapter dealt primarily with the mechanical elements of the new robotic drum kit, this chapter focuses on the system used to control it, beginning with an outline of the goals of the project before providing an overview of the entire control unit. The system is then broken down into subsystems such as the actuator control board, DSP board, and power supplies. The design considerations, construction and functionality of each of these are described in detail. Following that, the firmware written to carry out the closed-loop audio feedback-based continuous calibration functionality is outlined. Evaluations are undertaken to measure to what extent these developments improve the performance and utility of this new system compared to previous systems, and creative works that have been realised and performed with the system are presented and discussed.

5.1.1 Design Goals

The design goals of this control system primarily mirror those outlined in Section 1.3, in that the system should be able to automatically and continuously calibrate itself, making setup more straightforward, and performance more robust and reliable. Similar to the design considerations for the Closed-Loop Robotic Glockenspiel described in Section 3.3, the system should enhance the expressiveness of the musical robots it controls by allowing the confident programming of soft strikes, and precise, dynamic rhythms with latency compensation. As the functionality of passing through the picked-up audio and enabling the instruments to output MIDI messages proved useful in the composition of new works for that instrument, these features should also be possible in this modular percussion system.

Since this control unit is created to handle a not-entirely-specified array of robotic percussion instruments, both the hardware and firmware elements of the system must be more flexible than those of the Closed-Loop Robotic Glockenspiel. The actuator control circuitry must be capable of generating a range of different control signals while maintaining a high level of accuracy. The audio feedback system must also be capable of receiving audio signals from many different types of transducers, from electromagnetic pickups and piezoelectric discs, to electret, condensor, and dynamic microphones.

As the actuator and audio channel counts are not fixed, it is also important that the system is designed to be modular and must include adequate overhead in its specifications to allow straightforward expansion into larger and

more sophisticated arrays of robotic percussion instruments. This includes affordances for higher actuator control channel counts, higher audio feedback channel counts, and DSP hardware that is capable of handling these numbers of channels as well as control algorithms that have not yet been conceived.

Section 3.6 outlines several practical issues that were encountered over the course of composing music for, and delivering performances with, the Closed-Loop Robotic Glockenspiel. These include high noise floor of the audio output, audible whine of switch-mode power circuitry, insufficient accuracy of solenoid control pulses, and insufficient utility of commercially available development boards and environments. These issues must all be addressed in the design of this new percussion control system.

FIGURE 5.1: Diagram explaining the architecture of the robotic percussion array control system.

Figure 5.1 presents the basic architecture of the design created to fulfil the goals specified above. This design separates the implementation into preamp modules, a DSP board, and an actuator control board. By utilising individual modules for the preamps, channel counts for various versions, iterations, or applications of the system can be easily adjusted. Separate boards for audio analysis and actuator control allow these two functions to be carried out entirely in parallel by dedicated hardware optimised for these tasks.

While the implementation of the robotic percussion instruments and the audio transducers mentioned in Figure 5.1 are described in Chapter 4, the preamp modules, audio analysis board, actuator control board, the external audio and MIDI interfacing, and the power supply circuitry are described in this chapter. The specific design considerations (such as part selection, signal flow, and other characteristics) for each of these subsystems are addressed at the beginning of their corresponding sections.

5.2 Robotic Percussion Array Control Unit

Before describing each subsystem in detail, the entire control unit is presented in order to contextualise the discussion. Figure 5.2 presents photographs of the entire control unit, complete with close-up pictures of the back and front panels. The enclosure is constructed with 6 mm thick black laser-cut acrylic for the top and sides, with a 9 mm thick base for added strength, and 3 mm front and back panels to accommodate the interfacing hardware. The back panel image shows the standard IEC connector with fuse and illuminated switch for mains power input, three 5-pin DIN sockets for MIDI In, Out, and Thru connections, a USB-B socket for USB-MIDI communication, and two ADAT (Alesis Digital Audio Tape) optical transmitter sockets for audio output.

The front panel consists of 16 preamp channel strips, each with phantom power switches, gain potentiometers, red and green LED indicators, high impedance (Hi-Z) pushbuttons, -20 dB pad push buttons, and Neutrik XLR and TRS combo sockets. The design and functionality of these preamp channel strips are detailed in Section 5.4. This particular unit is equipped with 16 channels, as this number is sufficient to drive all of the actuators and audio channels of the new robotic drum kit described in Chapter 4, though the electronic hardware is designed to accommodate higher channel counts for expandability.

2.5 mm DC barrel sockets are mounted beneath each preamp channel, and these act as the solenoid control outputs. While they are vertically aligned with the preamp channels, there is no tight coupling necessary between the solenoid control channels and the audio feedback channels, with both single actuator to multiple microphone, and single microphone to multiple actuator configurations also possible.

FIGURE 5.2: Photographs of the robotic percussion array control unit including back (middle) and front (bottom) panels.

Figure 5.3 is an annotated photograph of the control unit taken from above, with the top panel of the enclosure removed to show the componentry inside. The left side of the image corresponds with the front of the enclosure, with the right side showing the back. On the left-hand side, the 16 preamp circuit boards can be seen mounted vertically, attached to the front panel by the machine screw connections of the Neutrik sockets and the potentiometer caps. The preamps are powered by the main power supply board in a daisy chain configuration starting from the bottom of the photograph to the top.

The line level analogue audio outputs from these preamps are routed to the blue DSP board shown in the top middle of the image, carried by white coloured hookup wire. The details of this board are presented in Section 5.5. Black hookup wire is used for the solenoid control lines from the DC barrel sockets to the green solenoid control board shown in the centre of the picture, which is described in Section 5.3. A short bidirectional UART connection is made between these two boards, and the USB, MIDI, and RS-232 connections are all made to the control board.

Since this is a 16 channel unit, two ADAT lines were chosen as a means of passing audio out of the device. Each of these lines is capable of streaming eight channels of 16 or 24 bit 44.1 kHz or 48 kHz digital audio via fibreoptic cable (Lightpipe). In the top middle of Figure 5.3, short lines of time domain multiplexed (TDM) audio are connected from the DSP board to a custom ADAT conversion board, which converts the data into the ADAT Lightpipe format that commercial audio hardware is capable of receiving. The design of this conversion board is outlined in Section 5.6.

The bottom middle and right of Figure 5.3 shows the power supply circuitry utilised to provide a number of different power rails at different currents to the various elements of the system. A 6 V AC, 1 A, EI style transformer feeds the red 5 V supply board which powers the digital circuitry of the unit. Another EI style 15 V AC, 1 A transformer provides ± 15 V rails for the analogue audio circuitry via the main power supply circuit board. A larger toroidal transformer with dual 60 V secondary windings also feeds the main power supply board, powering the solenoid outputs and the 48 V phantom power rail. The power supply architecture and circuit boards are described in more detail in Section 5.7.

FIGURE 5.3: Annotated photograph of the inside of the robotic percussion array control unit.

5.3 Actuator Control System

Creating expressive performances with musical robots generally involves dynamically varying the amount of power applied to actuators, when the power is applied, and the length of time for which the power is applied. The more accurately these variations can be programmed, the more precisely composers can write convincing and expressive performances for these instruments. When these parameters are less accurate, the result may be perceived as a variation in the robot's performance that makes the machine sound 'more natural' or 'more human', or conversely, when the parameters are unfavourably inaccurate, it can create negative issues with timing and dynamics.

It is for this reason that it is necessary to improve the accuracy of these signals, so that composers have the choice of whether or not to inject expressive characteristics into the performance of their pieces, and in which way. Unfortunately, due to the often high actuator count and signal processing requirements, conventional purely microcontroller-based methods of controlling actuators have proven insufficient at providing the accuracy required for precise and reproducible control. This is exemplified by the issues with the Closed-Loop Robotic Glockenspiel in which output pulse widths exhibited up to 100 microseconds of error, described in Section 3.6.

In order to address these issues, a new custom control board is created which separates actuator control from DSP tasks, dedicating hardware specifically to the task of creating accurate and consistent digital signals for controlling actuators. The design considerations of this board are outlined below, with the various elements and functionality described afterwards. Following that, the performance of the actuator control system is evaluated and comparisons are drawn between the new system and traditional methods.

5.3.1 Design Process

There are several design goals for this musical robot controller board. First and foremost, it should provide a large number of sub-microsecond accurate one-shot control pulses to an equal number of MOSFETs, and provide outputs to that many actuators. Comparable musical robotic actuator control systems such as Novalis provide up to 48 outputs [62]. Secondly, the board should be flexible as to which actuators, MOSFETs, and external hardware

can be connected to it, and thirdly, it should be straightforward to program, compatible with streamlined development environments such as Arduino and Mojo. This allows the board to be reprogrammed by non-engineers, improving the accessibility of the system. It should also be equipped with MIDI In, Out, Thru, and USB connections for control and feedback, along with options for other protocols such as Serial Peripheral Interface (SPI), I²C, and RS-232. Extra GPIOs and analogue inputs should be provided for expandability.

Since Chapter 4 found solenoids to be the most suitable actuator for general robotic percussion tasks, these are the control board's primary focus. The velocity response of solenoid actuators is highly non-linear, as discussed in detail in [59] and [110]. As a result, there are certain regions of the dynamic range which have increased significance, being particularly sensitive to minute variations in the width of control pulses. Unreliable pulse widths at the low end of the dynamic range are also particularly problematic, as they can result in potentially very musically important programmed notes not sounding at all. For these reasons, the accuracy, resolution, consistency, and reliability of control pulses are vital for creating musical robots with precisely controllable expressive ability.

While earlier methods often utilised Darlington transistors to switch current through and away from the electromagnets [59], this has since been superseded by the more efficient MOSFET-based circuit illustrated in [80]. The most common method of generating the required pulses that trigger those MOSFETs is utilising the GPIO pins of microcontrollers, as was implemented in the Closed-Loop Glockenspiel described in Chapter 3.

There are several potential approaches that could yield improved accuracy in this area. Utilising the timer peripherals of microcontrollers, configuring them to act as one-shot timers and routing their outputs directly to the GPIO pins is one approach. However, arrays of robotic percussion instruments require individual one-shot timers for every single actuator under control. A survey of available microcontrollers on the market did not produce an offering that provided the high number of high resolution timer peripherals that a controller of such an array would require.

Alternatively, multiple microcontrollers can be included, each controlling its own actuator, as was implemented by Craig Newswanger with his Percusibot Drum Controller Kit [112]. This solution carries a high cost, requires

significant board space, and necessitates individually programming each microcontroller separately for each change in firmware - a time consuming and error-prone workflow. Rather than individual microcontrollers, another possibility is utilising multiple programmable interval timer ICs such as the Intel 8254 mentioned by Raes [126]. However, these chips also consume a lot of physical space, are costly, and are inflexible in functionality. Other systems such as that described in [76] have forgone digital control, utilising analogue control voltages from synthesiser equipment instead. While this solution provides the lowest possible latency, achieving a high degree of temporal precision in these types of signals without digital control is impractical.

Upon investigation into these possibilities, it is decided that the most effective solution to achieve sub-microsecond accuracy of pulse width for a large numbers of channels simultaneously, independently of the work load of the rest of the system, is to utilise an FPGA. The FPGA could be configured to house as many high resolution hardware timers and other circuitry as required. The result is a single small board that can control large numbers of a variety of actuators with a high level of precision, while being cost-effective and easily reconfigurable. The following section describes the hybrid microcontroller-FPGA musical robot control board developed to implement this functionality.

5.3.2 The Circuit Board

This section outlines the design concepts of the robot control circuit board shown in Figure 5.4, complete with descriptions of the board's internal and external communication protocols and capabilities, how the board is programmed and powered, and how it can be used in a large variety of different musical robots.

The 130×100 mm board attempts to support as many use-cases as possible by making use of a modular design. This means that several different communication interfaces such as standard MIDI, USB-MIDI, UART, SPI, I²C and GPIO are included, and a channel count of 48 is offered. Any functionality that is not required may not be populated on the board. An example population is shown in Figure 5.5 where two banks of eight actuators are populated, and both USB and standard MIDI communication are utilised. This leaves the remaining four banks of buffers, MOSFETs, and flyback diodes unpopulated.

FIGURE 5.4: The new musical robot controller board. Fully populated rendering with main sections labeled.

FIGURE 5.5: Photograph of the new musical robot controller board with two banks of eight channels populated, as utilised in the new robotic drum kit controller.

5.3.3 Internal Communication

In response to external information received via the MIDI, GPIO, analogue inputs and other sources, the main microcontroller processes all of the control logic of the musical robot, including arranging the control of the actuators, processing any closed-loop input signals, and conducting any auto-calibration, latency compensation, and other functions. The Atmel Atmega1280 microcontroller is chosen due to its four UARTS, high number of GPIO ports, low cost, and direct compatibility with Atmel Studio, AVR-GCC, and the multi-platform and uncomplicated Arduino integrated development environment (IDE).

It is widely documented that latency is an important characteristic in the design and performance of digital musical instruments, and minimising latency is the subject of much research and development in the area [94]. In order to minimise the latency of the system, the main microcontroller communicates with the FPGA in a parallel configuration. The communication system, as shown in Figure 5.6, makes use of three 8-bit GPIO ports and an extra single pin to clock the data into the FPGA. This allows the microcontroller to transmit high resolution pulse width information along with corresponding actuator numbers in as little as eight clock cycles, or 1000 ns at 8 MHz. Additionally, a bidirectional UART connection between the microcontroller and the FPGA may also be utilised for less time-critical information transfers.

FIGURE 5.6: Diagram of the robot controller circuit board's signal flow.

5.3.4 Communication with the Outside World

As shown in Figure 5.4, traditional MIDI In, Out, and Thru connectivity is provided, adhering to the 2014 update to the MIDI 1.0 specification. MIDI In is primarily utilised for receiving musical control messages from a sequencer or live performance interface, MIDI Out can relay information about the robot and performance back to the sequencer, including notes played by

humans on the automated instruments, and the buffered MIDI Thru functionality can relay incoming commands to additional units should more than 48 channels be required. USB MIDI functionality is also provided on the top right side of the board by an Atmel Atmega8u2 microcontroller equipped with built in USB connectivity, loaded with a modified version of the HIDuino firmware [21].

16 general purpose analogue inputs line the top centre of the board, which may be used for reading sensors utilised in closed-loop control systems or other purposes. Additional bidirectional UART, SPI and I²C ports are broken out for external communication to and from the main microcontroller. An 8-pin GPIO port is also made accessible to accommodate potential future requirements. An RS-232 interface is included underneath the board for bidirectional communication with other devices outside the unit at baud rates higher than that of MIDI. In the case of this control system, the RS-232 interface is used to control the position of the hi-hat movement system described in Section 4.4.

5.3.5 The FPGA

In order to create high resolution, precisely timed signals for a large number of actuators simultaneously, a Xilinx Spartan 6 FPGA is used. The 6SLX9 device was selected due to its appropriate number of I/O ports, number of logic cells that exceeds the requirements of this project, and compatibility with both the Xilinx ISE tools and the streamlined Mojo IDE, with its simplified Lucid hardware description language.

The FPGA is highly flexible in that it can be reconfigured to provide control signals to drive a number of different musical robot actuators, from high resolution PWM signals for driving servomotors to DC and stepper motor control signals. It is also possible to precisely create complex signals such as dynamically changing PWM curves and pulse-wave audio tones for electromagnetic actuation techniques [95]. However, since the new robotic drum kit utilises solenoids, a configuration for solenoid control will be described henceforth.

While PWM is occasionally used in conjunction with solenoids, they are more commonly controlled by either one-shot or toggle style pulses. The castanet and egg-shaker robots described in [73] are examples of one-shot triggered

and toggle controlled instruments respectively. A custom configuration was designed in Verilog which contains an address decoder and an array of 48 one-shot timers, clocked by the FPGA's undivided 50 MHz oscillator, allowing the device to deliver pulses with a granularity of 20 ns. Upon receiving a positive edge on the Latch line, the corresponding output line is brought high, the 8 bit address bus is decoded, and the 16 bit pulse-width bus is loaded into the corresponding counter comparison register. In this case the pulse-width values are specified in microseconds, so the counter value is multiplied by a factor of 50 to match the 50 MHz master clock. The master clock then decrements the counter each cycle until it reaches zero, at which time the solenoid controlling output line is brought down to ground.

As the solenoids are powered by unregulated 60 V rails in this application, MOSFETs with a drain to source voltage limit of at least 100 V should be used to provide a safe overhead. These higher voltage MOSFETs require gate to source voltages of 5 V in order to achieve low R_{ON} levels, which allow most of the available current to be utilised by the solenoids rather than dissipated in the MOSFET. The output banks of currently available FPGA devices which are suitable for this application accommodate a maximum voltage rail of 3.3 V, so 3.3 V to 5 V translation is required.

As shown in Figure 5.6, the output banks of the FPGA are sent to up to six octal buffer chips which adhere to High-density Complementary Metal Oxide Semiconductor (HCMOS) logic standards, allowing 3.3 V logic to reliably register as 'high' in a 5 V powered buffer. These buffers serve the dual purposes of protecting the internal circuitry from potentially damaging voltage spikes from the outside, and lifting the logic level from 3.3 V to 5 V in order to activate these higher voltage MOSFETs. These 5 V levels are also potentially useful for controlling analogue servomotors and other devices that require 5 V control logic.

5.3.6 The Programmer

While the actuator control board is equipped with Joint Test Action Group (JTAG) pins that can be used to upload configurations to, and debug the on-board FPGA, a separate programmer is also created. This board is made to maintain compatibility with the Mojo IDE, and with it, the Lucid hardware description language, and to provide a low-cost and gentle learning curve option. The small 35×50 mm board shown in Figure 5.7 is built around an

In addition to the power that the electronics of the system requires, musical robots often require additional higher voltage supplies for their actuators. This can be just 5 or 6 V for servomotors, 12 or 24 V for stepper motors, or 48 or 60 V for driving solenoids with short pulses. In order to maintain flexibility and compatibility with a large number of musical robot systems, each MOSFET bank has independent higher voltage lines. This allows each bank to accommodate actuators with different supply voltage requirements, up to a maximum of six. However, if fewer or only a single higher voltage level is required, temporary or permanent jumpers can be placed between the supply lines of the banks, connecting them.

5.3.8 Evaluation

In order to evaluate whether or not the created system achieved the desired performance, a trial was conducted to determine the accuracy and consistency of the generated pulses. The output pulses were measured with an Agilent DSO-X 2024A digital oscilloscope at a sample rate of 1 Gigasample per second, and 100 measurements were taken to provide a margin of error for the trial.

FIGURE 5.8: Graph showing the accuracy of measured output pulses.

Figure 5.8 shows that of the 100 measurements of the pulses taken, 39 of them were measured to be of the correct width to the nanosecond, with

the rest of the measurements showing predominantly 1 and no more than 3 nanoseconds of deviation. This distribution can be explained by the fact that the measurement apparatus itself is accurate to the nanosecond and the unsynchronised phase relationship between the device under test and the measurement equipment's sampling. The accuracy achieved is two orders of magnitude higher than the required specification, and an improvement of five orders of magnitude over the Closed-Loop Robotic Glockenspiel.

This board is shown to achieve high levels of accuracy, consistency, and reliability for musical robot actuation, and presents a novel method of overcoming the difficulties experienced by previous systems by utilising the massively parallel capabilities of FPGAs. This implementation is capable of consistently generating control signals for musical robot actuators with a timing granularity of as low as 20 nanoseconds. This is a significant improvement over the microcontroller-only design utilised in Chapter 3's Closed-Loop Robotic Glockenspiel, which generates timing errors as high as 100 microseconds. The design goals outlined in Section 5.3.1 have been successfully achieved with a board that is more accurate than other systems, highly expandable, simple to use, and programmable by entry-level development environments.

5.4 Preamp Modules

While the analogue audio circuitry of the Closed-Loop Robotic Glockenspiel, described in Section 3.4.3, allows the unit to carry out its closed-loop continuous calibration functions, the noise floor of the instrument's audio is higher than desired. The requirements of expandability and flexibility for unpitched percussion arrays also means that a more modular preamp design is required. In order to accommodate the wide variety of transducers that may be used to receive audio from a range of percussion devices and achieve low-noise gain, custom preamp modules are created using low noise, low distortion instrumentation amplifier ICs. The top image of Figure 5.9 shows one of these modules fitted to a panel, and the bottom image depicts the circuit board. The full schematic of the preamp circuit board is also presented in Figure 5.10.

The preamp makes use of a TRS / XLR Neutrik input and can take balanced and unbalanced audio from a wide range of sound sources. The mounting

FIGURE 5.9: The custom preamp used to condition the audio signals. Top: Behind a panel, Bottom: Bare circuit board.

holes of this connector are also used to affix the preamp circuit board to the panel. An SP3T bias switch is mounted to the board so that it barely protrudes past the 3 mm thick panel, preventing accidental movement. This switch allows the user to activate different bias configurations with options for no bias for dynamic microphones, 9 V bias for electret or plug-in-power (PiP) microphones, and 48 V phantom power for condenser microphones.

A red -20 dB pad 4PDT toggle switch is provided to reduce the level of input signals by way of a differential voltage divider. This prevents distortion from occurring for signals that require less gain, and aids with gain staging in the system. A black 4PDT toggle switch is then included to route the differential signals through standard impedance AC coupled pathways, or circuitry that handles high impedance input signals. This provides support for magnetic pickups, coils, piezoelectric discs, and other such high impedance (Hi-Z) sources.

The signals are then routed to the positive and negative inputs of a Texas Instruments INA217 Low-Noise, Low-Distortion Instrumentation Amplifier. The gain setting pins of the INA217 are connected to the panel-mounted potentiometer, which dictates the level of gain applied to the signal before it is output from the board. An OPA137 FET input opamp is utilised to drive the INA217's reference input in a feedback loop in order to eliminate any potential output offset voltage.

A green LED is illuminated by the audio signal via an LM358 opamp configured as an amplifier with a gain of 10:1. This helps confirm to the user that the signal is being correctly received by the preamp from the utilised transducer, and the intensity of the glow visually reflects the level of that signal. A red LED is also driven by the other opamp in the LM358, except this one is configured as a comparator, with the LED shining only when the signal present at the positive input rises above that specified at the negative input by a voltage divider. When it does shine, it shines with full strength rather than a graduated response. This acts as a warning to the user that the gain is set too high, and may be distorted in the system. In the case of this robotic drum set controller, the voltage divider of the hard clipping indicator is tuned just below the clipping level of the ADCs on the DSP Board described in Section 5.5.

The result of these developments is a new preamp module that is specifically designed for implementing closed-loop audio feedback in robotic percussion arrays for performance and installation contexts. The slim form-factor allows

FIGURE 5.10: Schematic of the preamp module.

arrays of these units to be populated in a control system without sacrificing the flexibility of individual pad, Hi-Z, gain, and bias settings for each channel on the panel itself. Dual power connectors also allow the unit to be powered in a daisy-chain configuration, reducing the complexity of installation. The flexibility of this implementation allows this percussion array control system to utilise many different sound sources with different impedance and biasing requirements, boosting all of them to levels that the DSP board's analogue to digital converters require.

5.5 Multi-Channel Embedded MIR Board

The primary goals of this thesis involve making musical robots more reliable, integrated, and expressively capable, as discussed in Section 1.3. To achieve this, the new musical robots should be provided with the ability to listen to their own audible output, and automatically make adjustments to their playing technique to refine their musical actions. In order to implement this closed-loop audio feedback as defined in Section 2.5, it is necessary to equip the instruments with audio sensors such as pickups, coils, and microphones, audio acquisition circuitry, and digital signal processing hardware. In addition, in order to make musical sense of the audio that is being received, musical information retrieval algorithms must be implemented in the firmware running on that hardware, to extract pertinent data such as onset times, latency of actuation, envelope shape, and amplitude, among other features.

While a number of attempts have been made to create this functionality with commercially available hardware and personal computers, each of these options suffer from drawbacks that limit their utility in actual performance environments, as discussed in Section 2.6. The Closed-Loop Glockenspiel described in Chapter 3 was the first such device created with these capabilities integrated inside the instrument, and is an example of how these types of systems can be utilised in robotic polyphonic percussion instruments. Unpitched percussion arrays, like the new robotic drum kit introduced in Chapter 4, have different hardware and firmware requirements to the glockenspiel, so a novel solution is required for the closed-loop implementation. This section presents a new custom multi-channel embedded DSP system designed to address these issues. It starts by determining the requirements of the new embedded MIR system and using these specifications to guide the

selection of the various parts of the board. The development process is then outlined and the various subsystems of the design are described in detail.

5.5.1 Design Considerations

There are several ways in which the audio processing requirements of robotic unpitched percussion arrays differ from those of robotic polyphonic percussion instruments. Firstly, whereas the number of notes in a polyphonic percussion instrument is fixed, the number of channels of audio in unpitched percussion arrays may vary widely between small and large installations. In addition, the analytical requirements of the audio of each channel may also vary significantly, with some instruments requiring amplitude and latency detection, while other instruments may require other processes such as pitch detection or spectral analysis.

In addition to customising the solution to the needs of unpitched percussion arrays, some of the issues that were encountered with the implementation of the Closed-Loop Robotic Glockenspiel, described in Section 3.6, must be addressed and improved upon. In order to reduce extraneous noise on the audio lines, a high quality audio path, defined as CD quality 16 bit resolution and 44.1 kHz sample rate, is specified for the audio inputs and outputs. Instead of passing through the analogue audio directly, a high quality digital audio path allows the possibility of the DSP implementing functions such as digital noise gating and other programmable and musical content-informed noise reduction processes to be applied to the audio before it is passed out of the unit.

The onset detection of unpitched percussion instruments is also not able to take advantage of prior knowledge of the fundamental frequency of the expected audio in the way pitched percussion can. This means that the greater dynamic range and reduced noise floor of 16 bit audio is required in order for the onset detector to provide adequate performance during continuous real-time self calibration. With these considerations in mind, the following is a summary of the key capabilities with which this new embedded DSP board must be equipped:

- up to 32 channels of high quality and low latency audio input, to accommodate large installations or polyphonic instruments.

- a digital signal processor (DSP) capable of carrying out common MIR processes such as Finite Impulse Response (FIR) filters at sufficient sample rates.
- up to 32 channels of digital audio output, for forwarding the acquired audio to subsequent processors, digital audio workstations or debuggers.
- up to 32 channels of high quality and low latency analogue audio output, for use in distributing the acquired audio or driving closed-loop actuated or magnetic resonance instruments.
- a modular architecture to facilitate smaller or larger systems.
- a capability to be programmed in an accessible environment without requiring expensive proprietary development environments and hardware.

Figure 5.11 illustrates the signal flow of the system. A number of channels are processed by a DSP, and bidirectional serial communication between the DSP and outside systems enable it to share the musical information that has been retrieved from the audio streams with MIDI devices, solenoid control boards, or other hardware.

FIGURE 5.11: Signal flow diagram of the new embedded MIR system.

5.5.2 Part Selection

As mentioned in the previous section, high quality audio is regarded as utilising a sample rate of at least 44.1 kHz, and a bit depth of at least 16 bits, equal to CD audio quality. Low latency is considered to be in the sub-millisecond range in order to facilitate a responsive control loop frequency of at least 1 kHz. These requirements eliminate the possibility of choosing general purpose microcontroller boards, as their on-board ADCs and DACs rarely meet

these bit depths. The low latency requirements also eliminate the use of commercial laptops and audio interfaces whose drivers and operating systems generally require larger block sizes. The Bela cape for the Beaglebone Black single-board computer fulfills the latency and audio quality requirements, but lacks the number of channels required [94].

As a result, it is decided that creating a board based on an embedded digital signal processor that is optimised for high channel counts of high quality, low latency audio is the preferred solution. Several manufacturers such as Texas Instruments, Analog Devices, and Motorola offer microprocessors which are specifically targeted at audio digital signal processing. Though these chips generally utilise clock frequencies lower than single-board general-purpose computers, they are equipped with highly optimised instruction sets and peripherals that allow the processor to undertake several tasks related to processing audio in a single instruction cycle, as well as many related tasks in parallel.

After examining the architectures, peripheral sets, and characteristics of a range of digital signal processors, it was decided that the Analog Devices ranges of Blackfin and SHARC (Super-Harvard Architecture) chips would suit the application. Before committing to a single DSP, an applicable chip from each of these ranges was chosen for evaluation, with ADSP-BF592 chosen from the Blackfin range, and ADSP-21478 chosen from the SHARC range. The evaluation boards available for these chips from the manufacturer are very highly priced, with a large range of functionality that is not applicable to this project. As a result, small 5×5 cm two-layer PCBs were created with just the circuitry required to power, program, and debug the chips, with their I/O pins broken out to columns of headers that allow them to be plugged directly into breadboards. This allows a range of different ADCs, DACs and other hardware to be trialled with the chips via the breadboard.

These custom evaluation boards, shown in Figure 5.12, are driven by 5 V power rails that are regulated down to 3.3 V for the main supply, and the 1.4 V or 1.2 V core voltages required by the Blackfin and SHARC chips respectively. They also use reset supervisor chips that ensure that the power rails are correct before releasing the reset lines, and both are equipped with external M25P16 16 Mb Serial Flash Memory chips for firmware storage. Both boards are also equipped with JTAG interfacing circuitry, reset push-buttons, power and user controllable LEDs, and all additional pins broken out for easy access.

FIGURE 5.12: Custom Blackfin (left) and SHARC (right) development boards.

While both architectures are capable of fulfilling the functionality required, the ADSP-BF592 was selected for several reasons. The BF592 runs at a clock speed of 400 MHz, and has dual 16 bit multiply-accumulators (MACs) for up to 800 mega MACs (MMAC). It has the necessary communication peripherals for transmitting the results of analyses to external hardware, complete with a UART, I²C compliant two wire interface (TWI), and two SPI compatible ports. More significantly, it is equipped with two dual-channel, full-duplex synchronous serial ports (SPORTs) which are capable of supporting eight stereo I²S (Inter-IC Sound) channels, each supporting TDM with DMA. This results in a maximum 32 channels of audio input and 32 channels of audio output with very little processor overhead. It is also equipped with 64 Kilobytes of L1 instruction ROM, which contains user-accessible built-in optimised algorithms such as FFT and FIR filters.

While the SHARC chips also possess many of these peripherals and characteristics, the BF592 is also capable of booting from its UART port, and is compatible with the open source GNU Blackfin toolchain. This means that it is possible to program the chip using an inexpensive USB to UART adapter, without the potentially prohibitive cost involved in purchasing proprietary software development environments and in-circuit emulators. The cost of the Blackfin chip is also significantly lower than the SHARC for equivalent MMACs. These points make the BF592 a chip that is more accessible, one of the objectives of this thesis specified in Section 1.3.

There are a large number of ADCs and DACs available that are suitable for high quality, low latency audio. In the case of this system, in order to be

compatible with the choice of DSP, the ADCs and DACs must be capable of utilising eight-channel time domain multiplexed digital audio, and provide either I²C or SPI control interfaces. The Cirrus Logic CS5368 was chosen as the ADC, as it fulfills these requirements and is capable of sample rates of up to 192 kHz at up to 24 bit resolution. For the DAC, the Analog Devices AD1934 was selected with similar specifications.

The front and back of the resulting circuit board is depicted in Figure 5.13. There are a total of 1085 components on board, and each segment of the system is described in turn in the following subsection.

FIGURE 5.13: Rendering of the front (left) and back (right) of the circuit board.

5.5.3 The System's Hardware

The circuitry of the embedded MIR system is housed on a single, four-layer circuit board 100 mm in width and 150 mm in length. The layers consist of two signal layers, one ground plane, and one power plane split into several sections. The signal flow is structured similar to Figure 5.11 in that inputs are received on the left side, processed in the middle, and output on the right. These three stages are outlined in the following subsections, while referring

FIGURE 5.14: Annotated photograph of the top of the system's circuit board with circuit sections colour coded.

to the annotated photograph of the top side of the printed circuit board (PCB) provided in Figure 5.14. Since the ADCs and DACs are the most expensive components of the board, unused banks of inputs or outputs may be left unpopulated, saving monetary cost. The photograph of the inside of the robotic drum kit control system enclosure presented in Figure 5.3 shows a version of the board in which 16 channels of ADCs were installed, and no DACs.

Input Stage

The first section of the board contains 32 analogue inputs split into four, 8-contact Molex connectors. Banks of ground pins are interspersed with these connections to facilitate analogue audio connections with external equipment. These audio signals are then AC coupled, and passed through a conditioning stage based on the LMV722 dual op amp. This conforms the input

signals to the ADC's levels, and converts the single-ended inputs to the differential format that the ADC expects. Each ADC is configured via I²C on startup to run in 256 f_s mode (f_s being the sample rate), and clock their eight channels of digital TDM audio to the central processor, with bit and word clocks distributed by the DSP.

Central Processor

The circuitry shown in the top middle of Figure 5.14 provides regulated 3.3 V and 1.4 V power for the external and core supply lines respectively, and a reset supervisor IC ensures that the unit only boots once both supplies are stable. A jumper provides the options of either booting directly from UART or from the 16 Mb SPI flash memory chip on the board. Programming the DSP and the external flash memory can be achieved either via JTAG using Analog Devices' proprietary tools, or via UART using open source tools. A DPDT switch at the bottom of the board switches the UART connections between the programming port, which is equipped with MOSFET-based voltage converters for 5 V compatibility, and separate communication lines which may be used in the application. Further details about the booting process are provided in Section 5.8.1.

Upon booting, the DSP's firmware assumes control and proceeds to initialise its various peripherals before configuring each available ADC and DAC on the board for operation. This process takes just a few milliseconds and is essentially instantaneous from the perspective of the user. The chip then begins the continuous transfer of streams of audio in and out by way of DMA, with up to 32 channels of audio streamed in and out of the chip via the peripherals, without any load on the core's resources. The core is then available to conduct MIR, digital audio processing, synthesis and any other operations specified in firmware. The firmware used in the closed-loop audio-feedback based control of the new robotic drum kit application is described in Section 5.8.

Output Stage

The TDM output from the DSP's SPORTs is routed simultaneously to two destinations: the four AD1934 octal DACs, and four three-pin Molex connectors to their left in Figure 5.14. This allows both line level analogue audio

output directly from the board and lossless digital audio output to external circuitry. The analogue outputs from the DAC are then passed to an array of TL074 Quad Low-Noise JFET Input Op Amps, running at ± 15 V, that amplify the signals and provide low impedance, AC coupled audio via the banks of outputs on the right.

5.6 Digital Audio Output Board

The composition of *Broken Mirror* (described in Section 3.7.1) for the Closed-Loop Robotic Glockenspiel relied on the instrument's ability to pass the audio signals that the closed-loop control system acquired out of the unit for creative processing. It is desired for the robotic unpitched percussion array control system to be equipped with the same ability, though a different approach is required. Since this control system handles an array of different percussion instruments, rather than many notes of a single instrument, it is important that each of these audio channels can be output from the unit individually.

In order to reduce levels of extraneous noise and maintain a lossless signal path, digital methods of audio transfer are desirable. There are several digital audio transmission standards in use in professional audio equipment, such as AES/EBU or S/PDIF for one or two channel signals, ADAT for up to eight channel signals, MADI (Multichannel Audio Digital Interface) for up to 64 channel signals, AVB (Audio Video Bridging) for over 100 channels, and a number of other proprietary formats. This robotic unpitched percussion array control unit is configured for up to 16 channels of audio input, so ADAT (Lightpipe) is selected as the most appropriate format for digital output.

The ADAT protocol was patented in 1994 by the Alesis Corporation [5], and Wavefront Semiconductor is licensed to manufacture the AL1401AG (Opto-Gen) encoder, and AL1402G (OptoRec) decoder ICs utilised in implementing the protocol. Figure 5.15 shows a custom circuit board that was created to evaluate these chips and develop an ADAT interface. The board is powered by 5 V, and is split into input (right) and output (left) halves. The only connection between the two halves is the 5 V rail, so it is possible to cut the boards apart down the split point and create various combinations of ADAT input and output circuits.

FIGURE 5.15: Rendering of the first iteration of a custom ADAT I/O board.

Each side of the board is equipped with a power indicator LED, and the receive side also uses an extra LED to indicate whether an error has occurred in reception. All of the pins of each chip are broken out into 2.54 mm spaced pads so as to be compatible with breadboard and standard headers, and the board allows fixed parameter setting by attaching resistors or solder bridges to the 0603 sized footprints marked with silkscreen artwork.

After trialling this solution, a number of issues were encountered. Firstly, the AL1401AG and AL1402G parts are now categorised as ‘end of life’ parts not recommended for new designs, meaning that they are not a sustainable solution going forward. They also run with 5 V logic levels that are not compatible with modern DSP hardware, requiring extra level shifters to work reliably. In addition, they each utilise four streams of I²S digital audio to transfer audio information from DSPs and other audio circuitry, and the ADSP-BF592 SPORT peripherals are not capable of generating that number of I²S streams without external TDM to I²S conversion.

For these reasons it is decided that rather than utilising the officially licenced implementation, a custom solution should be designed that is capable of natively receiving TDM streams at a 3.3 V logic level directly from the DSP

board, and converting those streams into ADAT signals for forwarding out of the unit. A number of commercial businesses create integrated circuits capable of implementing the custom logic required to undertake this task, and the three largest FPGA manufacturers (Xilinx, Altera, and Lattice) all offer inexpensive parts that are capable of carrying out these operations. The fabless semiconductor company XMOS also offers multicore microcontrollers that are optimised for implementing custom ‘soft-peripherals’ that can be programmed to carry out such functions.

As each of these solutions is capable of implementing the required TDM to ADAT conversion, selection for this application is then based on the price point of the silicon, price and complexity of hardware implementation, and capability and subjective evaluation of the manufacturer’s synthesis software suite. As evaluation boards for these chips are priced very highly, a custom 10×10 cm two-layer evaluation board was created that splits into the four 5×5 cm boards shown in Figure 5.16. Similar to the custom DSP evaluation boards described in Section 5.5.2, these boards provide power conditioning, programming, and persistent memory circuitry to the target chips, and break out the remaining pins of the chips so that they can be accessed in the application. In addition, these boards also include user accessible push-buttons and LEDs, and Lightpipe transceivers for ADAT and S/PDIF applications.

The Lattice ICE40HX1K on the top right of Figure 5.16 is the least expensive of the devices, though it also has the least internal resources, which results in less expandability. Separate power conditioning circuitry is required for the core (1.2 V) and I/O banks (3.3 V), and external flash memory is necessary for persistent configuration storage, adding to the cost and complexity of implementation. In addition, Lattice’s ‘iCEcube’ development environment lacks much of the advanced functionality of the software supporting the other chips.

The Xilinx Spartan 3 XC3S50A on the top left of Figure 5.16 is second-to-least expensive of the devices, with more internal resources than the ICE40HX1K, but with almost identical supporting circuitry required. Synthesis for this chip is performed by the Xilinx ISE Design Suite, and while it is not as modern as the Vivado Design Suite utilised with their more expensive chips, it is much more feature-complete than Lattice’s ‘iCEcube’.

The Altera MAX10 FPGA 10M02SCE144I7G is shown on the bottom left of Figure 5.16. While this chip is slightly more expensive than the Lattice and

FIGURE 5.16: Photos of custom evaluation boards for ADAT interfacing. Top left: Xilinx Spartan 3 XC3S50A, top right: Lattice ICE40HX1K, bottom left: Altera MAX10 FPGA 10M02SCE144I7G, bottom right: XMOS XE-216-512-TQ128.

Xilinx chips used in this comparison, it is much more integrated, with persistent flash and core voltage regulation functionality provided inside the chip. This means that only 3.3 V regulation is required to run the circuit, resulting in a much simpler implementation, and lower cost of external hardware. The chip's internal resources are also greater than the Xilinx and Lattice devices trialled, and its 144 pin package provides a greater number of I/O lines. Altera's Quartus Prime software used as the development environment for the chip is also the most modern and feature-rich of the environments evaluated.

The XMOS XE-216-512-TQ128 16 core microcontroller is shown on the bottom right of Figure 5.16. While not an FPGA, this device provides high-speed programmable I/O pins and the capability of dedicating entire cores to the task of implementing specialist protocols such as ADAT. While the xTIME-composer Studio software used for developing with this chip is modern and capable, and the resources available to the chip are high, it is also the most expensive chip of the group. In addition, it requires external flash memory, multiple regulated power rails, and complex power sequencing, making it the most complicated and expensive implementation.

Through the comparison of the cost of the chips and necessary external circuitry, the resources available for future expandability, and the qualitative evaluation of the provided software suites from the various manufacturers, the Altera MAX10 FPGA 10M02SCE144I7G was chosen as the most appropriate solution for this application. The board is affixed to the robotic unpitched percussion control unit's back panel via machine screws threaded through the board's Lightpipe transmitters as shown on the top left of Figure 5.3, and TDM signals are routed from the DSP PCB to this board.

The TDM to ADAT conversion process loaded into the MAX10 chip consists of several sections: TDM decoding, ADAT formatting, serialising, and NRZI (non-return-to-zero inverted) encoding. First, the chip takes in a word clock input at the sample rate (f_s), a bit clock input at $256 f_s$, and two TDM data inputs, each carrying eight channels of digital audio clocked at $256 f_s$. Custom TDM reader modules sequentially sample the TDM streams to dedicated registers, decoding the audio information into individual 24 bit words. This information is formatted into 256 bit ADAT words via combinatorial ADAT formatting modules. Serialising modules clock those 256 bit ADAT formatted words out one bit at a time, and NRZI modules encode the signal for output directly to the Lightpipe transmitters. The result is a lossless digital stream of audio that is converted with just a single sample of latency,

capable of transmitting eight channels with a single cable (16 channels with two cables), to most modern audio interfaces, digital mixers, and other audio hardware for processing, amplification, or other functions.

5.7 Power Supplies

The robotic unpitched percussion array control unit utilises several different subsystems with a range of power requirements. Since one of the goals of the project is a high level of integration, it is desired that the only external power line is a standard IEC cable providing mains power. This means that the various regulated and unregulated power rails of different voltage levels and current requirements must all be derived by circuitry inside the unit from this single mains connection. Figure 5.17 illustrates the power architecture of the unit.

The IEC input socket used for mains power is equipped with an illuminating switch to turn the unit on and off, and a 10 A fuse. The connections made between the IEC socket and the primary sides of the EI transformers inside the enclosure are physically isolated from the rest of the unit by way of a 3D printed bracket which covers the high voltage sides of the transformers and the socket. This ensures that even if the enclosure is open while powered, tools, fingers, and other miscellaneous items will not come in contact with the high voltage lines. The primary cabling of the toroidal transformer is wired through very small gaps in the mains protection bracket, and the mounting machine screws of all the transformers are connected to mains earth.

The two 60 V and one 9 V secondaries of the toroidal transformer, and the 15 V secondary of the 15 V EI transformer, are wired to the main power supply board, whose image and schematic are presented in Figure 5.18. The 60 V lines are both passed through Positive Temperature Coefficient (PTC) resettable fuses that are rated at 3 A before undergoing full wave rectification via sets of 1N5404 high current diodes arranged in a bridge rectifier configuration. Since it could be possible for the solenoid control board to demand current loads higher than what the transformers and other components are rated to handle, these PTCs temporarily bring down the respective solenoid supply in over-current situations, automatically restoring the current after a time constant when the flow is reduced.

FIGURE 5.17: Block diagram of the power supply and distribution of the robotic unpitched percussion array control unit.

FIGURE 5.18: Photograph (top) and schematic (bottom) of the Main Power Supply board.

After rectification, the two 60 V supplies are smoothed with pairs of 2200 μF electrolytic capacitors rated at 100 V, and made available at the main power supply board's outputs. In the robotic unpitched percussion array control unit, each of these supply rails powers half of the solenoid outputs, balancing the load on the transformer's secondary windings, and allowing at least some instruments to keep playing in the case of an over-current situation on one of the rails' PTCs. One of the 60 V rails is also utilised for deriving the 48 V phantom power required for the use of commercial condensor microphones. This regulation is carried out by a TL783 High-voltage Adjustable Regulator with which the 48 V output rail is configured with a pair of resistors, and sent to the array of preamp boards for biasing the microphone inputs of channels that have phantom power selected.

The 15 V AC input with centre tap is rectified with lower current 1N4007 diodes, separated into positive and negative rails, and smoothed with large electrolytic capacitors. Both rails are then regulated with 7815 and 7915 positive and negative voltage regulators respectively before being made available at the outputs of the main power supply board for powering the dual rail analogue circuitry of the preamp boards and the DSP board. The power indicator LED is run from the -15 V rail as this is the least loaded. A 7809 positive voltage regulator also derives a 9 V rail from the 15 V rail, and this is sent to the preamps to be used as a selectable plug-in-power (PiP) for biasing electret microphones. A 5 V rail is then derived with a 7805 positive voltage regulator and made available at additional outputs.

Due to the cascading configuration of the 15, 9, and 5 V regulation of the main power supply board, high current loads should not be applied to the 5 V outputs on that board, as this current is drawn through all three regulators. As a result, in order to power the digital circuitry of the solenoid control board, the DSP board, and the ADAT conversion board, a separate 5 V supply is included. An annotated photograph and schematic of the dedicated 5 V power supply board are shown in Figure 5.19. It draws its input from the centre-tapped secondary coil of the 6 V transformer, is full-wave rectified with a pair of 1N4007 diodes, smoothed with 4700 μF electrolytic capacitors rated for 50 V, and regulated with a 7805 positive voltage regulator. The board is equipped with a power indicator LED, and four regulated 5 V outputs. The boards with digital logic then derive their own 3.3 V, 1.4 V and 1.2 V rails from these 5 V outputs with additional voltage regulators on the respective boards.

FIGURE 5.19: Photograph (top) and schematic (bottom) of the 5 V Power Supply board.

5.8 MIR Firmware

The heart of the robotic percussion array control unit is the closed-loop audio feedback system that receives the audio output by the sounding objects used and interprets the relevant musical information from that audio in order to inform subsequent actuation commands. This whole procedure is carried out by the combination of the DSP and the solenoid control boards, with the DSP board primarily in charge of deriving musical information from the audio signals, and the solenoid control board in charge of utilising that information to continuously calibrate the actuations of the robotic instruments. Bidirectional serial UART communication between these two boards allows the solenoid control board to relay configuration and control information to the DSP board, and the DSP board to send the results of its MIR analyses to the solenoid control board.

This section describes in detail how the firmware of both boards make the system work. It begins by explaining how programming and booting the DSP hardware is carried out and how both its internal and external peripherals are configured. It subsequently explains the onset, latency, amplitude, and envelope length detection algorithms that are used in the robotic percussion array control unit, and how the results of these real-time analyses are used to continuously calibrate the various characteristics of the system to produce more consistent and reliable robotic percussion instruments.

5.8.1 Booting the DSP

As outlined in Section 5.5.2, a significant feature of the ADSP-BF592 chip used in the DSP board is that it is compatible with the open source GCC compiler and can be programmed via its UART port without requiring costly proprietary compilers or programming hardware. Compiling can be carried out via command line software provided by the Bfin-ELF toolchain, which compiles an executable and optionally a loader file for permanent storage. This loader file can then be uploaded to the BF592 via a standard USB to UART interface using the BlueBoot software, as outlined in [55]. This entire process can be automated with a simple makefile for one-command programming.

The BF592's booting mode is specified by the logic level setting of its external BMODE2-0 pins, with a value of 0b100 assigned to 'SPI0 master boot from flash', and 0b110 assigned to 'Boot from UART host device'. This means that

an SPST switch or jumper on BMODE1 can be utilised to switch between the two modes. Quick programming can be done by resetting the chip, and commanding BlueBoot to compile the program and stream it to the chip via UART. This configuration is useful while developing with the chip, as new iterations of firmware can be compiled and run conveniently, though the uploaded firmware is not stored between power cycles. For persistent storage, boot mode 0b110 is also used; however, BlueBoot is instead commanded to upload a small kernel to the chip which then receives the loader file containing the firmware to be stored via the UART, and writes it to an external flash memory device connected via SPI. After this process is complete, when the chip is powered up or reset with boot mode 0b100 configured, it will automatically load and run its firmware from the external flash memory.

5.8.2 Configuring the Peripherals

Upon booting, the first subsystem that the firmware must configure is the onboard phase-locked loop (PLL) unit. This device receives a clocking signal from an external crystal oscillator, multiplies the frequency of the signal to the core clock, and subsequently divides it down to the system clock. In order to produce a sample rate for the entire DSP board of 44.1 kHz, an 11.2896 MHz crystal oscillator (which is 256×44.1 kHz) is utilised as the clock of the DSP, the ADCs, and the DACs on the board. This frequency also corresponds with the bit clock for the TDM streams, whose words consist of eight 32 bit values.

The ADSP-BF592 has a maximum core clock speed of 400 MHz and a maximum system clock speed of 100 MHz, therefore the signal from the crystal oscillator must be multiplied by the PLL to the highest value below 400 MHz for the core clock, and divided to the highest value below 100 MHz for the system clock. The SPORT peripherals then divide the system clock down to the 11.2896 MHz bit clock and the 44.1 kHz word clock. A multiplier of 35 would provide a core frequency of 395.136 MHz. However, since the SPORT is only capable of dividing the system clock by an even number, the highest available system clock for a bit clock of 11.2896 MHz is 90.3168 MHz, a multiplier of eight. This then means that a core clock PLL multiplier of 32 should be used to achieve a core clock of 361.2672 MHz.

Following the clock configuration, the rest of the on-board peripherals are initialised. The first of these are the GPIO ports, where a single pin is utilised

as an output to drive an on-board LED, and two extra pins are broken out on the board for debugging and expansion purposes. Each of these pins are configured by clearing the corresponding bit in their Function Enable Register to ensure that no peripheral functions are mapped to that pin, and the corresponding bit is set in their I/O Direction Register to enable output.

In cases where DAC functionality is required, the SPI peripheral must be initialised in order to communicate with and set the initial configuration of the AD1934 octal DAC units. However, since the DACs are not populated in the robotic unpitched percussion array control unit, the SPI and AD1934 configuration is omitted. Instead, the I²C peripheral is initialised in order to communicate with the CS5368 ADCs. A FIFO (First In First Out) buffer is established for the I²C transmissions, and a low-priority interrupt vector is assigned to the I²C interrupt handler routine, which services the peripheral. The I²C reference clock is derived from the system clock and should be as close to 10 MHz as possible without exceeding it, so a divider of ten is used for a reference of 9.03168 MHz. This reference frequency is then used to clock the high and low portions of the I²C clock, specified using the SCL Clock Divider Register in order to set both the pulse width of the SCL clock and its frequency (400 kHz in this case).

With the I²C ready for transmission, the CS5368 ADCs can be configured. The CS5368s are each equipped with two externally-set address pins on the circuit board, allowing four individual addresses for these ICs per I²C bus. I²C writes to the configuration registers are addressed to the ADCs' base address (1001100) plus the value applied to the address pins. The configuration begins by setting the ADCs to control port mode, the clock mode to 256 f_s , the data format to TDM, and the device to slave mode, to be clocked entirely by the BF592. The initialisation routine also activates internal high-pass filters on all of the channels to ensure proper AC coupling, and ensures that all channels are powered and un-muted.

In order to establish communication with the solenoid control board, the UART peripheral is also configured. A DPDT switch on the circuit board connects the UART port either to the programming connector via logic level translation MOSFETs, or directly to the serial communications port for communication with the solenoid control board. Similar to the I²C peripheral, the UART is configured for use with a FIFO buffer serviced by a low priority interrupt. The UART also derives its baud rate from the system clock, and is configured for a rate of 115200 baud. The format of communication is 8

bits with one stop bit and no parity bits. After the peripheral registers have been configured, initialisation is completed by mapping the peripheral's I/O to the BF592's external pins by enabling the corresponding Function Enable Registers, and selecting the UART peripheral in the function multiplexer.

The final peripheral to be initialised is the serial port (SPORT). The BF592 is equipped with two full duplex SPORTs that both have primary and secondary data channels, all capable of TDM audio streaming. This allows the device to receive four eight-channel streams and send four eight-channel streams, for a total of 32 channels of audio in and out. When the DSP board is fully populated, all of these channels are utilised, but in the case of the robotic unpitched percussion array control unit, just the primary and secondary transmitters and receivers of the first SPORT are configured for use. The SPORT is configured for 16 bit operation, 44.1 kHz word clock, 11.2896 MHz bit clock, and I²S style TDM operation. A high priority interrupt routine is activated at every word clock, the SPORT delivers its received samples, and DMA is configured to shift these samples from the SPORT FIFOs to user accessible memory automatically. This audio-rate interrupt is also where all of the MIR analysis takes place on the samples. An illustration of the audio-rate routines carried out by the DSP board is presented in Figure 5.20, and the analyses utilised in the robotic unpitched percussion array control unit are described in the following subsections.

5.8.3 Onset Detection

Onset detection is the process of determining the temporal start point of a musical event in a stream of audio. This start point can be utilised in subsequent analyses such as establishing the latency, amplitude, and envelope length, as described below. While the onset detection algorithm utilised in the Closed-Loop Robotic Glockenspiel and described in Section 3.4.4 made use of prior knowledge of the periodicity of the incoming signals, the robotic unpitched percussion array control system must handle instruments of unspecified pitches or no discernable periodicity.

In order to provide the DSP board with information as to when to expect a strike, the solenoid control board transmits a serial message to the board via UART whenever a strike begins. The DSP board then allows a short 20 ms stand-down period to pass before beginning to listen for a sharp increase in the amplitude of the signal above a set threshold. This delay acts to exclude

FIGURE 5.20: Flow diagram of the audio rate analysis routine running on the DSP board.

any interference from the electromagnets on the audio lines from being included in the analysis, and is set to be greater than the length of any pulses sent to the solenoids but shorter than the time taken for any strike to occur. The threshold is hard coded to sit just above the noise floor, plus a safety buffer to minimise the effect of bleed. During the window when a strike is expected, the onset is recorded as the time the amplitude threshold is crossed, and if the threshold is not crossed in more than 50 ms, it is considered that the instrument was not struck. The 50 ms value is determined by the longest expected latency of any striker plus a buffer.

This type of simplified onset detection can be effective because of the low noise floor of the high resolution ADCs and low noise instrumentation amplifiers used. However, since the audio of the instruments of a percussion array are likely to be picked up by microphones of some description rather than magnetic pickups, the potential for bleed to interfere with the continuous automatic calibration is increased. There are several approaches to minimise such bleed and reduce the issues caused by any remaining bleed that may be picked up. Firstly, bleed can be minimised by utilising various microphone placement strategies such as internal and close mic'ing as described in Chapter 4. When setting up the system, the level of bleed can be evaluated by carrying out loud strikes on each of the instruments and using the green signal indicator LEDs on the preamps to check for bleed signal on channels other than the intended one. By setting the onset detection amplitude threshold at the same signal level that activates the green LEDs, it can be conveniently ensured that the bleed level is acceptable if the only glowing green LED is the one associated with the target channel. If specific channels are causing unavoidable bleed issues with other channels, these channels can be marked so that analyses on the receiving channel can be automatically cancelled when the offending channels are actuated in close proximity before or after them. If certain channels possess incurable bleed issues, it is also possible to revert to automatic pre-performance calibration for those specific channels.

5.8.4 Latency Detection

While latency of audio processing is a significant issue in the wider area of digital music creation, its effects on musical robots are perhaps more complex than the wider field. In the case of robotic percussion arrays, a variety

of actuators and mechanisms are utilised, and each of these devices take different amounts of time to actuate their sounding object. This has the potential to create issues where strikes that are programmed to occur simultaneously are staggered, straight rhythms may sound swung, and other problems with timing can occur. In addition, each velocity value of each instrument creates a different amount of latency, with softer strikes taking longer to occur than stronger ones.

Without a system that can predict accurately when a human is going to command a strike, in performances where each individual actuation is triggered by a human live in real time, this latency cannot be avoided. However, many performances involve musical material that has been composed, sequenced, or generated ahead of time to some degree. In these cases, if the controller knows the latency of each robotic striker at each different velocity, compensation can be undertaken. This is carried out by adding a longer delay to quick strikes and less delay to slow strikes in order to make all latencies consistent. Once all strikes by all instruments at all velocities are performing at the same latency, the MIDI sequence can be made to play back that amount earlier and the above-mentioned timing issues are greatly reduced.

In order to acquire the latency information of the various instruments at their range of velocities, the DSP board starts a timer each time it receives notification of a strike beginning from the solenoid control board. When an onset is observed by the onset detector, the timer is stopped and its value is recorded as the latency of the strike. This latency information is then sent to the solenoid control board along with the velocity and channel information. When the solenoid control board receives this information, it amends its latency calibration tables with the newly measured value and carries out linear interpolation between that value and the previously measured value in order to estimate the values in-between. This process is illustrated on the right side of Figure 5.21.

5.8.5 Amplitude Detection

Amplitude detection is utilised to implement continuous calibration of the loudness of the robotic instruments' strikes in order to ensure the consistent and accurate performance of musical sequences performed with them. After a strike notification has arrived from the solenoid control board and an onset has been detected, the amplitude detector begins observing the audio's

envelope, recording the highest value that it reaches. When the envelope of the strike is complete, the highest observed value of the strike is noted as the amplitude of the strike.

Similar to the Closed-Loop Robotic Glockenspiel amplitude detection described in Section 3.4.4, this measurement is only taken on strikes to an instrument that has not been struck over the previous three seconds, to prevent a build-up of sound skewing the amplitude values. Another strike on the same instrument that interrupts the stand-down period or the amplitude measurement period will also void an analysis. If 50 ms elapses after strike notification and no onset was detected, a message to communicate this fact is sent from the DSP board to the solenoid control board, which raises the pulse-width value at the velocity used, interpolating all pulse widths below that value accordingly. If all of the above criteria are met and an uninterrupted amplitude value has been acquired, it is sent to the solenoid control board so that it can amend its velocity calibration tables in order to bring the velocity curve in line with an ideal loudness curve. The dynamics calibration process is illustrated on the left side of Figure 5.21.

5.8.6 Envelope Length Detection

Envelope length detection is a process which determines the length of a musical event from the point of onset to the point that the audio signal is no longer present above the noise floor. For most percussion instruments, the length of a strike's envelope is possible to estimate from the loudness of the transient, as the sound following a strike will decay in a predictable fashion. However, instruments such as the hi-hat of the new robotic drum kit described in Chapter 4 possess the ability to dynamically manipulate the timbral and temporal characteristics of their sound throughout a performance, making such estimates impossible. The hi-hat movement mechanism described in Section 4.4 enables this ability by precisely raising and lowering the top hi-hat cymbal upon receiving serial commands from the solenoid control board. Envelope length detection is the MIR mechanism by which closed-loop control of the hi-hat's height can be realised.

When a hi-hat is set up, there may occur slight variations in the distance between the cymbals. Though the positioning system can place the hi-hat between the initialised extremes, this does not provide musicians controlling the unit with information about what the musical result will be at a given

FIGURE 5.21: Flow diagram of the calibration routines running on the solenoid control board.

height. As the primary musical result of a hi-hat being opened and closed is changes in the length of the sound's envelope, closed-loop envelope length detection can match the solenoid power values output from the solenoid control board with envelope lengths in real time. This allows the user to specify the desired length of the envelope, and be confident that the robotic hi-hat will strike the instrument with the cymbal at a height that produces an envelope near that length. In practice, the solenoid control board is receiving 7 bit MIDI Control Change (CC) messages, mapped linearly from the shortest to longest desired envelope lengths. These values may be sequenced directly, or optionally presented to the user in an interface that displays these envelope length values in seconds explicitly.

To extract the envelope length information from the incoming audio, the algorithm undertakes several processes, illustrated in Figure 5.22 and listed below.

1. Since the audible activity of a hi-hat is predominantly present in the higher end of the audible spectrum, a simple high pass filter with a cut-off of 300 Hz is applied to remove any unrelated low frequency activity that may be present in the audio.
2. An envelope signal is then created by rectifying and averaging the last 1024 absolute values of the audio stream.
3. After a MIDI Note On message which corresponds to the striker of the robotic hi-hat is received, the algorithm expects a strike to follow soon after. When a sudden onset is observed on the averaged envelope signal, the algorithm starts a timer which measures the length of the envelope.
4. When the envelope drops below a certain threshold value, the timer is stopped, and the length of the envelope is recorded.

When an uninterrupted envelope length is observed by the designated hi-hat channel of the DSP board, the 16 bit length in milliseconds, together with the velocity and height used, is sent to the solenoid control board. That board then populates a table of solenoid power values that correspond with the desired envelope lengths (user's sent CC values) and interpolates between the measured values to generate a curve that creates a linear relationship between the two. The signal flow of the calibration algorithm is shown in the centre column of Figure 5.21. The algorithm automatically and continuously updates the calibration table when new, eligible envelope values are

FIGURE 5.22: Oscilloscope output of the envelope length detection function.

received, but because of the two-dimensional nature of the velocity of strikes and the position of the hi-hat, uninterrupted strikes of the same velocity and height may be few and far between. For types of music where this may be an issue, it is helpful to also carry out a short, automated initialisation of the unit on start-up, sending spaced out strikes to the hi-hat at full velocity, at several different heights.

5.8.7 Additional Potential Capabilities

While the previous subsections outlined firmware algorithms that are utilised in the closed-loop robotic percussion array control system, the expandability of the hardware and the extensible nature of software mean that this system also has a range of capabilities beyond the scope of this thesis. This subsection serves to outline a few of these additional potential capabilities in order to provide direction to future practitioners and researchers in this field.

The large number of DAC channels on the small DSP board create possibilities of integrating a range of systems that have heretofore predominantly relied on personal computers and racks of audio interfacing hardware. In particular, it can facilitate the integration of large scale magnetic resonance

instruments such as the Magnetic Resonator Piano [96] that utilises an audio-rate closed loop for synchronising the phases of the resonating bodies and the output waveforms used to excite them. In his *Techniques and Circuits for Electromagnetic Instrument Actuation*, McPherson states that “for poly-phonic instruments such as the piano, it is not feasible to have an audio output channel for every string.” As a result, the polyphony of the Magnetic Resonator Piano is limited and multiplexing techniques are used. By porting the PLL routines to the firmware of the DSP board described in this chapter, all 88 keys of the piano can be controlled simultaneously with just three DSP boards and an associated array of amplifiers, all integrated inside the instrument. This would remove the need for multiplexing, polyphony limitations, and external computing and audio interfacing hardware.

Other potential applications include:

- Actuated instruments with large arrays of loudspeakers and pickups.
- Massively multichannel synthesis, audio effects, and distribution.
- Digital mixing and spatialisation among acousmonia.
- Large arrays of audio-rate industrial style closed-loop systems.
- Other types of closed-loop musical robot creations utilising, for example, FFT-based timbre analysis, for recognising the distortion points of resonating bodies.

5.9 Evaluation of Functionality

In order to establish whether the closed-loop robotic percussion array control unit adequately performs the real-time continuous calibration functions outlined in the previous section, a series of tests are conducted. Similarly to the evaluations carried out in Section 3.5, the primary characteristics that are tested are dynamics and latency, but in this case the linearity of hi-hat expression is also examined. The testing methodologies of each of these three characteristics are outlined below, comparing performance between traditional open-loop and the novel closed-loop configurations.

5.9.1 Methodology

As a representative percussion mechanism to test the closed-loop control system with, the snare drum and associated striker described in Section 4.3.1 are chosen. An Audix i5 dynamic cardioid microphone is utilised as the audio transducer for the closed-loop feedback. It is attached to the rim of the snare drum using a specialist Audix clip that maintains its position pointing at the centre of the drum's head at a 45 degree angle, 50 mm from its surface. This microphone is connected via a balanced XLR cable to channel 1 of the robotic control unit, with the corresponding solenoid control signal connected to the same channel.

To provide an external measurement of the system, a second microphone is utilised, pointing directly down at the centre of the drum head, 100 mm from its surface. A Shure SM57 dynamic cardioid model is used for this purpose. The microphone's output is recorded into a DAW via a MOTU 828 mk3 Hybrid audio interface at a sample rate of 44.1 kHz and 16 bit resolution.

Measurements are first taken of the instrument in its non-calibrated 'open-loop' state, where the solenoid control board activates the striking mechanism, and auto-calibration and the functionality of the DSP board are disabled. Following the open-loop measurements, the closed-loop automatic calibration functionality of the DSP board and the latency compensation functions of the solenoid control board are enabled. A sequence of low, medium, and high velocity MIDI messages are sent to the instrument, simulating a sound check or run through of a musical performance. While each strike is taking place, the analyses described in Section 5.8 are being conducted and the instrument is being automatically calibrated. After this simulated sound check, the calibration is frozen so that an accurate snapshot of it can be acquired. The measurements of the loudness are then repeated with respect to the MIDI velocity values in order to establish the extent to which the system is capable of successfully mapping the velocity values to a linearly rising range of loudness. Measurements of the latency of the strikes are also taken, including the compensation that has been conducted by the DAW.

5.9.2 Dynamics

In order to observe the snare drum striking mechanism's dynamic response, recordings are made of it striking the snare drum, powered by a range of

pulse widths. Two-second renders of each strike are exported from the DAW, and the loudness of each strike is extracted using a script written in the Chuck musical programming language. The script carries out spectral RMS analyses on 1024-bin blocks of the audio and sums them, generating a loudness value (RMS) for each strike. Each of these measurements is undertaken ten times, and the standard deviation of these are illustrated as error bars in the graphs presented below.

Open-loop Response

Figure 5.23 presents a graph which shows the loudness of strikes powered by a range of pulse widths. It shows that pulse widths of $6000\ \mu\text{s}$ and $8000\ \mu\text{s}$ are too short for the striker to make contact with the drum head. The loudness of strikes rises from $10000\ \mu\text{s}$ to $16000\ \mu\text{s}$, and there is little change in loudness from $16000\ \mu\text{s}$ to $24000\ \mu\text{s}$. Pulse widths longer than $24000\ \mu\text{s}$ show much higher standard deviation values, and the relationship between higher pulse widths and loudness is broken. This is because these longer pulses prevent the drum stick from bouncing freely from the drum head, dampening the head's vibration to varying degrees. If working as intended, the automatic calibration mechanism should map MIDI velocity values to the range of approximately $9500\ \mu\text{s}$ to $16000\ \mu\text{s}$ to ensure linearly rising loudness for linearly rising velocity values.

FIGURE 5.23: Graph showing the loudness of strikes of the snare drum at a range of pulse widths of power.

Closed-loop Response

Figure 5.24 presents a graph which shows the loudness of strikes at a range of MIDI velocity values after closed-loop calibration has taken place. The graph shows a relatively linear relationship between increases in MIDI velocity and increases in the loudness of strikes to the snare. This shows that the system correctly mapped pulse widths to the range of MIDI velocity levels.

FIGURE 5.24: Graph showing the automatically calibrated loudness of strikes of the snare mapped to MIDI velocity values.

5.9.3 Latency

Using the same recordings from the dynamic response measurement trials, the latency of each strike is manually measured in a DAW as the time difference between the MIDI note leaving the DAW, and the onset of movement in the recorded audio.

Open-loop Response

Figure 5.25 presents a graph which shows the latency of strikes powered by a range of pulse widths. This latency graph shows higher latency for strikes of lower pulse width, decreasing as the loudness increases. Latency values are not recorded for pulse widths of 6000 μ s and 8000 μ s, as these values did not make contact with the drum. By automatically recording this latency curve,

the algorithm can apply corresponding amounts of extra delay to make the resulting latency consistent between strikes of varying velocity.

FIGURE 5.25: Graph showing the latency of strikes of varying strengths on the snare drum without calibration.

Closed-loop Response

Figure 5.26 shows the measured latency values of drum strikes over the range of MIDI velocity values after latency compensation has been conducted by the control board and the DAW. An ideal result would be a straight line with all values at zero, indicating completely consistent latency and compensated strikes achieving perfect timing. The results show that the latency detection, mapping, and compensation functionality works as intended, with all but two of the values less than 1 ms away from zero, and the remaining two less than 2 ms away. The standard deviation over the entire range is 0.84 ms. If it was possible to set an ‘ideal’ latency compensation value in the open-loop scenario that was at the average level of latency over its range, a standard deviation of 3.16 ms would be observed. This shows that the closed-loop system achieves a significant improvement over the traditional open-loop architecture in this area.

5.9.4 Hi-hat Height

The hi-hat evaluation configuration is similar to the snare drum striker evaluation described in the previous subsection, with the closed-loop transducer

FIGURE 5.26: Graph showing the latency of strikes of different MIDI velocities on the snare drum following automatic calibration and latency compensation.

being an Audix ADX-51 small diaphragm cardioid condensor microphone positioned 100 mm above the top hi-hat, pointed vertically downward at the edge of the cymbal. Solenoid power values are sent as 12-bit numbers from the solenoid control board to the hi-hat movement mechanism via RS-232. The movement mechanism's 12-bit DAC then outputs these values to its built-in amplifiers. The amplifiers convert the values to voltages ranging from 0 to 24 V which are applied to the solenoids.

The hi-hat calibration system measures the length of amplitude envelopes of strikes to the instrument in order to inform the positioning system. To verify its functionality, external measurement of the envelope lengths should be carried out. As manually distinguishing the end of an exponentially decaying envelope to a degree of accuracy in a DAW is problematic, the envelope lengths are verified using the pulse width measuring function of a Rigol MSO1104 mixed-signal oscilloscope. This technique provides much greater reliability of measurement than manual methods, though variation in the measured length tends to be proportional to total length, instead of absolute.

The microphone is connected to channel two of the robotic percussion array control unit, with the hi-hat striker control line connected to the same channel. All strikes of the hi-hat are powered by pulses 11 ms in length, the middle of the dynamic range of the hi-hat striker. The open-loop set of measurements are carried out with all of the automatic calibration functionality disabled. Following the open-loop measurements, the automatic calibration

functionality is enabled and a series of low, medium and high MIDI CC values are sent to the hi-hat, each followed by a strike. This process simulates a sound check or performance rehearsal, and during this time the closed-loop automatic calibration is taking place in the background. The calibration is then frozen in order to capture a snapshot of the system's response. Each measurement is taken ten times in order to provide standard deviation values, as shown in the error bars of the graphs below.

Open-loop Response

Figure 5.27 presents a graph which shows the envelope lengths of strikes to the hi-hat, with varying degrees of power supplied to the soft-shift solenoids in the hi-hat movement system. The lowest power value shows the envelope length of strikes to the top cymbal when it is making no contact with the bottom cymbal. The 400 and 800 values show the result of very slight and occasional contact between the top and bottom cymbals after a strike. In hi-hat performance practice, it is relatively rare to strike the top hat in a position of zero contact with the bottom cymbal, as the particular strength of the instrument is the adjustable contact between the two cymbals.

FIGURE 5.27: Graph showing the envelope lengths of a range of levels of power sent to the hihat movement mechanism without calibration.

After these values that show the hi-hat entirely and almost entirely open, there is a significant decrease in envelope length as the cymbals come into further contact. A higher standard deviation is also observed, pointing to some inconsistencies in the response of the soft-shift solenoids in their very

low power state. By the 1600 mark, the cymbals are making more stable contact, and the solenoids are in a more consistent area of their power curve. This is indicated by the much lower standard deviation values observed.

The 1600 to 3600 area of the graph shows a relatively linear decrease of envelope length to solenoid power, and this is where the majority of the musical expression of the instrument lies. At the 3600 point, the cymbals are tightly closed, and additional power does little to affect the sound of strikes and length of envelopes.

Closed-loop Response

Figure 5.28 presents a graph which shows the resultant envelope lengths of strikes to the hi-hat that have been mapped to MIDI CC values via the closed-loop automatic calibration system. Before calibration, CC values are mapped over the full solenoid power range from 0 to 4095, and the ideal envelope length range is specified as the MIDI CC value multiplied by five. The graph shows that the system was successfully able to scale the solenoid power levels to the MIDI CC values to resemble the specified ideal graph relatively closely.

FIGURE 5.28: Graph showing the envelope lengths of strikes at various MIDI CC values following auto calibration.

The flatness at the bottom of the graph is a result of reaching the same hard limit of tightness described in the open-loop trial, and similar to the open-loop trial, the lower values show very low standard deviations with the higher values exhibiting somewhat less consistency. Given that many widely

used general MIDI (GM) systems utilise just closed, open, and ‘chick’ hi-hat sounds, and the designers of many high-end virtual drumming software systems may sample up to 6 or 7 different hi-hat positions, these results show that this new robotic hi-hat system provides a higher level of expression than such solutions.

5.10 Creative Outputs

In order to make use of the new robotic drum kit in a live concert situation, a performance was given as part of the 2017 TEDxWellington event², held at the St. James Theatre, Wellington, with an audience of 1000. Figure 5.29 presents a photograph of the new robotic drum kit and the robotic percussion array control unit on stage at the event during the performance. The Susan Rhind Award in Composition³ commissioned a new work to be written, resulting in the creation of *The Impossible Drum Solo* for the event.

FIGURE 5.29: The robotic drum kit being controlled by the robotic percussion array control unit, performing *The Impossible Drum Solo* at TEDxWellington.

²<http://www.tedxwellington.com>

³https://www.victoria.ac.nz/study/student-finance/scholarships/browse-prizes/prize-details?result_307328_result_page=134

The piece is split into two parts, each with differing technical and creative focuses. The structure of the performance was such that after a video introduction, part one of *The Impossible Drum Solo* was performed, followed by the composer presenting a talk about some of the concepts and motivations behind the creation of the work. After the talk, part two of the piece was played and the performance concluded. An audiovisual recording of the performance can be found at <https://www.youtube.com/watch?v=hxCMSyx-j0E>. Descriptions of both the technical content and creative considerations of the first and second parts of the piece are provided in the following two subsections.

5.10.1 The Impossible Drum Solo Part One

The purpose of *The Impossible Drum Solo* is to introduce newcomers to the field of musical robotics and expand common conceptions about what it means for music to be ‘robotic’. It demonstrates a combination of expressive musicality and superhuman virtuosity in robotic performance. TEDxWellington is the ideal venue to host such a performance, as the audience is predominantly made up of curious, intellectual, non-musicians.

An audience that is more accustomed to the conventions of contemporary popular music than those of experimental music bring with them certain preconceptions of drum solos and drum kit performance practice. In order to cater to this reality, *The Impossible Drum Solo* utilises a range of idiomatic techniques from popular music styles and extends them into superhuman and experimental territories gradually over the course of the two parts of the piece. The first part is written just for the kit itself, with no sound reinforcement or processing.

Figure 5.30 shows the MIDI sequence of an excerpt of an early section of part one. The middle third of the image shows the MIDI CC sequence that controls the position of the hi-hat, the bottom third shows the velocities of the programmed strikes, and the top third of the figure shows the MIDI note sequence, with each note corresponding with a different actuator of the percussion array. From the bottom to the top, those notes are assigned to Kick, Snare Left, Snare Right, Hi Tom Left, Hi Tom Right, Mid Tom Left, Mid Tom Right, Floor Tom Left, Floor Tom Right, Hi-hat Left, Hi-hat Right, Ride Bow, Ride Bell, and Crash.

FIGURE 5.30: An extract from the piano roll during the first minute of *The Impossible Drum Solo Part One*. The top third shows the drum sequence, middle third shows the hi-hat position, and bottom third shows the velocities of the drum strikes.

At the TEDx performance, the audience expects to see a human presenter on the stage, with a lack of human on stage likely to indicate a waiting period. In order to focus the audience's attention on the robotic drum kit without any human gestures, the piece begins with a snare roll of the type that might introduce a sports team or circus act. The snare roll transforms into a rhythmic snare solo of the type that is common in marching bands, and the three toms join the snare one by one with the same rhythm, but staggering accents. The tail end of that section can be seen on the left of Figure 5.30.

Following the climax of that section, the kit commences a section of idiomatic kit drumming inspired by improvisational live jungle / drum 'n bass practice. This high speed virtuosic section provides a stark contrast in drumming style to the preceding material. This reveals to the audience a particular strength of robotic drum kits: their ability to perform styles of drumming that require the coordination of large numbers of people and are impossible for solo performers, yet also recreate grooves that solo performers excel at, and large groups of performers struggle to perform with a high degree of expression. The piece progresses to a half-time, highly swung, ride cymbal-based style, referential of jazz drumming practice, shown on the right of Figure 5.30. This loosening of quantisation and addition of precisely programmed expressive timing and dynamics challenges the audience's notion of 'robotic' music, a concept which is explored in depth in [66].

FIGURE 5.31: An extract from the piano roll during the second minute of *The Impossible Drum Solo Part One*.

Following the early sections that introduce the robotic instruments to the audience by drawing inspiration from a range of popular music styles, the middle of the piece moves toward more experimental territory as shown in Figure 5.31. This develops the B section of the piece, bringing the audience unfamiliar sonic content designed to expand their perspective of what a drum solo can be. The various instruments carry out rolls that adhere to time-domain logarithmic curves, first individually and then in combinations and in different directions, while at the same time being segmented into double-measure-length passages that allow the audience to remain oriented with regards to the expected pulse.

These types of musical passages reference Nancarrow's work, but apply it to unpitched percussion, creating a sound-world that neither a solo drummer nor a large group of human percussionists would practically be able to realise. The greater temporal accuracy that the closed-loop automatic calibration system implements enables these passages to be performed as they were composed, with smooth dynamic and rhythmic lines. These logarithmic curves are subsequently sporadically combined with idiomatic kick, hi-hat, and snare rhythms before transitioning toward the final section of part one, shown in Figure 5.32. This final section recaps elements of the marching band rhythms from the beginning of the piece, completing the simple ternary structure and bringing the audience back to somewhat quantised and familiar territory. However, elements from each of the preceding sections are also

included, and the tempo is continuously accelerated to a point where the majority of the instruments are playing simultaneously and nearing the upper limits of their speed capabilities. Part one of the piece concludes with superhumanly fast tom fills and unquantised kick and cymbal rolling as a direct reference to the end of rock anthems, providing the audience with a clear indication of conclusion and the expectation of applause in preparation for the talking section of the performance.

FIGURE 5.32: An extract from the piano roll during the third minute of The Impossible Drum Solo Part One.

5.10.2 The Impossible Drum Solo Part Two

With part one of the piece and the composer's talk introducing the audience to the concept of musical robots and demonstrating their capabilities for expressive and superhuman performance, part two extends these new possibilities with live audio processing and more experimental content. The talk emphasises the fact that these musical robots are not created to replace human performers or loudspeakers, rather to enhance the ability of composers to create new, previously impossible musical experiences. The use of live audio processing in part two demonstrates that both musical robots and loudspeakers have their own complimentary strengths and that by leveraging these in composition, the two paradigms can create novel and interesting music. Since the sound from the robotic drum kit is being picked up and delivered to the robotic unpitched percussion array control box via the preamps described in Section 5.4, and to the DSP board described in Section 5.5 for

closed-loop audio-feedback control, it can be passed out of the unit via the ADAT converter described in Section 5.6.

After the audio signals have left the unit, it is possible to route these channels of audio to a large variety of commercial audio equipment, and, in the case of this piece, it was streamed via a commercial audio interface to digital audio workstation software running on a laptop computer. From there, various types of creative processing can be applied to the incoming audio channels, composed in automation lanes. Appendix A shows a screenshot of a digital audio workstation loaded with *The Impossible Drum Solo Part Two*'s sequencing and automation score, followed by explanations of the purpose and functionality of each of the channels.

The Impossible Drum Solo Part Two seeks to further explore many of the concepts of liveness, expressive space, and gestural surrogacy that were examined through the process of composing *Broken Mirror* for the Closed-Loop Robotic Glockenspiel, described in Section 3.7.1. This piece builds upon the robotic instrument and automated live electronics paradigm introduced in *Broken Mirror* with the novel creative capabilities provided by the provision of high quality digital audio streams of all of the individual instruments.

Beyond mere extension, there are also several key differences in the realisation of these two pieces. Unlike the glockenspiel, the drum kit is an instrument that is distributed throughout space, and the high sound pressure level of the instruments dictates the types of space that may be used with it. Intimate space is unusable without the provision of hearing protection for the listeners, with the instrument more suitable for ensemble and theatre sized spaces. The St. James provides a partially reverberant indoor amphitheatre where the acoustic sound of the drum kit can propagate, and stand on equal footing with the electroacoustic processed parts. Through the use of dynamic reverberation processing in tracks eight, 14, and 28, the spacial characteristics of the theatre can be extended with creative use of transcendant space.

Though both *Broken Mirror* and *The Impossible Drum Solo* share similar paradigms, the perceptions of gestural surrogacy between the two pieces differ somewhat. The actuation of the glockenspiel is undertaken from beneath its keys, a method which does not echo the traditional ways that humans play the instrument. While there are drumming robots such as <Snr> that apply similar actuation approaches [131], a conscious decision was made with the design of the new robotic drum kit to adhere to the familiar drum kit performance paradigm of striking the instruments from above with drum sticks

and beaters. This has the effect of providing memetic gestures that are imbued with an increased level of relatability and sense of 'effort', that evoke a proprioceptive response in the audience.

A limitation of utilising the drum kit in a solo work is that it is generally incapable of realising melodic and harmonic musical content. The composition of *The Impossible Drum Solo Part Two* circumvents this firstly by tuning the tom drums to a precise harmonic relationship (perfect 4ths between the drums), and secondly by utilising a sampler in track 15 to record certain strikes of the mid tom, and play it back at different pitches to create harmonic and melodic themes in the piece. The preconception of static pitch and envelope content in drum kit performances is also overcome by using a spectral freeze process in tracks 14 and 17 that indefinitely sustains a channel of sound, at which point the pitch of the sound can be bent, and the envelope transformed.

The process of creating *The Impossible Drum Solo* for the new robotic drum kit and its associated control system demonstrates that the system enables many new musical possibilities in several different areas, and it is clear that only a small percentage could be explored in a single work. The piece acts as a conduit through which the accurate implementation of compositions of super-human speed, complexity, and synchronisation can be trialled with physical acoustic musical instruments. It also expands our concepts of what a drum solo can be and acts to demonstrate the efficacy of the technology developed in this thesis.

5.11 Conclusions

This chapter has presented a robotic unpitched percussion array control unit that includes a new solenoid control board that provides nanosecond accurate pulses for actuators, a novel embedded DSP board with 32 high quality audio inputs and outputs and MIR capabilities, new flexible preamp modules, ADAT converters, and supporting hardware. This new closed-loop audio feedback system makes robotic percussion arrays more reliable, accurate, robust, and integrated, fulfilling several stated goals of this thesis. The embedded MIR system implements continuous dynamics, latency, and hi-hat position calibration that can be invisible to the end user, and the integration is such that deploying the robotic drum kit is not much more complicated

than traditional setups. By adhering to audio connectivity standards, the system can be used in conjunction with a range of commercial hardware, and it features a level of expandability that opens up the system's components to a wide range of uses outside those examined in this thesis.

Evaluations were carried out to measure the system's performance in automatically measuring the dynamics, latency, and envelope lengths of percussion robots, and utilising that information to provide continuous calibration of the instruments' velocity curves, latency levels, and hi-hat movement. The results of the evaluations show that the system is capable of dynamically mapping RMS levels to MIDI velocities in order to provide linearly increasing loudness for linearly increasing velocity commands. They also demonstrate the system's ability to measure the latency curve of an instrument and use that information to carry out latency compensation, significantly reducing potential timing irregularities in performances. Evaluations of the audio-based closed-loop calibration of the hi-hat movement system showed that the unit can realise a relatively linear relationship between the MIDI CC values used to control the system and hi-hat envelope lengths.

The composition and performance of a new two-part work for the system has demonstrated its enhanced capabilities and acted as a vehicle through which to explore new creative processes and workflows. As the second generation closed-loop robotic percussion system after the Closed-Loop Robotic Glockenspiel, it makes significant improvements in the deficient areas of that first system, and it provides a platform for the creation of a large range of new types of works for percussion arrays. It is hoped that the creation of this new hardware and software will push the state of the art of robotic percussion forward, so that these types of instruments can come closer to matching the capabilities and prevalence of non-robotic electronic instruments.

The research presented in this chapter is featured in the following publication:

J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "Improving the Accuracy of Musical Robots with Programmable Logic Circuits". In: *Proceedings of the Electronics New Zealand Conference, ENZCon*. Wellington, New Zealand: Victoria University of Wellington, 2016

Chapter 6

Closed-Loop Chordophones

6.1 Overview

While the categories of pitched and unpitched percussion robots represent the majority of instruments in the musical robotics field, the next most significant group of instruments are robotic chordophones. The use of chordophones offers composers a wide range of timbres and modes of expression that cannot be achieved with percussion units, but their implementation is generally more complex, presenting unique challenges to practitioners. Some of these challenges relate to issues with accuracy, reliability, and robustness, some of the central themes of this thesis as outlined in Chapter 1. Embedded closed-loop audio feedback systems have the potential to make significant improvements in these areas.

Robotic chordophones come in a large variety of forms ranging from bowed instruments such as violin playing robots [52] and modifications of traditional stringed instruments [75], to augmented electric and bass guitars [31], new forms of string plucking instruments [143], and instruments that utilise magnetic resonance techniques [138]. As their development has progressed, these mechatronic instruments have been able to surpass human capability in parameters like picking speed [164], and have steadily been adding new expressive capabilities to their repertoire [108].

However, as outlined in Section 2.6, there is a key component of robotic chordophones that is yet to be fully realised. Labelled ‘kansei’, or sensibility, this is the capability of a robot to ‘hear’ the sound that it is making during a performance, and adjust its technique accordingly, in a manner akin to a human musician. This is important for ensuring that when a robotic instrument

plays out of tune, or with an incorrect timing, timbre, or loudness, it is capable of fixing these issues automatically and in real time. While these parameters are all applicable to chordophones, the most significant issue for these types of instruments that was not addressed in previous chapters is that of tuning. The nature of chordophones is that their strings tend to go out of tune due to environmental and performance conditions, making pre-performance and mid-performance tuning an unavoidable necessity in concert situations. While it is an inconvenience for players of these instruments, the problem is magnified in the case of robotic chordophones because they perform autonomously, often without direct supervision by a human.

There are a number of different methods musical roboticists have devised to deal with the issue of robotic chordophones losing their tuning. Trimpin's *If VI Were IX* installation, which consists of a large structure of roboticised chordophones, utilises an array of commercially available guitar tuners to aid in the automatic tuning of the exhibition [31]. Another example is the Gibson Robot Guitar. While not an automatically playing instrument, it embeds tuning functionality inside the instrument and utilises a set of motors to turn the guitar's tuning pegs on command until the strings are in tune.

While the above instruments are fretted devices, fretless robotic chordophones have additional techniques at their disposal. Murphy's Swivel robotic slide guitar can be tuned by connecting the audio output to custom tuning software running on a laptop. This software then provides the instrument with an array of 'virtual frets', positions for the fretting mechanism to travel to in order to achieve the desired notes without requiring tuning pegs to be adjusted [107]. Shawn Trail's STARI robot approaches this problem from a different direction by implementing a pre-performance tuning algorithm in PureData, and embedding the software inside the robot on a Beaglebone Black single-board computer [159]. The implementation that comes closest to the goal of embedded pitch extraction in robotic chordophones is Barton et al.'s Cyther V2, which includes audio feedback in its control system [6]. Cyther utilises a tension sensing mechanism to estimate the current pitch of the respective string, with a frequency analysis used to confirm this estimation. However, the observed accuracy and latency of its pitch extraction leaves significant room for improvement.

All of the above mentioned systems focus on *pre-performance* tuning of robotic chordophones. This chapter introduces a new system, based on the work carried out in the previous chapters, that seeks to fill this gap in the research by

adding continuous, real-time pitch calibration that prevents a robotic chordophone from drifting out of tune during a performance. This is especially useful for performances that are particularly long or vigorous and for gallery and installation works that may need to run for days, weeks, or months at a time, as these works are most likely to be affected by tuning issues. In implementing these systems, a compromise must be made between the speed and accuracy of correction, as larger averaging buffers for pitch detection provide greater accuracy, but slower response. Also, the larger the error in the string's tuning, the longer it may take to correct. Through the course of working with these units, a reasonable goal for the system is to ensure that an out-of-tune note played by the instrument can be automatically adjusted to be in tune to within 10 cents of the target frequency in less than a second.

A smaller-scale embedded DSP board based on the work described in Chapter 5 is presented that continuously monitors the audible output of a robotic chordophone during its performance, analyses the information in real time, and utilises the resulting musical data to automatically maintain the instrument's tuning. In order to provide an example use case for the DSP system, an implementation with *ServoSlide*, a robotic chordophone recently built by Neill Skelly [145], is presented. This chapter begins by discussing the design of that instrument. A description of the DSP board's hardware is subsequently provided before outlining how it interfaces with the mechatronic chordophone. The firmware of the DSP board is then described, including its real-time pitch extraction algorithm. Following that, evaluations are undertaken that quantify how effective the system is at keeping itself in tune, and the chapter concludes by discussing some of the wider ramifications of this work.

6.2 *ServoSlide*

As mentioned above, in order to provide an example use case for the closed-loop audio feedback system, an implementation was created with Neill Skelly's new *ServoSlide* robotic chordophone. Shown in Figure 6.1, the instrument is one of the most recent of several robotic chordophones produced by graduate students of Victoria University of Wellington, which include *MechBass* [97], *Swivel* [104], and *StrumBot* [165].

FIGURE 6.1: ServoSlide, entire instrument (top), fret positioning mechanism (bottom left), string clumper (bottom middle), and picking mechanism (bottom right).

ServoSlide provides a number of capabilities that make it a suitable candidate for implementing closed-loop audio feedback, and the mechanisms that make up the instrument are described in turn below.

6.2.1 Picking Mechanism

The picking mechanism as shown on the bottom right of Figure 6.1 consists of an MKS95i servo with a plectrum attached to its output spline via a small 3D printed bracket. The servo is mounted to a 3D printed ABS housing which also contains the instrument's tuning peg and bridge. Rotating the servo in either direction forces the pick through the string, exciting it. Also mounted in close proximity to the picking mechanism is a commercially produced Fishman passive electromagnetic pickup, which enables the robot's string vibration to be amplified, recorded, and processed for performance situations. The choice of a small servo for picking the string, as opposed to a solenoid or stepper motor, results in minimal interference observed in the pickup coil. In addition, this pickup is also the source of the audio information received by the DSP board described in Section 6.3.

6.2.2 Fretting Mechanism

Fretting the instrument is carried out by the combined efforts of both the positioning mechanism shown in the bottom left of Figure 6.1 and the clamping mechanism shown on the bottom middle of the same figure. The positioning section utilises a Dynamixel MX-64 digital servo motor mechanically coupled to an overhead arm in a manner similar to StrumBot [165]. The lower and upper halves of the arm are connected via bearing assemblies to allow free rotation, and the opposite end of the arm is connected to a sliding carriage. This translates the servo's rotational motion to a linear travel along the string for the clamping device.

The clamping carriage is transported along a length of Servo City X-Rail aluminium extrusion, making use of four V-Wheels that are designed specifically for the X-Rail. The wheels are manufactured in Delrin, a plastic characterised by its high rigidity, toughness, and low friction coefficient, resulting in a much lower level of extraneous acoustic noise generated during locomotion than previous units [145].

The clamper itself is actuated by an MKS95i servo powered at 6 V. Attached to the servo's horn are dual clamping and damping fixtures, so that when the servo is rotated anti-clockwise, it is capable of damping the vibration of the string to varying degrees, and when rotated clockwise, it clamps the string with an adjustable force. The clamper utilises a pair of cylindrical brass lugs mounted in parallel. These lugs are mounted on bearing assemblies that allow them to rotate freely, a development which also acts to decrease extraneous noise and allow for the clamper to roll along the string while clamping. The high resolution positioning servo coupled with the rolling clamp mechanism makes this instrument highly suitable for an embedded audio feedback system that keeps it in tune by sliding to the intended frequency while a note is played. This ability allows the instrument to correct the pitch of a note as soon as an error is observed rather than needing to wait for subsequent notes.

6.2.3 Existing Closed Loops

There are several non-audio closed-loop control systems at work in this machine. Firstly, each of the MKS DS95i servos contains their own closed-loop control system, activated by pulses from 380 to 1100 μs at a rate of 560 Hz.

The parameters of the tuning of these closed systems are not available to be adjusted by the user. However, the Dynamixel MX-64 utilises a half-duplex serial communication line to allow users to adjust the internal PID parameters of its own closed-loop mechanisms, and feed information back to the user to enable closed-loop control within the user-defined control system. These internal parameters are left in their default settings for this project.

In addition to the closed-loop possibilities provided by the servos, the ServoSlide's control board implements an additional feedback parameter. By connecting the clamping servo's return path to ground through a $1\ \Omega$ resistor and feeding that signal to an analogue to digital converter, a reading of the current consumption of the servo can be conducted. This reading is then utilised during the initialisation of the robot to populate clamping servo positions that correspond to adequate pressure being exerted on the string to achieve a successful clamp. This servo pressure calibration is a development by Neill Skelly outlined in [145], and is not the focus of this chapter.

While these various closed-loop abilities are helpful in ensuring the clamping performance and positioning accuracy of the robot, the fact that none of these systems observes the musical output of the instrument means that in order for it to keep itself in tune, the extra hardware described in the following section is also necessary.

6.3 The DSP Board

FIGURE 6.2: Rendered images of the top side (left) and bottom side (right) of the fully populated embedded DSP board.

As robotic chordophones often utilise several different actuators and require more simultaneous types of MIR processing per audio channel than robotic percussion units, a new DSP board customised for this task is created. Rather than the star configuration utilised by the DSP and actuator control boards described in Chapter 5, this new board, shown in Figure 6.2, is designed to be integrated into a single robotic chordophone unit, and carry out all of the control, closed-loop audio analysis, audio conditioning, and communication with the outside world. In addition to carrying out all of these functions, the board must also be small in size so that it may be permanently attached to a robotic chordophone unobtrusively. This section describes the design and functionality of the board, and how it is integrated with ServoSlide.

6.3.1 Board Design

This small DSP board utilises many of the elements of the DSP board described in Section 5.5, using the same processor, flash storage, power conditioning, and programming and booting methods. However, it condenses the design down to a four layer, 5 cm×5 cm circuit board that has stereo audio input and output, and adds a variety of interfacing circuitry. Sections of that interfacing circuitry that are not immediately required do not need to be populated on the board, and Figure 6.3 shows a photograph of a mostly populated board, with annotations that label the various sections.

The core of the board is the ADSP-BF592 DSP and the stereo audio codec, which are both clocked by the crystal oscillator positioned between them. A Wolfson WM8731 was chosen for the codec, as it provides up to 24 bit resolution at up to 96 kHz sample rates, is controllable by I²C, and can communicate directly with the Blackfin via I²S digital audio streams. There is also analogue audio conditioning circuitry on board that utilises TL074 and TL072 quad and dual low-noise JFET-input opamps to scale and bias the incoming audio in similar ways to the conditioning circuitry described in Section 3.4.3.

The audio signals enter the board on the right-hand side of Figure 6.3 and pass through the input level adjustment surface mount trimmer potentiometers before proceeding to the audio conditioning circuitry and subsequently to the WM8731. After the audio is digitised, the codec passes the digital signal to the DSP via an I²S stream, and the DSP drives the signal indicator LEDs next to the trimmers. Each channel has one green LED to indicate the presence of signal on that channel and a red LED to indicate clipping. The

FIGURE 6.3: Annotated photograph of the DSP board with all headers, SPI ADC, and UART selection switch unpopulated.

thresholds for these are set in firmware. This configuration allows the user to quickly and easily set appropriate levels, ensuring that the closed-loop audio feedback functions operate effectively.

Flexible communication capabilities are facilitated by including MIDI In, Out, and Thru circuitry, along with MOSFET-based level shifters on the UART, as shown on the top left of Figure 6.3. This allows the board to handle all musical communications and also carry out serial communication at 1.8 V, 3.3 V or 5 V logic levels. A DPDT toggle switch is included which allows the user to select MIDI or UART interfacing. A boot jumper near the centre of the board also enables booting the processor via the UART port, allowing programming of the board using an open source compiler, without requiring the use of expensive proprietary programming hardware. The programming and booting process is described in Section 5.8.1.

The second SPORT, I²C, multiple GPIOs, and the output pins of two dedicated high speed 32-bit PWM timers are made available to the user, as shown on the left of Figure 6.3. These can be used to pass digital audio in and out of the board, control various types of actuators such as solenoids, stepper motors, and servos, receive digital information from buttons or sensors, or other such functionality that may be required when integrating the system with a robotic chordophone. An array of analogue voltage inputs are also provided, serviced by a MAX11613 quad I²C SAR ADC and an ADC124S101 quad SPI SAR ADC. These cater to potential non-audio analogue signals that may need to be acquired by the system, such as potentiometers, control voltage inputs, analogue sensors, and other signals.

6.3.2 Integration with the Robot

Some use cases with more complex actuator control requirements would see this two-channel DSP board conducting the MIR analyses and sending the resultant information to the existing controller hardware of a robot, similar to the architecture of the system described in Chapter 5. However, in the case of integrating it with the ServoSlide robotic chordophone, the DSP board is capable of carrying out both control and analysis functionality.

Control of the MX-64 positioning servo is performed by the DSP's UART at a baud rate of 1 MHz, with the half-duplex bidirectional communication

switching conducted by utilising a GPIO pin to switch a CD74HCT241E octal buffer chip. Both of the clamping and picking servos are driven by PWM signals directly from the DSP's timers, and the current sensing inputs are fed into the on-board I²C ADC. The audio output from ServoSlide's pickup is fed into one of the codec inputs, leaving the audio outputs available for producing any applied audio processing and effects, or for debugging purposes.

6.4 Firmware

In order to implement both the actuator control and closed-loop audio feedback functionality, firmware is written for the BF592 that boots the chip, configures its clocks, and initialises its internal and external peripherals. The firmware also carries out all of the real-time audio analysis for continuous calibration of the instrument's pitch. A description of how the firmware is implemented which addresses these various functions is provided henceforth.

6.4.1 Booting and Configuration

Both temporary and persistent program compilation, storage, and loading is carried out identically to the multi-channel system described in Section 5.8.1. However, some of the peripherals used differ. The clock is the first subsystem configured. The ADSP-BF592 has a maximum core clock speed of 400 MHz and a maximum system clock speed of 100 MHz, so the signal from the crystal oscillator must be multiplied by the PLL to the highest value below 400 MHz for the core clock, and divided to the highest value below 100 MHz for the system clock. An 11.2896 MHz crystal oscillator is used for the clock input, and the target audio sample rate is 44.1 kHz (f_s). The I²S protocol requires a bit clock of $64 f_s$, equating to 2.8224 MHz. Multiplying the input clock by 35 provides a core frequency of 395.136 MHz, and this frequency is divided by five to achieve a system clock of 79.0272 MHz. A divisor of five instead of four is chosen because the SPORT peripheral's bit clock (2.8224 MHz) must be an even factor of the system clock, and a divisor of four would result in a system clock of 98.784 MHz, an odd multiple of 2.8224 MHz.

The GPIO lines are then configured in a way similar to that outlined in Section 5.8.2, though four LEDs are used instead of one. Firmware code is also reused to initialise the I²C and UART peripherals, though a baud rate of 1 MHz is used for communication with the MX-64 servo. The WM8731 audio codec is then initialised via I²C, activating the line inputs and disabling the mic input, setting the volumes of all channels to 0 dB unity gain, disabling all mutes, enabling AC coupling on the audio inputs, utilising de-emphasis at 44.1 kHz, and powering down unused modules in the unit. The digital audio interface is set to 16 bit I²S format with a master clock at $256 f_s$ (11.2896 MHz, directly from the same crystal oscillator that clocks the DSP).

The MAX11613 12 bit ADC is then also initialised via I²C, set to unipolar mode, and configured as a single channel in single-ended mode. This ADC takes the current sensing input from ServoSlide. Following that, the SPORT is configured to match the settings described for the codec above, and digital audio begins streaming between the codec and the DSP. A high priority interrupt is triggered at each word clock, and the audio analysis described in Section 6.4.2 is conducted.

In order to control ServoSlide's clamping and picking analogue servos, two of the BF592's timers are configured as PWM outputs with a frequency of 560 Hz, and pulse widths that range from 380 to 1100 μs , with the width dictating the position of the corresponding servo. A medium priority level interrupt routine is then attached to a third timer. Set at 1 kHz, this function acts as a 'control-rate' routine where the logic that processes non-audio inputs and controls the actuators is located. The lowest priority processing such as logic that drives the onboard LEDs is then placed in the main loop.

6.4.2 Audio Analysis

In addition to the tasks of controlling the robot's actuators, the firmware loaded onto the DSP undertakes regular pitch analyses of the instrument's sound to check that the desired pitch is indeed being played, and if not, to automatically adjust the instrument's fretting position and calibration tables to the correct values. In order to carry out these functions quickly, reliably, and with adequate resolution, a three step method is conducted that consists of envelope detection, signal conditioning, and zero-crossing-based pitch extraction. These steps are outlined in the following sections, with reference to Figure 6.4.

Closed-Loop Pitch Extraction Continuous Calibration

FIGURE 6.4: Diagram illustrating the pitch extraction and continuous calibration routines of the Closed-Loop Chordophone system.

6.4.3 Envelope Detection

Section 5.8.6 outlines a method of detecting the length of a sound's envelope for the purposes of associating a robotic hi-hat's sound with its cymbal height. The same process described in that section is conducted with the robotic chordophone, but for a different purpose. Instead, the process is carried out in order to define the period during which pitch detection should take place. As carrying out pitch detection on a very weak signal that is near the noise floor of the system is likely to generate erroneous results, the envelope detector works to only extract pitch information from signal that is confirmed to be a part of a plucked note, and has not decayed to a level near the noise floor.

6.4.4 Conditioning the Signal

Since the rich harmonic content of a plucked string can generate a high number of unwanted zero-crossings, a separate FIR filter in a low-pass configuration is applied to the signal, with a cut-off frequency just beyond the maximum playable fundamental frequency of the robot. The result of this processing is shown in the transformation of the audio input to the filtered signal in Figure 6.5.

FIGURE 6.5: Oscilloscope output illustrating the filtering and zero-crossing extraction.

6.4.5 Pitch Detection

There are several methods of extracting fundamental frequency information from an audio signal, each with benefits and drawbacks. The performance requirements of this system dictate that it should be low latency and have high frequency resolution. Though FFT-based methods are effective at differentiating between the fundamental frequency and other harmonics of a signal, they also introduce latency to the process, and FFTs with high resolution require very large block sizes. It is for these reasons that a zero-crossing based method similar to the one outlined in [119] would be more suitable.

In order to improve on the reliability of the pitch extraction algorithm, adaptive frequency selection is applied. Since octave-higher errors are common in zero-crossing algorithms, a frequency eligibility band slightly below an octave in width is defined, with the currently calibrated pitch of the string at that position situated in the middle. This band is then defined around the current position of the fretting mechanism, moving with it, so as to accommodate a realistic level of string detuning while eliminating octave displacement errors.

In order to improve on the frequency resolution and reduce the effect of outliers, the algorithm averages the frequency measurements over a user-definable number of adjacent wavelengths. Averaging higher numbers of wavelengths increases the reliability of the measurement, but also increases latency, while lower numbers are less stable but more frequent. Through experimentation, a value of 16 wavelengths was arrived upon as a reasonable compromise. To reduce the susceptibility of this process to erroneous measurements skewing the average, each individual wavelength must be within the frequency eligibility band mentioned above. If a zero-crossing is observed earlier than expected, it is discarded, and if no crossing is observed past the lower extreme of the band, a timeout occurs, restarting the measurement process.

The frequency information extracted via the described algorithm is then utilised in dynamically redefining the calibration table of the instrument. If the desired pitch is elsewhere on the string, the positioning mechanism can shift the clamper in the direction of the correct pitch, stopping when it reaches it. In most cases this movement is so subtle as to not be noticed by an audience, but produces the benefit of the instrument staying in tune during a performance or installation presentation.

6.4.6 Detecting Other Characteristics

With the ServoSlide, the primary parameter for audio-based feedback control is the instrument's pitch, but the hardware presented is equally capable of implementing a wide variety of closed-loop MIR algorithms for the improvement of robotic chordophones. It can apply the algorithms described in Chapters 3 and 5 to allow instruments like MechBass to dynamically calibrate the loudness of their picking, and to measure and compensate for the latency of the instrument's actuations. Other instruments like robotic violins, which have parameters such as bowing speed and angle that manipulate their own timbre, could also benefit from real-time FFT-based spectral analysis to continuously calibrate their actions.

There are also a number of creative applications for continuous pitch-extraction of external sound sources during a performance. An example could be having the chordophone dynamically stay in tune with a human performer, electronic sound source, resonating acoustic environment, or other robotic musical instrument. Another use case could see the instrument automatically identifying non-standard scales and tuning systems and dynamically adjusting its performance to match them. By utilising this system to detect a wide range of musical characteristics, a large number of extended musical techniques can be achieved and novel musical performance paradigms are made possible.

6.5 Evaluations

In order to establish whether the closed-loop chordophone system adequately performs the real-time continuous calibration functions outlined in the previous section, evaluations are carried out. Since ServoSlide is incapable of altering its dynamics, and latency compensation was covered in Chapters 3 and 5, this evaluation focuses specifically on pitch recognition and correction. The testing methodology is outlined below, comparing the performance of the traditional open-loop configuration with the novel closed-loop one.

Before testing of the instrument begins, the string is first characterised. This is carried out once by first tuning the string to middle C (261.6 Hz) and making

a recording of the string being fretted and plucked at each of the approximately 1000 fretter servo positions. This recording is scanned by a custom ChuckK script that applies a 32768-bin FFT to each pluck and outputs the value of the frequency with the highest amplitude. The resulting table of servo positions to resulting frequencies acts as a curve from which the desired musical scale can be accurately specified.

For these evaluations, the ideal frequency of the open string is set as 261.6 Hz, middle C, and since the fret positioning mechanism of ServoSlide is incapable of reaching the nut, an ascending, equally tempered D major scale with a pitch reference of 440 Hz is used. Four tests are carried out, each of which consists of the DSP board being commanded via MIDI to perform an ascending D major scale, with notes being plucked at two-second intervals. The notes are recorded into a DAW via a Focusrite Scarlett 18i20 first generation audio interface at a sample rate of 44.1 kHz and a resolution of 24 bits. The fundamental frequency of these notes is then established by running the recordings through the ChuckK script outlined above.

Open-loop Response

The first test establishes a baseline of tuning by demonstrating the performance of the instrument when configured in the traditional open-loop configuration. The string is manually tuned to middle C and the array of virtual frets are populated with values that correspond with the desired D major scale. The blue line of Figure 6.6 shows the measured frequencies when performing this scale, and these correspond very closely with their ideal values.

A second test is then undertaken that is identical to the first, except that before it is carried out, the string is manually tuned down one semitone to the B below middle C. This simulates the slight detuning of strings that commonly occurs during performances with chordophone instruments. The orange line in Figure 6.6 shows the pitches of the resulting notes output by the instrument. All of the notes are significantly flat. This shows that when strings lose their tuning during a chordophone performance, all of the subsequently performed notes will be out of tune, and the situation will not be remedied until either a manual or automated offline tuning takes place.

FIGURE 6.6: Graph showing the frequency of each note played without calibration, with the instrument manually tuned correctly (blue) and detuned by a semitone (orange).

Closed-loop Response

Following the open-loop tests, the closed-loop continuous calibration functionality of the instrument is activated. Replicating the first test, the open string is tuned to middle C and the trial is run. Figure 6.7 shows the outcome of frequency analysis of the performed notes. The values are almost identical to the blue line of Figure 6.6, showing that the D major scale is performed correctly and remains in tune with continuous calibration activated.

FIGURE 6.7: Graph showing the frequency of each note played with calibration, with the instrument manually tuned correctly.

For the fourth test, the instrument is detuned a semitone to the B below middle C, similar to the second test, and the continuous calibration functionality remains activated. Figure 6.8 shows that despite the manual detuning of the string, the frequency curve of the instrument remains close to the in-tune curve of the blue line of Figure 6.6. Figure 6.9 shows the difference between these two graphs in cents. All of the middle values are in tune to within 10 cents, with the bottom D recorded as 14 cents out, as that was the first note in the sequence, where the majority of the automatic adjustment took place. The top C $\sharp$ note is also flatter than its ideal tuning. This is because the semitone rise in the servo positions required to accommodate the detuning has pushed the fret positioning servo to the edge of its travel and it is unable to reach higher notes.

FIGURE 6.8: Graph showing the frequency of each note played with automatic calibration, with the instrument detuned by a semitone.

FIGURE 6.9: Graph showing the difference in cents between test one and test four.

While the low D of Figure 6.8 is recorded by the FFT-based pitch-extraction

Chuck script as being almost in tune, perceptually, a short pitch bend is observed by the listener. Figure 6.10 shows a melodic-range spectrogram image of the recording of the first note of the fourth test. The image is zoomed in to the second harmonic as it most clearly demonstrates the effect. The zero seconds mark is when the MIDI Note-On D command is sent to ServoSlide and the string is plucked soon after. The initial frequency of the string is close to C $^\sharp$ rather than D because of the detuning that has taken place.

Following a short static period while frequency measurements are being automatically carried out by the DSP board, the pitch begins to rise. It continues upwards until it reaches the correct frequency, and after reaching it, the pitch remains in place. Figure 6.10 shows that the goal of reaching the correct pitch to within a tenth of a semitone in less than a second is achieved in this case. Since the automatic calibration routine shifts the positions of all of the virtual frets simultaneously, after this short pitch bend, subsequent notes are performed in tune. In real-world performance situations, detuning of the string is likely to occur more gradually than the entire semitone jump performed in these evaluations, so the magnitude of such pitch bends are less audible to audience members and the speed of reaching the correct frequency is increased.

FIGURE 6.10: Melodic Range Spectrogram view of the first note of the closed-loop evaluation.

6.6 Conclusions

This chapter has presented a small, integrated embedded DSP system for carrying out continuous real-time MIR tasks for improving the reliability, usability, and utility of robotic chordophones. By combining these systems with existing robotic instruments, they require less calibration and maintenance, allowing composers and performers to concentrate on creating compelling music with these instruments instead of constantly supervising them.

Following an outline of the capabilities of the ServoSlide robotic chordophone, the new DSP board's design, layout, and functionality were detailed. The process of integrating the board with the existing robotic instrument was described and the configuration process of the DSP board outlined. The firmware written to carry out continuous, real-time audio filtering, envelope detection, and pitch extraction was also detailed.

The evaluations carried out demonstrate that this novel system is capable of ensuring that the notes played by a robotic chordophone during the course of a performance or installation stay in tune, even as the strings of the instruments lose their tuning. While this functionality is useful in short musical robotics performances, it is especially beneficial for installation works or other long-form musical practices which more often suffer from tuning issues. By demonstrating that embedded DSP-based closed-loop systems can be successfully applied to robotic musical instruments with continuous pitch and envelope characteristics, this research opens the door to these systems being implemented in a majority of the musical robots in existence, potentially providing significant improvements in the performance practice of these instruments.

The research presented in this chapter is featured in the following publication:

J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "Improving Robotic Chordophones with Closed-Loop Embedded Audio Feedback Systems". In: *Proceedings of the International Computer Music Conference, ICMC*. Shanghai, China: Shanghai Conservatory of Music, 2017

Chapter 7

Conclusions

In the first chapter of this document it was noted that while the field of musical robotics is experiencing a rise in popularity and accessibility, there remain many barriers that prevent musicians, composers, and other makers and artists from utilising them in their work. It was observed that many musical robots suffer from reliability and accuracy issues, require very specialised technical knowledge to build and play, and may lack desired expressive capabilities. It was hypothesised that by implementing audio-based closed-loop feedback control systems with musical robots that simulate the way a human performer listens to their instrument while they play, it would be possible to increase their robustness, minimise setup and maintenance requirements, and add new modes of musical expression to them.

In order to explore the configuration of these systems with a variety of different types of musical robot, implementations were created for three of the most prevalent categories of robotic musical instruments: pitched percussion, unpitched percussion, and chordophones. First, an existing robotic glockenspiel was upgraded with an integrated DSP system that processed individual channels of audio from each of its 21 keys in real time, to implement continuous closed-loop calibration of its latency and amplitude characteristics. This functionality significantly reduced the length and complexity of the instrument's setup and increased its reliability. The extra capabilities of the system also enabled the creation of an interactive installation and a new musical work, both of which introduce novel paradigms and experiences to electroacoustic music performance.

Following the evaluation of the Closed-Loop Robotic Glockenspiel in testing and real-world environments, it was observed that custom circuit board designs with higher specifications than existing systems could increase the accuracy and consistency of the closed-loop system, reduce noise in the audio

path, and provide increased functionality. As a result, for the Closed-Loop Unpitched Percussion Array Controller, a fully customised system was created consisting of power supplies, an actuator controller board, a DSP board, preamp modules, and digital audio conversion boards that were all designed specifically for this purpose. A study of existing robotic percussion mechanisms was undertaken, and the most appropriate design was iterated upon to create a new robotic drum kit for use with the closed-loop control system. These new developments created a set of instruments with much higher consistency, that were able to automatically calibrate their latency and amplitude characteristics with less prior knowledge about the incoming signals. It was also capable of closed-loop automatic calibration of hi-hat height by extracting the envelope length of strikes to the instrument, and feeding back that information to the control system. The system was evaluated in laboratory and live performance environments, and these new developments were observed to significantly simplify setup and calibration, increase accuracy and robustness, and provide additional parameters for expression via multi-channel audio output. These novel capabilities were put to use in a new two-part composition for robotic drum kit and loudspeakers.

Since the needs of robotic chordophones differ somewhat to those of robotic percussion arrays, a new, smaller form factor, dedicated DSP controller board was designed and integrated with an existing robotic chordophone instrument. By implementing fast and accurate pitch extraction with the embedded DSP system, it allowed the chordophone to continuously listen to its own pitch and adjust its fretting position and calibration dynamically. Laboratory evaluations showed that the system was capable of automatically fixing discrepancies in the tuning of the instrument quickly and accurately, even with a dramatic and sudden detuning of the string. The result is an instrument that allows composers to focus on writing the musical content that they desire, without worrying about whether the robot will be able to perform it accurately or not.

As a result of exploring and evaluating the outcomes of the closed-loop implementation with three different categories of robotic musical instruments, it is possible to answer all three of the research questions presented in Section 2.8.

1. Yes, the implementation of real-time closed-loop feedback based calibration systems for musical robots can result in higher accuracy in performance than manually calibrated systems.

2. Yes, closed-loop musical robotic systems can reduce the frequency of calibration and tuning required of open-loop systems.
3. Yes, real-time closed-loop feedback systems can provide robots with added musical capabilities that would otherwise be impossible.

7.1 Key Contributions

This thesis makes a number of both technical and musical novel contributions to the field of musical robotics. In order to make these clear and present them in an organised manner, each of the primary contributions is headlined and outlined individually below.

A Survey of Closed-loop Control in Musical Robotics

In Chapter 2, the first comprehensive survey of closed-loop control with regards to the field of musical robotics is presented. The survey examines and clarifies the often ambiguous terminology used in the area and defines audio-based closed-loop control and where the area sits between the fields of general robotic control systems, hyper-instruments, musical information retrieval, and pre-performance calibration of musical robots.

The Closed-Loop Robotic Glockenspiel

The Closed-Loop Robotic Glockenspiel is the first integrated robotic instrument in existence that is capable of continuously calibrating its own velocity and latency compensation characteristics in real time during its performance.

A Methodology for Evaluating Robotic Percussion Mechanisms

In Chapter 4, the first prescribed methodology for evaluating the performance of robotic percussion mechanisms is presented. This methodology is also used to carry out the first study that evaluates such mechanisms that includes pneumatic and servo-based robotic percussion devices in addition to linear and rotary solenoid-based ones.

A New Robotic Drum Kit

Informed by the study of robotic percussion mechanisms outlined above, a new robotic drum kit was designed and built with novel striker and mounting apparatuses. This new robotic drum kit is also the first created with specific consideration given to implementing audio-based closed-loop feedback, with custom internal mic'ing systems included.

An Expressive Hi-hat Movement System

Chapter 4 presents a new hi-hat movement system that utilises soft-shift solenoids driven by novel-to-this-device power amplifier-based driver circuits and microcontrollers, to enable subtle and dynamic control of the position of the top hi-hat cymbal in relation to the bottom one. The system provides fine-grained control over the height differential between the cymbals and pressure applied between the cymbals. It is also capable of performing 'chick' sounds and implementing a wide range of envelopes from very short tight strikes, to wide open strikes where the cymbals are not in contact with each other.

A Closed-Loop Modular Robotic Percussion Control System

Chapter 5 presents the first example of a modular robotic percussion control system that is created from the ground up to implement audio feedback-based closed-loop control with a wide range of different robotic percussion apparatuses.

A Highly Accurate FPGA-based Musical Robot Control Board

A new, hybrid, microcontroller-FPGA-based control board that makes significant improvements over the accuracy of traditional solutions is presented. By creating electronic pulses with widths that are accurate to within several nanoseconds, the control board achieves accuracy in musical robot performance that has heretofore been impossible. The board was also designed to work with the Mojo and Arduino development environments to increase accessibility to the wider field of musical robotics practitioners.

A New Modular Preamp Board

Because of the wide range of audio transduction solutions that may be required of a robotic percussion array, novel preamp modules were designed that implement padding, accept standard and high impedance inputs, accept multiple connector types, and include multiple selectable plug-in power standards.

A 32-Channel Input and Output Embedded DSP Board

To accommodate the large channel counts of high quality, low latency, embedded audio processing that closed-loop audio feedback systems necessitate, a novel DSP board with 32 channels of audio input and output was created. The board is customisable to suit various use cases, and programmable by a free, open source compiler.

A 16-Channel TDM to ADAT Conversion Module

To enable audio output in a commonly utilised commercial format, four new TDM to ADAT conversion boards were designed and trialled, with an Altera MAX10-based board being selected as the most appropriate for the unpitched percussion array controller. Custom IP was created that accepts two, eight-channel streams of I²S-style TDM signals, and converts these to ADAT format with a latency of a single sample.

Novel Algorithms for Embedded MIR for Musical Robots

New varieties of onset detection, amplitude detection, latency detection, pitch detection, and envelope length detection algorithms were created that are customised specifically for the implementation of audio-based closed-loop functionality with musical robots. These algorithms are utilised to implement built-in, continuous, real-time, low latency calibration of pitch, velocity, and hi-hat height, as well as to enable dynamic latency compensation in musical robots.

A DSP board for Space-Constrained Closed-Loop Musical Robotics

Due to the specific needs of robotic chordophones, much of the functionality of the larger 32-channel DSP board was condensed into a new 5 cm × 5 cm board. By combining actuator control and MIR functionality, this board can be integrated within musical robots with much smaller form-factors or in other space-constrained situations.

New Musical Works for Closed-loop Robotic Instruments

Coupled with the technical contributions outlined above, a number of original musical compositions were written to make specific use of the new instruments' capabilities. These musical pieces and installations have been performed and installed in a number of different venues from libraries, to art galleries and auditoriums. These new musical works act to demonstrate the novel capabilities provided by these developments and explore the new musical paradigms that the developments described in this thesis introduce.

7.2 Future Work

The creation of musical robots that have integrated audio-based closed-loop feedback systems opens up many new possibilities for these instruments, and makes possible many future developments. As all musical robots create sound in one way or another, any such instrument that has issues with manual calibration, reliability, or musical accuracy may benefit from the integration of a closed-loop audio feedback system. While this thesis worked on 'closing the loop' with regards to robotic instruments' amplitude, pitch, envelope length, and latency, much work is still to be done with other musical characteristics such as timbre and rhythm.

This thesis evaluated the effectiveness of these systems using a glockenspiel, a drum kit, and a chordophone as case studies. However, there were many candidates for implementation that did not get built, such as a robotic rattle drum whose rhythm is closed-loop, a robotic singing bowl whose timbre is closed-loop, and applications with wind instruments. There are also many other instruments that this system can be applied to directly with little change. For example, much of the Closed-Loop Robotic Glockenspiel system could be applied to xylophones, vibraphones, and marimba with little

change, the Closed-Loop Modular Percussion Array system can be applied to a wide range of instruments and actuators, and the Closed-Loop Chordophone system can be utilised with all manners of robotic guitars, zithers, and other string instruments.

While the hardware developed throughout the course of this research was targetted specifically at implementing audio-based closed loops with musical robots, there are applications outside of this field where these developments can also be utilised. In particular the hybrid actuator control board can be used with many different systems that require precise digital control of electromagnetic actuators, and the TDM to ADAT conversion boards can be adapted for use in a variety of digital audio systems. The multi-channel DSP board can also be utilised in large scale magnetic resonance and actuated instrument projects, as well as for implementing multi-channel synthesis and sound distribution.

There is also still much work to be done to realise the full potential of utilising this hardware to create novel robotic, actuated, and magnetic resonance instruments, along with hybrids of these three. For the ‘low level’ applications trialled thus far, the resources of the chosen DSP are sufficient and higher level MIR such as phrase and pattern detection also appear entirely possible. Future iterations of the board could also utilise more powerful, expensive DSPs in order to conduct more intensive analyses on many channels simultaneously. An example is the ADSP-BF7xx series that has superseded the BF5xx series during the course of this research. It is hoped that by further developing this dedicated embedded hardware, and continuing to optimise the software with more efficient and sophisticated algorithms, advances in the field of MIR, machine learning and AI can be utilised to a greater extent in musical robotics, pushing the capabilities of musical machines into new territory.

7.3 Research Outputs

Throughout the course of this research, a number of publicly visible outputs have been realised. These include several peer-reviewed publications, awards, public performances and installation works, and features in the media. Lists of these, delimited by category of output, are provided below.

7.3.1 Awards

OCTOBER 2017	International Computer Music Association Student Scholarship
SEPTEMBER 2017	Lilburn Trust NZSM Composers Competition 1 st Place
MARCH 2017	Susan Rhind Award in Composition
NOVEMBER 2016	Electronics New Zealand Conference Student Scholarship
SEPTEMBER 2015	International Computer Music Association Student Scholarship
JULY 2015	Freemasons Lankhuyzen / Whetu-Kairangi Grant
MAY 2015	D.F. McKenzie Award
MAY 2015	Victoria University of Wellington Doctoral Scholarship
APRIL 2015	Freemasons Lankhuyzen / Whetu-Kairangi Music Award
NOVEMBER 2014	Electronics New Zealand Conference Special Mention Award

7.3.2 Peer-Reviewed Publications

Throughout the various stages of the research presented in this thesis, a number of articles and conference papers were written. The topics range from reviews of related historical and contemporary literature, preliminary studies undertaken to establish gaps in current musical robotics research, experiments with alternative methods of controlling musical machines, and papers that describe the development of new robotic musical instruments. These publications demonstrate the various direct and tangential contributions this work has made to wider field of musical robotics and non-loudspeaker based sound-art. All of these research outputs were accepted by their respective conference and journal committees by way of peer-review processes, and are enumerated below.

- J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "Improving Robotic Chordophones with Closed-Loop Embedded Audio Feedback Systems". In: *Proceedings of the International Computer Music Conference, ICMC*. Shanghai, China: Shanghai Conservatory of Music, 2017
- J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "A Closed-Loop Control System for Robotic Hi-hats". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Copenhagen, Denmark: Aalborg University, 2017
- J. Long, J. W. Murphy, A. Kapur, and D.A. Carnegie. "Loudspeakers Optional: A History of Non-Loudspeaker-Based Electro-Acoustic Music". In: *Organised Sound* 22.2 (2017)

- J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "Improving the Accuracy of Musical Robots with Programmable Logic Circuits". In: *Proceedings of the Electronics New Zealand Conference, ENZCon*. Wellington, New Zealand: Victoria University of Wellington, 2016
- J. Long, A. Kapur, and D.A. Carnegie. "The Closed-Loop Robotic Glockenspiel: Improving Musical Robots Using Embedded Musical Information Retrieval". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Brisbane, Australia: Griffith University, 2016
- J. Long, A. Kapur, and D.A. Carnegie. "An Analogue Interface for Musical Robots". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Brisbane, Australia: Griffith University, 2016
- J. Long, J. Bailey, J. McVay, D.A. Carnegie, and A. Kapur. "Improving the Musical Expressiveness of Tesla Coils with Software". In: *Proceedings of the International Computer Music Conference, ICMC*. Denton, Texas: University of North Texas, 2015
- J. Long. "Augmenting Virtual Worlds with Musical Robotics". In: *Proceedings of the International Symposium of Electronic Art, ISEA*. Vancouver, Canada: Simon Fraser University, 2015
- J. Long, J. Murphy, A. Kapur, and D.A. Carnegie. "A Methodology for Evaluating Robotic Striking Mechanisms for Musical Contexts". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Baton Rouge, Louisiana: Louisiana State University, 2015
- J. Long, J. W. Murphy, A. Kapur, and D.A. Carnegie. "A Comparative Evaluation of Percussion Mechanisms for Musical Robotics Applications". In: *6th International Conference on Automation, Robotics and Applications, ICARA*. IEEE. 2015, pp. 173–178
- J. Long, J. W. Murphy, A. Kapur, and D.A. Carnegie. "The Robotic Taishogoto: A Systems Overview and Evaluation". In: *Proceedings of the Electronics New Zealand Conference, ENZCon*. Hamilton, New Zealand: University of Waikato, 2014
- J. Long. "The Robotic Taishogoto : A New Plug 'n Play Desktop Performance Instrument". In: *Proceedings of the International Conference on*

New Interfaces for Musical Expression, NIME. London, United Kingdom: Goldsmiths University of London, 2014

7.3.3 Musical Performances and Installations

- **Homage to Al-Jazari** for Robotic Taishogoto. *Engineering and Computer Science Showcase*, Victoria University of Wellington, 17th October 2014.
- **Stardust Emersion** for 3 Tesla coils, robotic bass guitar, and robotic drum set. *Chime Red Concert*, Victoria University of Wellington, 17th November 2014.
- **Homage to Al-Jazari** for Robot Ensemble. *Composers' Forum*, Adam Concert Room, New Zealand School of Music, 4th March 2015.
- **Robot Mario** Installation for Nintendo NES and robot ensemble. *The Hub*, Victoria University of Wellington, 21st July 2015.
- **Cicadas, Oscillators, and Tree Frogs** Arranged for 3 EMS Synthi AKS analogue synthesisers. *Lunchtime Concert Series*, National Library of New Zealand, 30th July 2015.
- **Robot Mario** Installation for Nintendo NES and robot ensemble. *The Hub*, Victoria University of Wellington, 28th August 2015.
- **Robot Mario** Installation for Nintendo NES and robot ensemble. *Science and Technology Festival*, Onslow College, 14th May 2016.
- **Broken Mirror** for Closed-Loop Robotic Glockenspiel and Quadraphonic Electronics. *Lunchtime Concert Series*, National Library of New Zealand, 4th August 2016.
- **Broken Mirror** for Closed-Loop Robotic Glockenspiel and Quadraphonic Electronics. *Lunchtime Concert*, Adam Art Gallery, 5th August 2016.
- **Broken Mirror** for Closed-Loop Robotic Glockenspiel and Quadraphonic Electronics. *Resonance*, Dowse Art Museum, 12th November 2016.
- **Broken Mirror** for Closed-Loop Robotic Glockenspiel and Quadraphonic Electronics. *Creativity Week*, The Hub, Victoria University of Wellington, 25th November 2016.

- **AutoReich**, Installation for Closed-Loop Robotic Glockenspiel. *Creativity Week*, The Hub, 25th November 2016.
- **The Impossible Drum Solo** for Robotic Drum Kit and Loudspeakers. *TEDxWellington*, St. James Theatre, 18th June 2017.
- **Broken Mirror** for Closed-Loop Robotic Glockenspiel and Quadraphonic Electronics. *Lilburn Trust NZSM Composers Competition*, Adam Concert Room, Victoria University of Wellington, 30th September 2017.

7.3.4 Media, Press

Radio Coverage

- Interview and Musical Robotics Performance on *Music 101*, Radio New Zealand. 18th October 2014.
- Interview on *Standing Room Only: Robotics and Music of the Future*, Radio New Zealand. 16th November 2014.
- Interviewed by Alison Ballance on *Our Changing World*, Radio New Zealand. 20th November 2014.
- Interviewed by Emma Smith on *Music 101 Pocket Edition 22*, Radio New Zealand. 6th December 2014.
- Interview on *Radio New Zealand*, aired 4:30pm 5th September 2015.

Internet Coverage

- Interview on *The Wireless*, <http://thewireless.co.nz/articles/bright-sparks-making-music> 21st November 2014.
- Interviewed by Sonia Sly on *Sly on the Wall: Making Musical Robots*, <http://slyonthewall.com/2014/12/18/making-musical-robots/> 18th December 2014.
- Featured on *Adafruit*, <https://blog.adafruit.com/2015/03/30/> 30th March 2015.

- Featured on *Synthtopia*, <http://www.synthtopia.com/content/2015/04/05/drum-bass-tesla-coils-music-robots/>, April 5th 2015.
- Interview on *The Wireless*, <http://www.radionz.co.nz/national/programmes/nat-music/audio/201769401/super-mario-musical-robots> 5th September 2015.
- Featured on *TEDxWellington* <https://www.youtube.com/watch?v=hxCMSyx-j0E> 27th June 2017.

7.4 Concluding Remarks

Throughout the past millenium and longer, people have invented machines to create new sonic experiences and to express themselves musically. In the present era, the development of musical robots has continued this rich tradition by allowing us to realise exciting and interesting new music in the acoustic sound space with precise, interactive, and synchronous computer control. This thesis began by identifying one of the key weaknesses of musical robots when compared with humans: they currently cannot listen to, understand, and improve their own performance in real time while they play.

The proposed solution to this issue is implementing audio feedback-based closed-loop control systems in musical robots to simulate the biological closed loop that occurs between a human and their musical instrument while they play. This thesis has presented several new robotic musical instruments from three instrument categories, and described the creation of novel control subsystems and algorithms that implement these developments. It has also evaluated the effectiveness of these systems through laboratory analyses and performances in a variety of real-world installation and performance environments.

The results of this research show that audio feedback-based embedded closed-loop control systems can provide significant improvements in the performance of musical robots. It can make them more robust, accurate, and reliable, easier to set up and make music with, and can provide additional expressive parameters for musical performance. These improvements can help to advance the state of the art in the field of musical robotics and make

---

these machines more useful, accessible, and musical. However, this document merely marks an early point in the development of these systems and it is hoped that musical roboticists will be inspired to duplicate and extend these systems to create exciting new music in the future.

Appendix A

Track Descriptions

Chapters 3 and 5 present original compositions created for the Closed-Loop Robotic Glockenspiel and Closed-Loop Robotic Drum Kit respectively. Both of the works are realised using digital audio workstations. In order to demonstrate the technical implementation of the works, screenshots of the project files of the pieces are presented in this appendix and descriptions are provided that outline the functionality of each of the tracks used.

A.1 Broken Mirror

1. This channel is the main MIDI track of the piece, where all of the individual non-effected Note On commands with velocity information for the glockenspiel are sequenced. An excerpt from this channel is shown in Figure 3.25.
2. The MIDI sequences in channel 2 are also sent to the instrument, but are sent via a MIDI arpeggiator plugin that distributes held chords among their constituent notes in various patterns and at dynamic speeds.
3. The live audio input channel from the glockenspiel. This channel contains a number of insert effects such as denoising and EQ that condition the signal before creative processing. This channel is accompanied by 4 automation lanes, each of which controls the level of the audio signal that is sent to each of the 4 FX channels outlined henceforth.
4. FX channel 1 - This channel contains a 'Prism' processor that resembles a chorus and pitch bending effect.
5. FX channel 2 - A delay effect is utilised to add specialised echoes to specific areas of the piece, and the delay time and feedback levels are

FIGURE A.1: Screenshot of the Broken Mirror DAW Project.

creatively automated to speed up, slow down, and alter the length and intensity of echo effects. Two automation lanes are also utilised to pan the sound being processed in FX channel 2 with precise positions and trajectories throughout 2 dimensional space within the quadraphonic loudspeaker array.

6. FX channel 3 - This channel is primarily used as a reverberation effect, simulating a range of acoustic environments from small confined spaces, to large open ones. The combination of send level and reverb decay time automation allows a number of environmental possibilities and independent quadraphonic panning capability is also provided on this channel. A dynamic pitch shifter is also used on this channel to bend the pitch of certain reverb tails, and bypass and bend amount automation lanes are shown.
7. FX channel 4 - A pair of automatable samplers are utilised on this channel to carry out a range of functions relating to recording segments of the live sound, and replaying those recordings at different times and at different pitches. Given that this glockenspiel has a limited 21 note range, additional pitches can be achieved by recording live notes from the instrument, and playing them back at pitches outside its range. The automation lanes attached to this channel are concerned with sending audio to each sampler and specialising the output audio via the loudspeaker array.
8. This MIDI channel controls the first sampler housed in FX channel 4. This sampler is set to a reverse configuration so that after sound information has been recorded into it, subsequent playbacks of the sound are reversed in time. A certain MIDI note command is dedicated to triggering and terminating a recording event, associating the recorded note with another designated MIDI note. Subsequent notes then trigger playback of the audio at a pitch that is transposed in relation to that designated MIDI note.
9. This MIDI channel controls the second sampler housed in FX channel 4. This channel behaves identically to channel 8 with the exception that the sampler is set to a forward playback configuration instead of reverse.

A.2 The Impossible Drum Solo: Part Two

FIGURE A.2: Screenshot of The Impossible Drum Solo Part Two DAW Project.

1. The main MIDI track of the piece, where all of the Note On and MIDI CC commands with velocity information for the robotic drum kit are sequenced. The content inside this track resembles the piano rolls shown in Figures 5.30, 5.31, and 5.32.
2. The kick drum audio input track. Channel strip type effects such as EQ and compression are applied before the audio is routed to track 3.

3. Track that houses the creative effects applied to the kick, and routes the kick audio to a sampler.
4. MIDI channel that controls the kick sampler. The sampler records the kick's audio at specific points, and plays back the sound in reverse, at various pitches.
5. Audio input for the microphone pointed at the snare drum's top head. Channel strip type effects are applied here.
6. Audio input for the microphone pointed at the snare drum's bottom head. Channel strip type effects are applied here.
7. The snare top and bottom tracks are routed to this group track to mix and process them together. Two automation lanes control this track's sends, dynamically routing the signal in varying amounts to channels 8 and 10. This track also houses the snare's sampler.
8. Snare reverb track. When the snare is routed to this track, it applies different levels of reverberation to the snare instrument specifically. Two automation lanes manipulate the amount of reverberation.
9. MIDI channel that controls the snare sampler. The sampler records the snare's audio at specific points, and plays back the sound in reverse, at various pitches.
10. Snare delay effects track. When the snare is routed to this track, it applies delay and echo effects to it. Two automation lanes control the delay's feedback and delay-time parameters.
11. Hi tom audio input track. This track applies channel strip type effects, and utilises an automation lane to send the signal in varying amounts to track 12.
12. Hi tom effects track. When the hi tom is routed to this track, it applies panning and delay effects to it. Three automation lanes control the ping-pong delay's feedback and timing parameters, and a left-right panner.
13. Mid tom audio input track. This track applies channel strip type effects, and utilises an automation lane to send the signal in varying amounts to track 14.
14. Mid tom effects track. When the mid tom is routed to this track, it applies spectral hold, live pitch shifting, and reverberation effects to

- it. Four attached automation lanes control the hold switch, pitch, and reverb parameters.
15. MIDI channel that controls the mid tom sampler. The sampler records the mid tom's audio at specific points, and plays back the sound at various pitches.
  16. Floor tom audio input track. This track applies channel strip type effects, and utilises an automation lane to send the signal in varying amounts to track 17.
  17. Floor tom effects track. When the floor tom is routed to this track, it applies spectral hold and live pitch-shifting effects to it. Two attached automation lanes control the hold switch and pitch parameters.
  18. Group track that receives all of the tom tracks for group processing.
  19. Hi-hat audio input track. This track applies channel strip type effects, and utilises an automation lane to send the signal in varying amounts to track 20.
  20. Track that houses the creative effects applied to the hi-hat.
  21. Ride cymbal audio input track. This track applies channel strip type effects, and utilises an automation lane to send the signal in varying amounts to track 22.
  22. Track that houses the creative effects applied to the ride cymbal.
  23. Crash cymbal audio input track. This track applies channel strip type effects, and utilises an automation lane to send the signal in varying amounts to track 24.
  24. Track that houses the creative effects applied to the crash, and routes the crash audio to a sampler.
  25. MIDI channel that controls the crash sampler. The sampler records the crash's audio at specific points, and plays back the sound in reverse, at various pitches.
  26. Group channel where all of the channels of the kit are mixed for combined processing. Also contains a sampler.
  27. MIDI channel that controls the sampler of the entire kit. The sampler records the kit's audio at specific points, and plays back the sound in reverse.

28. General purpose reverb. Various amounts of signal are sent from various channels to the fixed reverb to provide the piece with cohesive space among the manipulated space.

Bibliography

- [1] Aphex Twin. *Computer Controlled Acoustic Instruments pt2*. EP. Warp Records. 2015.
- [2] Mituhiko Araki. "PID control". In: *Control systems, robotics and automation 2* (2002), pp. 1–23.
- [3] Massimo Banzi. *Getting Started with Arduino*. Ill. Sebastopol, CA: Make Books - Imprint of: O'Reilly Media, 2008. ISBN: 0596155514, 9780596155513.
- [4] Frank Barnes, Markus Kolb, and Stock Plum. *Building a Robot Lead Vocalist and Producing First Album*. URL: <https://www.kickstarter.com/projects/1335681368/building-robot-lead-vocalist-and-producing-first-a> (visited on 08/07/2017).
- [5] K. Barr, A. Zak, M. Ryle, D. Brown, and C. Lafky. *Method and apparatus for providing a digital audio interface protocol*. US Patent 5,297,181. 1994. URL: <http://www.google.com/patents/US5297181>.
- [6] Scott Barton, Ethan Prihar, and Paulo Carvalho. "Cyther: a human-playable, self-tuning robotic zither". In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. Copenhagen, Denmark: Aalborg University, 2017.
- [7] Filip Biedrzycki, Jakub Knast, Mariusz Nowak, and Jakub Paszkowski. "AutoPRK-Automatic Drum Player". In: *14th Computer Information Systems and Industrial Management (CISIM)*. Springer. 2015, pp. 592–600.
- [8] Björk. *Biophilia*. Album. One Little Indian Records. 2011.
- [9] Sinan Bökesoy and Patrick Adler. "1city1001vibrations: development of a interactive sound installation with robotic instrument performance". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Oslo, Norway: University of Oslo, 2011.
- [10] S. Bonner. *The History and Organology of the Aeolian Harp: Text*. Aeolian harp. Bois de Boulogne, 1970. ISBN: 9780900998119.
- [11] Bruce Bower. "Stone age flutes found in Germany". In: *A+A* 176.2 (2009), p. 13.

- [12] Mason Bretan and Gil Weinberg. "A Survey of Robotic Musicianship". In: *Communications of the ACM* 59.5 (2016), pp. 100–109.
- [13] Mason Bretan and Gil Weinberg. "Chronicles of a Robotic Musical Companion". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. London, United Kingdom: Goldsmiths University of London, 2014.
- [14] Mason Bretan, Deepak Gopinath, Philip Mullins, and Gil Weinberg. "A Robotic Prosthesis for an Amputee Drummer". In: *CoRR* abs/1612.04391 (2016). URL: <http://arxiv.org/abs/1612.04391>.
- [15] N. Cameron Britt, Jeff Snyder, and Andrew McPherson. "The EMvibe: an Electromagnetically Actuated Vibraphone". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Ann Arbor, Michigan: University of Michigan, 2012.
- [16] Hannah I.T. Brown. "Classics on Command". In: *Stanford Alumni Magazine* (2015).
- [17] Ken Caulkins. *Ragtime Automated Music*. URL: <http://www.ragtimewest.com/> (visited on 07/21/2015).
- [18] Dustin Cohen. *The Chalice Symphony for Stella Artois (Full Doc)*. URL: <https://vimeo.com/105684492> (visited on 07/27/2015).
- [19] F.B. Crane. *Athanasius Kircher, Musurgia Universalis* (Rome, 1650). 2014.
- [20] John Croft. "Theses on liveness". In: *Organised Sound* 12.1 (2007), pp. 59–66.
- [21] Dimitri Diakopoulos and A. Kapur. "HIDUINO: A firmware for building driverless USB-MIDI devices using the Arduino microcontroller". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Oslo, Norway: University of Oslo, 2011.
- [22] Simon Dixon. "Onset Detection Revisited". In: *Proceedings of the International Conference on Digital Audio Effects*. DAFX. Montreal, Canada, 2006.
- [23] Deepak Edakkattil Gopinath. "Enhancing Stroke Generation and Expressivity in Robotic Drummers-A Generative Physics Model Approach". Masters Thesis. 2015.
- [24] Alice Eldridge and Chris Kiefer. "The Self-resonating Feedback Cello: Interfacing gestural and generative processes in improvised performance". In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. Copenhagen, Denmark: Aalborg University, 2017, pp. 25–29.

- [25] S. Emmerson. *The Language of Electroacoustic Music*. Macmillan, 1986. ISBN: 9780333397596. URL: <https://books.google.co.nz/books?id=ZusXAQAIAAJ>.
- [26] *Encyclopædia Britannica Online*, s. v. "barrel piano". URL: <http://global.britannica.com/art/barrel-piano> (visited on 07/21/2015).
- [27] Peter Esmonde. *Trimpin - The Sound of Invention*. Motion picture. Microcinema International, DVD. 2011.
- [28] Michael A. Fabio. "The Chandelier: An Exploration in Robotic Musical Instrument Design". Masters Thesis, Massachusetts Institute of Technology. 2007.
- [29] Christian Faubel. "Rhythm Apparatus on Overhead". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. London, United Kingdom: Goldsmiths University of London, 2014.
- [30] Asbjorn Blokkum Flo and Hans Wilmers. "Doppelganger: A solenoid-based large scale sound installation." In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Baton Rouge, Louisiana: Louisiana State University, 2015.
- [31] Anne Focke ed. *Trimpin: Contraptions for Art and Sound*. Seattle, Washington: Marquand Books, 2011. ISBN: 0295991097.
- [32] Charles B Fowler. "The Museum of Music: A History of Mechanical Instruments". In: *Music Educators Journal* (1967), pp. 45–49.
- [33] Anders Friberg and Johan Sundberg. "Time discrimination in a monotonic, isochronous sequence". In: *The Journal of the Acoustical Society of America* 98.5 (1995), pp. 2524–2531.
- [34] *Full Spectrum Laser*. URL: <https://fslaser.com/> (visited on 07/15/2015).
- [35] S. Futterman. *Player Violin Introduced*. URL: http://www.qgrsmusic.com/press_details.asp?rid=19 (visited on 09/07/2016).
- [36] Nicolas Gadd and D. A. Carnegie. "A High Performance Drumming Robot". In: *Proceedings of the Electronics New Zealand Conference, ENZ-Con*. Hamilton, New Zealand: University of Waikato, 2014.
- [37] Marritt Gally. "The Autophone". In: *Scientific American* 40.23 (1879), p. 354.
- [38] K. Gann. *The Music of Conlon Nancarrow*. Cambridge University Press, 1995.

- [39] Moritz Simon Geist. *The MR-808 robotic HiHat*. URL: <http://learning.sonicrobots.com/2016/08/27/the-mr-808-robotic-hihat/> (visited on 07/27/2015).
- [40] Asif Ghias, Jonathan Logan, David Chamberlin, and Brian C. Smith. "Query by Humming: Musical Information Retrieval in an Audio Database". In: *Proceedings of the third ACM international conference on Multimedia*. ACM. 1995, pp. 231–236.
- [41] Dan Grunebaum. "Maywa Denki: Nobumichi Tosa's whimsical creations sound the battle squeak against digital music". In: *Metropolis Magazine* 870 (2010).
- [42] Michael Gurevich. "Distributed Control in a Mechatronic Musical Instrument." In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. NIME. Baton Rouge, Louisiana: Louisiana State University, 2015.
- [43] T.L. Hankins and R.J. Silverman. *Instruments and the Imagination*. Princeton University Press, 1999. ISBN: 9780691005492.
- [44] Jiffer Harriman. "Feedback Lapsteel: Exploring Tactile Transducers as String Actuators." In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. NIME. Baton Rouge, Louisiana: Louisiana State University, 2015.
- [45] Jiffer Harriman, Michael Theodore, and Mark Gross. "The Kitsch-Instrument: Hackable Robotic Music". In: *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*. TEI '15. Stanford, California, USA: ACM, 2015, pp. 141–144. ISBN: 978-1-4503-3305-4. DOI: 10.1145/2677199.2680593. URL: <http://doi.acm.org/10.1145/2677199.2680593>.
- [46] Jiffer Harriman, Matthew Bethancourt, Abhishek Narula, Michael Theodore, and Mark Gross. "SolidNoise: Tools For Making Musical Robots". In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM. 2016, pp. 3683–3686.
- [47] Jingyin He, A. Kapur, and D.A. Carnegie. "Developing a Physical Gesture Acquisition System for Guqin Performance". In: *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME. Baton Rouge, Louisiana: Louisiana State University, 2015.
- [48] J.W. Humphrey, J.P. Oleson, and A.N. Sherwood. *Greek and Roman Technology: A Sourcebook : Annotated Translations of Greek and Latin Texts and Documents*. Classics. Ancient history. Routledge, 1998. ISBN: 9780415061377.

- [49] Andy Hunt and Marcelo M. Wanderley. "Mapping performer parameters to synthesis engines". In: *Organised sound* 7.02 (2002), pp. 97–108.
- [50] I.R. Jazarī. *The Book of Knowledge of Ingenious Mechanical Devices*. Reidel, 1974. ISBN: 9789027703293.
- [51] Wonse Jo, Bumjoo Lee, and Donghan Kim. "Development of Auditory Feedback System for Violin Playing Robot". In: *International Journal of Precision Engineering and Manufacturing* 17.6 (2016), pp. 717–724. ISSN: 2005-4602. DOI: 10.1007/s12541-016-0089-6. URL: <http://dx.doi.org/10.1007/s12541-016-0089-6>.
- [52] Wonse Jo, Hyeonjun Park, Bumjoo Lee, and Donghan Kim. "A Study on Improving Sound Quality of Violin Playing Robot". In: *Proceedings of the 6th International Conference on Automation, Robotics and Applications, ICARA*. Queenstown, New Zealand, 2015.
- [53] C. Johns. *Sex Or Symbol?: Erotic Images of Greece and Rome*. Routledge, 1982. ISBN: 9780415925679.
- [54] Sergi Jordà. "Afasia: the Ultimate Homeric One-man-multimedia-band". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Singapore: National University of Singapore, 2002.
- [55] Shinya Kaneko, Tatsuya Sono, Shinichiro Nakamura, and Junichi Sakaguchi. *Oto Asobi! Blackfin DSP Kiban de Dijitaru Shingo Shori Shotaiken*. CQ Publishing, 2015.
- [56] A Kapur. "A History of Robotic Musical Instruments". In: *Proceedings of the International Computer Music Conference, ICMC*. Barcelona, Spain: Escola Superior de Música de Catalunya, 2005.
- [57] A. Kapur. *Digitizing North Indian Music: Preservation and Extension Using Multimodal Sensor Systems, Machine Learning and Robotics*. University of Victoria (Canada), 2007. ISBN: 9780494374412.
- [58] A. Kapur and Michael Darling. "A Pedagogical Paradigm for Musical Robotics." In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. NIME. Sydney, Australia: University of Technology Sydney, 2010, pp. 162–165.
- [59] A. Kapur, Trimpin, E. Singer, A. Suleman, and G. Tzanetakis. "A Comparison of Solenoid-Based Strategies for Robotic Drumming". In: *Proceedings of the International Computer Music Conference (ICMC)*. Copenhagen, Denmark, 2007.
- [60] A. Kapur, Michael Darling, Meason Wiley, Ben Womick, John Aspinall, Roger Marcelo, and Duncan Woodbury. *Breakbot*. URL: [http:](http://)

- / / www.karmetik.com/robot/breakbot (visited on 08/07/2017).
- [61] A. Kapur, Eric Singer, Manjinder S. Benning, George Tzanetakis, and Trimpin. "Integrating HyperInstruments, Musical Robots & Machine Musicianship for North Indian Classical Music". In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. NIME. New York, New York: New York University, 2007, pp. 238–241.
  - [62] A. Kapur, Murphy J. W., Michael Darling, Eric Heep, Bruce Lott, and Ness Morris. "MalletOTon and the Modulets: Modular and Extensible Musical Robots". In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. Vol. 16. Brisbane, Australia: Queensland Conservatorium Griffith University, 2016, pp. 69–72. ISBN: 978-1-925455-13-7. URL: http://www.nime.org/proceedings/2016/nime2016_paper0015.pdf.
  - [63] A. Kapur, Jordan Hochenbaum, Michael Darling, Dimitri Diakopoulos, and J. W. Murphy. "The karmetik notomoton: A new breed of musical robot for teaching and performance". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Oslo, Norway: University of Oslo, 2011, pp. 228–31.
  - [64] A. Kapur, Michael Darling, Dimitri Diakopoulos, J. W. Murphy, Jordan Hochenbaum, Owen Vallis, and Curtis Bahn. "The Machine Orchestra: An Ensemble of Human Laptop Performers and Robotic Musical Instruments". In: *Computer Music Journal* 35.4 (2011), pp. 49–63.
  - [65] Steven Kemper. "Composing for Musical Robots: Aesthetics of electromechanical music". In: *Proceedings of the Korean Electro-Acoustic Music Society's Annual Conference, KEAMSAC*. Seoul, Korea: Seoul National University, 2014.
  - [66] Steven Kemper and Rebecca Cypess. "Can Musical Machines Be Expressive? Views from the Enlightenment and Today". In: *Leonardo Music Journal* (). URL: http://dx.doi.org/10.1162/LEON_a_01477.
  - [67] Keisuke Kimoto. *Kimoto*. URL: http://onkansotsuten.com/2016/6_kimoto/works.html (visited on 05/11/2017).
  - [68] Thaddeus Kochanny. *A Music Box History*. URL: http://www.amica.org/Live/Instruments/Music_Boxes/Music_box_history.htm (visited on 07/19/2015).

- [69] Teun Koetsier. "On the prehistory of programmable machines: musical automata, looms, calculators". In: *Mechanism and Machine theory* 36.5 (2001), pp. 589–603.
- [70] Michael Krzyzaniak. *Timbral Learning for Musical Robots*. PhD Thesis, Arizona State University. 2016.
- [71] Sasha Leitman. "Trimpin: An Interview". In: *Computer Music Journal* 35.4 (2011), pp. 12–27.
- [72] Dan Livingstone and Eduardo Miranda. "ORB3–MUSICAL ROBOTS WITHIN AN ADAPTIVE SOCIAL COMPOSITION SYSTEM". In: *Proceedings of the International Computer Music Conference, ICMC*. Barcelona, Spain: Escola Superior de Música de Catalunya, 2005.
- [73] J. Long. "Augmenting Virtual Worlds with Musical Robotics". In: *Proceedings of the International Symposium of Electronic Art, ISEA*. Vancouver, Canada: Simon Fraser University, 2015.
- [74] J. Long. "The Development of an Ensemble of Automatic Instruments and their Possibilities". Masters Thesis, Tokyo University of the Arts. 2014.
- [75] J. Long. "The Robotic Taishogoto : A New Plug 'n Play Desktop Performance Instrument". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. London, United Kingdom: Goldsmiths University of London, 2014.
- [76] J. Long, A. Kapur, and D.A. Carnegie. "An Analogue Interface for Musical Robots". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Brisbane, Australia: Griffith University, 2016.
- [77] J. Long, A. Kapur, and D.A. Carnegie. "The Closed-Loop Robotic Glockenspiel: Improving Musical Robots Using Embedded Musical Information Retrieval". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Brisbane, Australia: Griffith University, 2016.
- [78] J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "A Closed-Loop Control System for Robotic Hi-hats". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Copenhagen, Denmark: Aalborg University, 2017.
- [79] J. Long, J. W. Murphy, A. Kapur, and D.A. Carnegie. "A Comparative Evaluation of Percussion Mechanisms for Musical Robotics Applications". In: *6th International Conference on Automation, Robotics and Applications, ICARA*. IEEE. 2015, pp. 173–178.

- [80] J. Long, J. Murphy, A. Kapur, and D.A. Carnegie. "A Methodology for Evaluating Robotic Striking Mechanisms for Musical Contexts". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Baton Rouge, Louisiana: Louisiana State University, 2015.
- [81] J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "Improving Robotic Chordophones with Closed-Loop Embedded Audio Feedback Systems". In: *Proceedings of the International Computer Music Conference, ICMC*. Shanghai, China: Shanghai Conservatory of Music, 2017.
- [82] J. Long, J. W. Murphy, D.A. Carnegie, and A. Kapur. "Improving the Accuracy of Musical Robots with Programmable Logic Circuits". In: *Proceedings of the Electronics New Zealand Conference, ENZCon*. Wellington, New Zealand: Victoria University of Wellington, 2016.
- [83] J. Long, J. Bailey, J. McVay, D.A. Carnegie, and A. Kapur. "Improving the Musical Expressiveness of Tesla Coils with Software". In: *Proceedings of the International Computer Music Conference, ICMC*. Denton, Texas: University of North Texas, 2015.
- [84] J. Long, J. W. Murphy, A. Kapur, and D.A. Carnegie. "Loudspeakers Optional: A History of Non-Loudspeaker-Based Electro-Acoustic Music". In: *Organised Sound* 22.2 (2017).
- [85] J. Long, J. W. Murphy, A. Kapur, and D.A. Carnegie. "The Robotic Taishogoto: A Systems Overview and Evaluation". In: *Proceedings of the Electronics New Zealand Conference, ENZCon*. Hamilton, New Zealand: University of Waikato, 2014.
- [86] Tod Machover and Joe Chung. "Hyperinstruments: Musically Intelligent and Interactive Performance and Creativity systems". In: *Proceedings of the International Computer Music Conference, ICMC*. San Francisco, California, 1989.
- [87] Laura Maes, Godfried-Willem Raes, and Troy Rogers. "The Man and Machine Robot Orchestra at Logos". In: *Computer Music Journal* 35.4 (2011), pp. 28–48.
- [88] *Makerbot Thingiverse*. URL: <http://www.thingiverse.com> (visited on 07/15/2015).
- [89] Leonard M. Marcus. *Encyclopaedia Britannica Online*, s.v. "music recording". URL: <https://www.britannica.com/topic/music-recording> (visited on 07/20/2015).

- [90] Pat Matheny. *The Orchestrion Project*. Blu-ray. Directed by Pierre and Francois Lamoureux, Eagle Rock Entertainment. 2012.
- [91] Paul Mathews, Ness Morris, J. W. Murphy, A. Kapur, and D.A. Carnegie. "Tangle: a Flexible Framework for Performance With Advanced Robotic Musical Instruments". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. London, United Kingdom: Goldsmiths University of London, 2014.
- [92] Louis J. McCallum. "Friend Me Your Ears: A Musical Approach to Human-Robot Relationships". PhD Thesis, Queen Mary University of London. 2016.
- [93] Louis J. McCallum. "Interactions With A Robot Percussionist". Masters Thesis, University of Sussex. 2010.
- [94] A. McPherson, R. Jack, and G. Moro. "Action-Sound Latency: Are Our Tools Fast Enough?" In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Brisbane, Australia: Griffith University, 2016.
- [95] Andrew McPherson. "Techniques and Circuits for Electromagnetic Instrument Actuation." In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. Ann Arbor, Michigan: University of Michigan, 2012.
- [96] Andrew McPherson. "The Magnetic Resonator Piano: Electronic Augmentation of an Acoustic Grand Piano". In: *Journal of New Music Research* 39.3 (2010), pp. 189–202.
- [97] J. McVay, D.A. Carnegie, and J. W. Murphy. "An Overview of MechBass: A Four String Robotic Bass Guitar". In: *Proceedings of the 6th International Conference on Automation, Robotics and Applications, ICARA*. 2015, pp. 179–184.
- [98] James McVay, D.A. Carnegie, J.W. Murphy, and A. Kapur. "Mechbass: A systems overview of a new four-stringed robotic bass guitar". In: *Proceedings of the 19th Electronics New Zealand Conference, ENZCon*. Palmerston North, New Zealand: Massey University, 2011.
- [99] Carolina Brum Medeiros and Marcelo M Wanderley. "A Comprehensive Review of Sensors and Instrumentation Methods in Devices for Musical Expression". In: *Sensors* 14.8 (2014).
- [100] Pete Millett. "Power Transformers for Audio Equipment". In: *audioXpress*. 2001.
- [101] Robert A. Moog and Thomas L. Rhea. "Evolution of the Keyboard Interface: The Bösendorfer 290 SE Recording Piano and the Moog

- Multiply-Touch-Sensitive Keyboards". In: *Computer Music Journal* 14.2 (1990), pp. 52–60. ISSN: 01489267. URL: <http://www.jstor.org/stable/3679712>.
- [102] J. W. Murphy. "Expressive Musical Robots: Building, Evaluating, and Interfacing with an Ensemble of Mechatronic Instruments". PhD Thesis, Victoria University of Wellington. 2014.
- [103] J. W. Murphy, D.A. Carnegie, and A. Kapur. "Little Drummer Bot: Building, Testing, and Interfacing With a New Expressive Mechatronic Drum System". In: *Proceedings of the International Computer Music Conference, ICMC*. Athens, Greece: National and Kapodistrian University of Athens, 2014.
- [104] J. W. Murphy and A. Kapur. "Swivel: Analysis and Systems Overview of a New Robotic Guitar". In: *Proceedings of the International Computer Music Conference, ICMC*. Perth, Western Australia: Tura New Music, 2013.
- [105] J. W. Murphy, A. Kapur, and D.A. Carnegie. "Better Drumming Through Calibration: Techniques for Pre-Performance Robotic Percussion Optimization". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Ann Arbor, Michigan: University of Michigan, 2012.
- [106] J. W. Murphy, A. Kapur, and D.A. Carnegie. "Musical Robotics in a Loudspeaker World: Developments in Alternative Approaches to Localization and Spatialization". In: *Leonardo Music Journal* 22 (2012), pp. 41–48.
- [107] J. W. Murphy, A. Kapur, and D.A. Carnegie. "Swivel 2: A Systems Overview and Evaluation of a New Mechatronic Slide Guitar". In: *Proceedings of the Engineering New Zealand Conference, ENZCon*. Auckland, New Zealand: Massey University, 2013.
- [108] J. W. Murphy, J. McVay, P. Mathews, D. A. Carnegie, and A. Kapur. "Expressive Robotic Guitars: Developments in Musical Robotics for Chordophones". In: *Computer Music Journal* (2015).
- [109] J. W. Murphy, Paul Mathews, A. Kapur, and D.A. Carnegie. "Robot, Tune Yourself: Automatic Tuning in Musical Robotics". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. London, United Kingdom: Goldsmiths University of London, 2014.
- [110] Jim Murphy, A. Kapur, and D.A. Carnegie. "Interacting With Solenoid Drummers: A Quantitative Approach to Composing and Performing

- With Open-Loop Solenoid-Based Percussion Systems". In: *Proceedings of Australasian Computer Music Conference, ACMC*. Brisbane, Queensland, Australia: Griffith University, 2012.
- [111] Steven R. Ness, Shawn Trail, Peter F. Driessen, Andrew Schloss, and George Tzanetakis. "Music Information Robotics: Coping Strategies for Musically Challenged Robots." In: *Proceedings of the International Society for Music Information Retrieval Conference, ISMIR*. Miami, Florida: University of Miami, 2011.
- [112] Craig Newswanger. *Percussibot Drum Controller Kit, Available Soon*. URL: <http://www.resonancestudio.com/craigsblog/?p=327> (visited on 07/25/2017).
- [113] Craig Newswanger. *Resonance Studio Workshop: Midi Controlled Drum Set*. URL: http://www.resonancestudio.com/craigsblog/?page_id=174 (visited on 08/07/2017).
- [114] NIME '07: *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*. New York, New York: ACM, 2007.
- [115] Hyunjoo Oh, Jiffer Harriman, Abhishek Narula, Mark D. Gross, Michael Eisenberg, and Sherry Hsi. "Crafting Mechatronic Percussion with Everyday Materials". In: *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM. 2016, pp. 340–348.
- [116] A.W.J.G. Ord-Hume. *The Musical Box: A Guide for Collectors, Including a Guide to Values*. Schiffer Pub., 1995. ISBN: 9780887407642.
- [117] Dan Overholt, Edgar Berdahl, and Robert Hamilton. "Advancements in Actuated Musical Instruments". In: *Organised Sound* 16.02 (2011), pp. 154–165.
- [118] Daniel Overholt. "The Overtone Fiddle: an Actuated Acoustic Instrument". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Oslo, Norway: University of Oslo, 2011.
- [119] Mark Owen. *Practical Signal Processing*. Cambridge University Press, 2007.
- [120] Cléo Palacio-Quintin. "The hyper-flute". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Singapore: National University of Singapore, 2003.
- [121] R. Palmieri. *The Piano: An Encyclopedia*. Taylor & Francis, 2004. ISBN: 9781135949631.
- [122] T. Patteson. *Instruments for New Music: Sound, Technology, and Modernism*. University of California Press, 2015.

- [123] Klaus Petersen, Jorge Solis, Takeshi Ninomiya, Tetsuro Yamamoto, Masaki Takeuchi, and Atsuo Takanishi. "Development of the anthropomorphic saxophonist robot WAS-1: Mechanical design of the lip, tonguing, fingers and air pump mechanisms". In: *Proceedings of the International Conference on Robotics and Automation*. IEEE. 2009, pp. 3043–3048.
- [124] J.A. Pfirrmann. *The Age of the Music-box*. Passaic County Historical Society, 1969.
- [125] Andrew Piepenbrink and Matthew Wright. "The Bistable Resonator Cymbal: An Actuated Acoustic Instrument Displaying Physical Audio Effects". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Baton Rouge, Louisiana: Louisiana State University, 2015.
- [126] Godfried Willem Raes. *Expression control in automated musical instruments*. URL: http://logosfoundation.org/g_texts/expression-control.html (visited on 09/02/2017).
- [127] Godfried Willem Raes. *<HAT>: a hit any thing percussion robot*. URL: http://logosfoundation.org/instrum_gwr/HAT.html (visited on 08/07/2017).
- [128] Godfried Willem Raes. *Logos: A quick composers guide to the Logos Robot orchestra*. URL: http://logosfoundation.org/instrum_gwr/manual.html (visited on 08/17/2015).
- [129] Godfried Willem Raes. *Logos M&M: The World's Largest Robot Orchestra*. URL: <http://logosfoundation.org/mnm/index.htm> (visited on 07/15/2015).
- [130] Godfried Willem Raes. *<Rodo>: an automated set of bronze tines*. URL: http://logosfoundation.org/instrum_gwr/rodo.html (visited on 08/01/2015).
- [131] Godfried Willem Raes. *<Snar> an automated snare drum*. URL: http://logosfoundation.org/instrum_gwr/snar.html (visited on 08/09/2017).
- [132] Godfried Willem Raes. *<Snar_2> an improved automated snare drum made to the order of Aphex Twin*. URL: http://logosfoundation.org/instrum_gwr/snar2.html (visited on 07/13/2017).
- [133] Godfried Willem Raes. *<Vibi>: an automated vibraphone*. URL: http://logosfoundation.org/instrum_gwr/vibi.html (visited on 08/01/2015).

- [134] Colin Raffel, Brian McFee, Eric J. Humphrey, Justin Salamon, Oriol Nieto, Dawen Liang, and Daniel P. W. Ellis. “mir_eval: a Transparent Implementation of Common MIR Metrics”. In: *Proceedings of the International Society for Music Information Retrieval Conference, ISMIR*. Taipei, Taiwan: National Taiwan University, 2014.
- [135] A.A. Reblitz and Q.D. Bowers. *The Golden Age of Automatic Musical Instruments: Remarkable Music Machines and Their Stories*. Mechanical Music Press, 2001. ISBN: 9780970595102.
- [136] David Rector and Spencer Topel. “EMdrum: An Electromagnetically Actuated Drum”. In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. London, United Kingdom: Goldsmiths University of London, 2014.
- [137] EG Richardson. “Aeolian tones”. In: *Proceedings of the Physical Society of London* 36.1 (1923), p. 153.
- [138] Troy Rogers, Steven Kemper, and Scott Barton. “MARIE: Monochord-Aerophone Robotic Instrument Ensemble”. In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Baton Rouge, Louisiana: Louisiana State University, 2015.
- [139] M.M.I. Shakir, A.M. Shakir, H.M.I. Shakir, and D.R. Hill. *The Book of Ingenious Devices*. Springer, 1979. ISBN: 9789027708335.
- [140] Greg Shear. “The Electromagnetically Sustained Rhodes Piano”. Masters Thesis, University of California Santa Barbara. 2011.
- [141] Eric Sheffield and Michael Gurevich. “Distributed Mechanical Actuation of Percussion Instruments”. In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Baton Rouge, Louisiana: Louisiana State University, 2015.
- [142] K. Shibuya, Y. Chikaoka, T. Koyama, and S. Sugano. “The planning of violin playing robot with KANSEI information-algorithm to decide bowing parameters from timbre”. In: *6th IEEE International Workshop on Robot and Human Communication*. 1997, pp. 230–235.
- [143] Eric Singer, Kevin Larke, and David Bianciardi. “LEMUR GuitarBot: MIDI Robotic String Instrument”. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. National University of Singapore. 2003, pp. 188–191.
- [144] Eric Singer, Jeff Feddersen, Chad Redmon, and Bil Bowen. “LEMUR’s Musical Robots”. In: *Proceedings of the Conference on New Interfaces for Musical Expression, NIME*. Hamamatsu, Shizuoka, Japan: Shizuoka University of Art and Culture, 2004, pp. 181–184.

- [145] Neill Skelly. *Mechatronic Guitar Pitch Shifter*. Wellington, New Zealand: Technical Report, Victoria University of Wellington School of Engineering and Computer Science, 2016.
- [146] Julius O. Smith. *Spectral Audio Signal Processing*. URL: [http : / / ccrma.stanford.edu/~jos/sasp/](http://ccrma.stanford.edu/~jos/sasp/) (visited on 02/18/2017).
- [147] M. Smith. *Meet Z-Machines, Squarepusher's new robot band*. URL: [http : / / www . theguardian . com / music / 2014 / apr / 04 / squarepusher - z - machines - music - for - robots](http://www.theguardian.com/music/2014/apr/04/squarepusher-z-machines-music-for-robots) (visited on 07/27/2015).
- [148] Tarek M. Sobh, Bei Wang, and Kurt W. Coble. "Experimental Robot Musicians". In: *Journal of Intelligent and Robotic Systems* 38.2 (2003), pp. 197–212.
- [149] Jorge Solis and Kia Ng. *Musical robots and interactive multimodal systems*. Vol. 74. Springer, 2011.
- [150] Jorge Solis, Takeshi Ninomiya, Klaus Petersen, and Atsuo Takanishi. "Anthropomorphic Musical Performance Robots at Waseda University: Increasing Understanding of the Nature of Human Musical Interaction". In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Pittsburgh, Pennsylvania: Carnegie Mellon, 2009.
- [151] Jorge Solis, Koichi Taniguchi, Takeshi Ninomiya, Klaus Petersen, Teturo Yamamoto, and Atsuo Takanishi. "Implementation of an Auditory Feedback Control System on an Anthropomorphic Flutist Robot Inspired on the Performance of a Professional Flutist". In: *Advanced robotics* 23.14 (2009), pp. 1849–1871.
- [152] Jorge Solis, Keisuke Chida, Koichi Taniguchi, Shinichiro Michael Hashimoto, Kei Suefuji, and Atsuo Takanishi. "The Waseda Flutist Robot WF-4RII in Comparison with a Professional Flutist". In: *Computer Music Journal* 30.4 (2006), pp. 12–27.
- [153] Squarepusher. *Music for Robots*. EP. Warp Records. 2014.
- [154] Shigeki Sugano and Ichiro Kato. "WABOT-2: Autonomous Robot with Dexterous Finger-arm - Finger-arm Coordination Control in Keyboard Performance". In: *Proceedings of the International Conference on Robotics and Automation*. Vol. 4. IEEE. 1987, pp. 90–97.
- [155] S Takashima and T Miyawaki. "Control of an Automatic Performance Robot of Saxophone: Performance control using standard MIDI files". In: *Proceedings of the IEEE/RSJ International Conference on Intelligent*

- Robots and Systems Workshop: Musical Performance Robots and Its Applications*. 2006, pp. 30–35.
- [156] Adam Tindale, A. Kapur, and George Tzanetakis. “Training Surrogate Sensors in Musical Gesture Acquisition Systems”. In: *Transactions on Multimedia* 13.1 (2011), pp. 50–59.
- [157] Tonya Yirka. *Ancient Wind Chimes*. URL: [http : / / www . windchimesaustralia . com . au / history - of - wind - chimes . php](http://www.windchimesaustralia.com.au/history-of-wind-chimes.php) (visited on 07/16/2015).
- [158] Shawn Trail, Leonardo Jenkins, and Peter Driessen. “DSRmarimba: low-cost, open source actuated acoustic marimba framework”. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. 2014.
- [159] Shawn Trail, Lee Jenkins, Duncan MacConnell, George Tzanetakis, Ming Cheng, and Peter Driessen. “STARI: A Self Tuning Automonochord Robotic Instrument”. In: *Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*. IEEE. 2013, pp. 405–409.
- [160] George Tzanetakis and Perry Cook. “MARSYAS: a framework for audio analysis”. In: *Organised sound* 4.03 (2000), pp. 169–175.
- [161] George Tzanetakis, Andrey Ermolinskyi, and Perry Cook. “Pitch Histograms in Audio and Symbolic Music Information Retrieval”. In: *Journal of New Music Research* 32.2 (2003), pp. 143–152.
- [162] Jay Vance. *Captured! By Robots*. URL: [http : / / www . capturedbyrobots . com](http://www.capturedbyrobots.com) (visited on 08/17/2015).
- [163] Menno van der Veen, Francisco de Leon, Brian Gladstone, and Valeriu Tatu. “Measuring Acoustic Noise Emitted by Power Transformers”. In: *Proceedings of the Audio Engineering Society Convention 109*. 2000.
- [164] R. G. Vindriis, D. A. Carnegie, and A. Kapur. “A Comparison of Pick-Based Strategies for Robotic Bass Playing”. In: *Proceedings of the Engineering New Zealand Conference, ENZCon*. Palmerston North, New Zealand: Massey University, 2011.
- [165] Richard Vindriis and D. A. Carnegie. “StrumBot – An Overview of a Strumming Guitar Robot”. In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. Brisbane, Australia: Griffith University, 2016.
- [166] Gil Weinberg and Scott Driscoll. “The Design of a Robotic Marimba Player: Introducing Pitch into Robotic Musicianship”. In: *Proceedings*

- of the *International Conference on New Interfaces for Musical Expression*. NIME. 2007, pp. 228–233.
- [167] Gil Weinberg and Scott Driscoll. “Toward Robotic Musicianship”. In: *Computer Music Journal* 30.4 (2006), pp. 28–45.
- [168] Gil Weinberg, Scott Driscoll, and Mitchell Parry. “Haile - an Interactive Robot Percussionist”. In: *Proceedings of the International Computer Music Conference, ICMC*. Barcelona, Spain.: Escola Superior de Música de Catalunya, 2005.
- [169] Emil Welte. *Mechanical Musical Instrument*. US Patent No. 287.599.
- [170] J. Whiting. *Yo-Yo Ma: A Biography*. Greenwood biographies. Greenwood Press, 2008. ISBN: 9780313344862.
- [171] *Wind Chimes Australia: History of Wind Chimes*. URL: <http://www.windchimesaustralia.com.au/history-of-wind-chimes.php> (visited on 07/16/2015).
- [172] Xiao Xiao. “MirrorFugue III: Conjuring the Recorded Pianist”. In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. London, United Kingdom: Goldsmiths University of London, 2014.
- [173] Yamaha. *Yamaha History*. URL: http://www.yamaha.com/about_yamaha/corporate/history/index.html (visited on 07/26/2015).
- [174] Diana Young. “A Methodology for Investigation of Bowed String Performance Through Measurement of Violin Bowing Technique”. PhD Thesis, Massachusetts Institute of Technology. 2007.
- [175] M. H. Zareei, D. A. Carnegie, and A. Kapur. “Rasper: a Mechatronic Noise-intoner”. In: *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME*. London, United Kingdom: Goldsmiths University of London, 2014.