Automation, Collaboration, & E-Services

ACES

Gil Weinberg
Mason Bretan
Guy Hoffman
Scott Driscoll

Robotic Musicianship

Embodied Artificial Creativity and Mechatronic Musical Expression

Springer

Automation, Collaboration, & E-Services

Volume 8

Series Editor

Shimon Y. Nof, PRISM Center, Grissom, Purdue University,
West Lafayette Indiana, IN, USA

The Automation, Collaboration, & E-Services series (ACES) publishes new developments and advances in the fields of Automation, collaboration and e-services; rapidly and informally but with a high quality. It captures the scientific and engineering theories and techniques addressing challenges of the megatrends of automation, and collaboration. These trends, defining the scope of the ACES Series, are evident with wireless communication, Internetworking, multi-agent systems, sensor networks, cyber-physical collaborative systems, interactive-collaborative devices, and social robotics—all enabled by collaborative e-Services. Within the scope of the series are monographs, lecture notes, selected contributions from specialized conferences and workshops.

More information about this series at http://www.springer.com/series/8393

Gil Weinberg · Mason Bretan ·
Guy Hoffman · Scott Driscoll

Robotic Musicianship

Embodied Artificial Creativity and Mechatronic Musical Expression

Springer

Gil Weinberg
Georgia Institute of Technology
Atlanta, GA, USA

Guy Hoffman
Cornell University
Ithaca, NY, USA

Mason Bretan
Novato, CA, USA

Scott Driscoll
Atlanta, GA, USA

ISSN 2193-472X ISSN 2193-4738 (electronic)
Automation, Collaboration, & E-Services
ISBN 978-3-030-38929-1 ISBN 978-3-030-38930-7 (eBook)
https://doi.org/10.1007/978-3-030-38930-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To all of our collaborators who made this work possible.

Foreword

Robotic Musicianship does, as the authors point out, sound like a contradiction in terms. “Robotic” is one of the most dismissive descriptions imaginable of a human musician’s playing style. A “robotic” performance is one that is considered inexpressive, that incorporates none of the microinflections of loudness, timbre, and pitch that make a performance, as it is said, “come to life.” Our human idioms give away the game here as to what our expectations for music performance are. Beyond that, the way we talk about them is not very respectful toward robots.

The fact of the matter is that robots, for the most part, do what we tell them to do. Or, as is more commonly the case in contemporary research, they do what we have helped them learn how to do. If they do not perform expressively, it is because we have not understood musical expression sufficiently to make its real-time implementation computationally tractable. Or, that we have not collected and represented data in such a way that we can provide meaningful targets from which a computer (the brains of a robot, after all) could learn expression.

There could be no better team to blast through the facile misconceptions attaching to *Robotic Musicianship* than the authors of the present volume. Gil Weinberg, Guy Hoffman, Mason Bretan, Scott Driscol and their group at the Georgia Institute of Technology have been the recognized leaders in building music robots for decades now. Their research program has been squarely directed at designing, implementing, and demonstrating in high-profile venues the possibilities of robotic musicianship, with results that are undeniably compelling.

The visual component of their work is critical: others have worked on the design of interactive music systems that react in a convincing way to a real-time musical context (this is hard enough), but pure through an audio output. We know, however, that much of human musical interaction is guided by musicians watching one another, particularly during improvisation. Moreover, seeing a robot perform the output of these systems shows an audience exactly what the computational system is adding to the whole—something that can be difficult to parse out from audio alone, particularly in the context of ensemble playing.

Chapters in this text demonstrate the human-centric and musically motivated strategies of their agenda: they aim to help the robots “listen like a human”, but “play like a machine”. To “listen like a human” requires a sophisticated understanding of music cognition, and the traditions and research that have developed in that field over a period of decades. To “play like a machine” requires, indeed, that a machine be built. And then that the machine be able to execute the (expressively nuanced) musical gestures that will provide a meaningful counterpart to live human musicians.

That is a tall order of business. We see emulation of human cognition, musical insights sufficient to compose compelling and appropriate additions to an ongoing musical texture, and the mechanical engineering necessary to complete all of this in real time, in the context of a live public performance. Certainly, this is a challenge on the order of, say, building a self-driving car, or a robot that could navigate a busy sidewalk.

And that challenge is met here with acumen, rigor, insight, and strategic design. This is an engineering text. The authors, scientists, and engineers all, present this consummate feat of design and implementation here in a way that makes their advances available and reproducible by those following in their considerable wake. Much of what they present here has been developed with the benefit of input from human subjects, tests that provide important insights for the psychology of music. Signal processing, genetic algorithms, machine learning, robotics, algorithmic composition, wearables, music cognition and artificial intelligence are all involved, and all presented in a way that is practicable for the expert yet approachable by the layman.

Robotic performance is not only a matter of emulating what a human might do—indeed, a machine can perform material that would be quite impossible for a human musician. The authors demonstrate how their creations can not only render music expressively, but can expressively perform passages that could not be produced by their human counterparts. This is but one of the many ways in which their work is an extension and elaboration and contribution to human musical expression, rather than an effort of substitution. Robotic musicianship is not about replacing human musicians, but a celebration and riotous new expansion of the possibilities of music-making. These musician/engineer authors have spent decades building at the cutting edge of interactive music, performed by robots, in real time, before a live audience. What they have built is no gee-whiz science-fair project, but substantial engineering and real contribution to the possibilities of music-making. We will all be richer for it. Seeing (and hearing) is believing—just watch, and listen to them.

New York City, NY Robert Rowe

Preface

We arrived at Robotic Musicianship from different directions. Gil is an improvising musician, interested in bringing his algorithmic improvisation ideas into the acoustic world; Scott is a mechanical engineer looking to bring his technical and musical interests together; Guy is a Human Robot Interaction (HRI) researcher who finds music a rich domain to explore social human-robot interaction; and Mason is a machine learning scientist and a musician whose research combines both disciplines. This diversity in backgrounds and interests exemplifies the wide interdisciplinary nature of the field of Robotic Musicianship. Since 2005 we have brought together this wide range of perspectives, collaborating among ourselves and others at the Georgia Tech Center for Music Technology, to develop the robotic musicians described in this book.

The first robot, Haile, was the product of a collaboration between Gil, a new assistant professor of music technology at Georgia Tech, and Scott, a mechanical engineering master's student. The unfunded project led to a hand-built wooden robot, probably one of the few wooden robots out there, with exposed electronics and a rugged look. A video posted online of Scott improvising with Haile was noticed by an NSF director, who emailed to request a proposal for developing a more advanced robot that would help public perception of robots through music. The proposal that ensued led to the development of Shimon, a marimba playing robot, which helped facilitate new research directions into melodic and harmonic robotic improvisation.

A coincidental meeting between Gil and Guy at a conference led to the idea of a social head for Shimon. Guy then joined the Robotic Musicianship group at GTCMT as a postdoctoral researcher to develop gestural HRI applications for Shimon. After Guy was invited to present his work at a conference in Germany, a decision was made to create a small and portable version of Shimon for future conferences. The new personal robot, called Shimi (aka Travis), was developed in a record speed of six weeks to be ready for future demonstrations. It was then selected as one of the few finalists at the prestigious TechCrunch Disrupt startup competition, although the commercialization effort that followed was unsuccessful.

When Mason joined the Robotic Musicianship group as a Masters and later a Ph. D. student, he helped bring Shimon and Shimi to the next level by developing machine learning based improvisation and emotional gestures for the robots. Mason was also the main researcher behind the development of our wearable robotic musicianship platforms, following a request by amputee drummer Jason Barnes. After losing his arm in an electrical accident, Jason asked us to build a musical robotic prosthetic that would allow him to play drums again. Using EMG, ultrasound, and machine learning, we developed different prosthetic arms for Jason that not only brought back his lost abilities but also allowed him to play like no other human can. Jason is now featured as the “Fastest Drummer (using Prosthetic)” in the 2020 Guinness World Book of Records.

In addition to pursuing our respective research interests, we have also been drawn to Robotic Musicianship by the exciting performance opportunities the field offers. With our robotic musicianship platforms we have traveled across the world, performing dozens of concerts across four continents. Some of our performances and presentations were commissioned by prestigious venues such as The Kennedy Center for the Arts, Aspen Ideas Festival and the World Economic Forum in Davos. Others took place in festivals in cities such as St. Petersburg, Brisbane, Istanbul, and Shanghai. We also performed in small venues in front of small audiences such as in an art gallery in Paris, a private event in Berlin and a nightclub in Atlanta. A concert actually led to this book, when we met our Editor, Shimon Nof, who asked us to write this book after a concert he attended at Purdue University.

Traveling and presenting newly developed prototypes that have just come out of the lab can also lead to memorable, yet anxiety ensuing experiences. Today we can look back fondly at stressful experiences such as a performance in Paris, where Haile stopped playing due to a broken wire, leading Gil and Scott to continue improvising as a human duo while calculating escape routes out of the venue; or the time when a few hours before a live presentation on The TODAY show, Shimon stopped working, sending Gil to the streets of New York in a frenzied search for a replacement USB hub, with Mason frantically debugging Shimon in the studio (the live demonstration went well at the end). Guy’s nervous performance in at an Atlanta Concert Venue started before the audiences entered the hall. A different performance in a prestigious festival in Munich ended with Guy not being able to return to Georgia Tech for 3 months due to visa issues.

These concerts and performances are an important part of our Robotic Musicianship work and we encourage readers to look online for videos that could help demonstrate the musical outcome of the work described in the book. We hope readers will find both the technical and artistic aspects of our work as interesting and inspiring as we do.

Sincerely,

Atlanta, GA, USA — Gil Weinberg
Atlanta, GA, USA — Scott Driscoll
Ithaca, NY, USA — Guy Hoffman
Novato, CA, USA — Mason Bretan

Acknowledgements We would like to thank our collaborators: Roberto Aimi, Ryan Nikoloaidis, Mark Godfrey, Alex Rae, Brian Blosser, Travis Thatcher, Mitch Parry, Sisi Sun, Trishul Malikarjuna, Annie Zhang, Jay Clark, Marcelo Cicconnet, Philip Mullins, Deepak Gopinath, Roozbeh Khodambashi, Tyler White, Lea Ikkache, Zach Kondak, and Richard Savery.

Contents

Chapter 1
Introduction

1.1 Abstract

The purpose of this book is to introduce the concepts that define robotic musicianship and the underlying technology that supports them. Research in robotic musicianship focuses on the construction of machines that can produce sound, analyze and generate musical input, and interact with humans in a musically meaningful manner. In this chapter, we first discuss motivations for pursuing robotic musicianship, and review past and ongoing research efforts. In subsequent chapters we dive more deeply into these topics and describe specific designs, algorithms, and evaluations that we have developed and applied to our own robotic musicians.

1.2 Why Robotic Musicianship

The term '*robotic musicianship*' may seem like a contradiction-in-terms. The word 'robotic' often bears negative connotations in art, typically referring to a lack of expression. Conversely, 'musicianship' is usually used to describe varying levels of an individual's ability to apply musical skills and knowledge in order to convey artistry, creativity, and expression beyond the facets of merely reading notes from a score. Together these two words describe a growing field of research that merges science, art, and technology.

There are two primary research areas that constitute the field of Robotic Musicianship: (1) *Musical Mechatronics* —the study and construction of physical devices that generate sound through mechanical means [1] (2) *Machine musicianship*—the development of algorithms and cognitive models of music perception, composition, performance, and theory [2]. Robotic musicianship brings these two disciplines together. Researchers within this space design musical robots that have the underlying musical intelligence to support both expressive solo performance and interaction with human musicians [3].

G. Weinberg et al., *Robotic Musicianship*, Automation, Collaboration, & E-Services 8, https://doi.org/10.1007/978-3-030-38930-7_1

Research in robotic musicianship has seen growing interest from scientists, musicians, engineers and artists in recent years. This can be attributed to the interdisciplinary nature of the field that is amenable to a wide range of scientific and artistic endeavors. First, music provides a common platform for research because of its cross-cultural elements such as rhythm and pitch. Second, robotic musicianship requires the integration of multiple disciplines including engineering, computation, music, and psychology, bringing together a wide range of scholars to address the challenges involved in making a machine artistically expressive. Third, different aspects of research in robotic musicianship, such as timing, anticipation, expression, mechanical dexterity, and social interaction have the potential to lead to broader impacts in other fields from cognitive science, human-robot interaction (HRI) to mechanics. Moreover, music is one of the most time-sensitive mediums where events separated by mere milliseconds are perceptually noticeable. Developing HRI algorithms for anticipation or synchronization in a musical context can, therefore, have impact on other HRI scenarios, where accurate timing is required. By building and designing robotic musicians, scholars can also better understand the sophisticated interactions between the cognitive and physical processes in human music making. By studying and reconstructing sound production mechanisms, robotic musicians can help shed light on the role of embodiment in music making, exploring how music perception and cognition interact with human anatomy.

Other motivations to develop robotic musicians are cultural and artistic in nature. Robert Rowe, a pioneer of the field of interactive music systems, writes, "if computers interact with people in a musically meaningful way, that experience will bolster and extend the musicianship already fostered by traditional forms of music education…and expand human musical culture" [2]. Similarly, the goal of robotic musicianship is to supplement human creativity and enrich the musical experience for humans, rather than imitate or replace it. By bringing together perceptual and generative computation with physical sound generators we can create systems capable of (1) rich, acoustic sound production, (2) intuitive, physics-based visual cues from sound producing movements, and (3) expressive physical behaviors through sound accompanying body movements. Moreover, there is artistic potential in the non-human characteristics of music-generating machines including increased precision, freedom of physical design, and the ability to perform humanly impossible fast computations. Discovering how best to leverage these characteristics to benefit music performance, composition, consumption, and education is a worthy effort as music is an essential component of human society and culture.

As social robots begin to penetrate public spaces, music can play an important role in aiding the process of integrating personal robotics into everyday life. Music is shared, understood, and appreciated by billions of people across many different cultures. Such ubiquity can be leveraged to yield a sense of familiarity between robots and people by using music to encourage and facilitate human-robotic interaction. As more people begin to integrate robots into their everyday lives, music can help establish trust and confidence in this new technology.

But the recent interest in robotic musicianship has not come without popular criticism and misconceptions, some of which focus on the concern that the ultimate outcome of this research might be to replace human musicians.[1,2,3] Though much of this concern may stem from the natural awe and unease accompanied by the introduction of new technology, it is important to reiterate both our own and Rowe's sentiment that by placing humans alongside artistically capable machines we are establishing an environment in which human creativity can grow, thus, enriching human musical culture rather than replacing it.

A related criticism stems from the belief that the creation of music—one of the most unique, expressive, and creative human activities—could never be replicated by automated "soul-less" machines, and that the product of such efforts will never lead to "good music".[4] We address this potential criticism by emphasizing its great potential for enhancing human musical creativity and expression. For example, although in some cases the need for live musicians diminished when new technology was introduced (such as the introduction of talking films, or the MIDI protocol), technology has also introduced new opportunities in sound and music making, such as with sound design and recording arts. Similarly, disk jockeys introduced new methods for expressive musical performance using turn tables and pre-recorded sounds. Further, robotic musicianship has the potential to advance the art form of music by creating novel musical experiences that would encourage humans to create, perform, and think about music in new ways. From utilizing compositional and improvisational algorithms that humans cannot process in a timely manner (using processes such as fractals or cellular automata) to exploring physical sound production capabilities that humans do not possess (from speed to timbre control), robotic musicians bear the promise of creating music that humans could never create by themselves and inspire humans to explore new musical experiences, invent new genres, expand virtuosity, and bring musical expression and creativity to uncharted domains.

The rest of this chapter will review past and ongoing research in robotic musicianship, starting with exploring different methods for sound production, moving to machine listening, and then to generative algorithms that enable the robot to complete different musical tasks such as improvisation or accompaniment. The chapter concludes with the design of social interactions and visual cues based on embodiment, anthropomorphism, and expressive gestures.

1.3 Sound Production and Design—Survey

Acoustic musical instruments produce sound in a variety of ways. For example, resonating chambers can be excited by bowing strings, striking a covering membrane, and blowing air. When designing autonomous sound generators it is important to

---

[1] http://www.asylum.com/2011/01/19/robot-rock-bands-expressive-machines-musical-instruments/.

[2] http://nymag.com/news/intelligencer/robot-jobs-2013-6/.

[3] http://www.wired.com/2012/12/ff-robots-are-already-replacing-us/2/.

[4] http://www.smithsonianmag.com/history/musicians-wage-war-against-evil-robots-92702721/?no-ist.

consider the acoustic properties of the resonator. Robots that play traditional acoustic instruments can be classified by those which emulate the human body and natural mechanics applied to playing an instrument and those which augment acoustic instruments to produce sound and create a physical affordance that is not possible in the instrument's archetypal state, such as in the case of prepared piano [1].

1.3.1 Traditional Instruments

Traditional acoustic instruments are designed to be played by humans in terms of their size, ergonomic design and input affordances that accommodate the human body. Robots, on the other hand, are not constrained to human form and can, thus, be designed to play musical instruments in a different manner than humans do. While emulating people can be mechanically challenging, some musical robot designers find it helpful to reference and build on human physiology and the various physical playing techniques employed by musicians. Other designers and researchers attempt to break free from the constraints of human form and explore aspects of instrument playability related to speed, localization, and dynamic variations that are not attainable by humans. In this section we will explore examples of different design methodology for different instruments and discuss their inherent advantages and disadvantages.

1.3.1.1 Percussive Instruments

The mechanics involved in human percussion playing are complex and require controlled movement of and coordination between the wrist, elbow, and fingers, and with some instruments, the legs and feet. In the case of hand drumming, subtle changes in the tightness of a hand grip on the drumstick, rate at which the wrist moves up and down, and distance between the stick and drum head provide drummers with expressive control over their performance. Human drummers use different kinds of strokes to achieve a variety of timbres and playing speeds. For example, for creating two distinct hits separated by a large temporal interval, drummers usually use a single-stroke in which the wrist moves up and down for each strike. For a double hit separated by a very small temporal interval, drummers tend to use a multiple-bounce stroke in which the wrist moves up and down once to generate the first strike, while subsequent strikes are achieved by adjusting the fingers' grip on the drumstick to generate a bounce. Multiple motors, gears, solenoids, and a sophisticated control mechanism are required to imitate such functionality.

Several robotic systems have been designed to imitate velocity curves of single-stroke wrist and arm movements. The *Cog* robot from MIT uses oscillators for smoother rhythmic arm control [4, 5]. Another approach to smooth motion control is to use hydraulics as done by Mitsuo Kawato in his humanoid drummer [6]. Some robotic designs are inspired by the general physical shape of humans rather than specific human muscle and joint control. The humanoid keyboard playing *WABOT* has two hands for pressing keys and feet for pushing pedals [7]. More recently, Yoky Matsuoka developed an anatomically correct piano-playing hand [8].

Designers of robotic percussion instruments can also achieve their sonic and expressive goals by using simpler engineering methods than those that are required to replicate human physical control. A common approach for such percussive robotic devices is to use solenoids that strike percussion instrument with sticks or other actuators. Erik Singer's LEMUR bots [9] and portions of *Orchestrion* (designed for Pat Metheny[5]) use such an approach. Singer's projects use non-anthropomorphic designs in which individual solenoids are used to strike individual drums.

In addition to a simplified actuation design, such non-humanoid robots can produce musical results that are not humanly possible. For example, in Orchestrion each note on a marimba can be played faster than human speed because a solenoid is placed on each one of the keys. This approach allows for increased flexibility in choosing different types of push and pull solenoids that can be used to achieve high speed and dynamic variability [10]. Unlike human percussionists, the single strokes generated by solenoids or motors can move at such great speeds so that it is not necessary to model human double-stroke functionality.

While solenoid based designs are mechanically simple and capable of a wide range of expression, their small range of movement cannot provide useful visual cues to humans. Humanoid robots, on the other hand, can provide human co-players and audience members with a framework that they are more familiar with, thus, allowing them to better understand the relationship between sound generation and the accompanying robotic movements and gestures. Additionally, the visual cues provided by the larger human-size motion and gestural movements allow the human musicians to anticipate the robot's actions and improve synchronization. Some robots combine functions of both design methods in an attempt to simultaneously gain the benefits of the inhuman sound producing capabilities and important visual cues. The marimba playing robot, *Shimon* (Fig. 1.1), has four arms each with two solenoid activators for striking the keys. The movements of the arms combined with the eight fast moving solenoids provides advantages from both design methods [11]. Additional percussive systems are listed in Table 1.1.

1.3.1.2 Stringed Instruments

Similarly to percussive robots, many string playing robots focus on achieving simple, yet robust excitation and pitch manipulation rather than mimetic control and anthropomorphism. Sergi Jorda's *Afasia* system [17] is one such example, where solenoids are used to pluck strings, while pitch is controlled by either push solenoids placed along the bridge or motors that move the bridge itself. Singer's *GuitarBot*, on the other hand, uses a wheel embedded with guitar picks that can rotate at variable

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

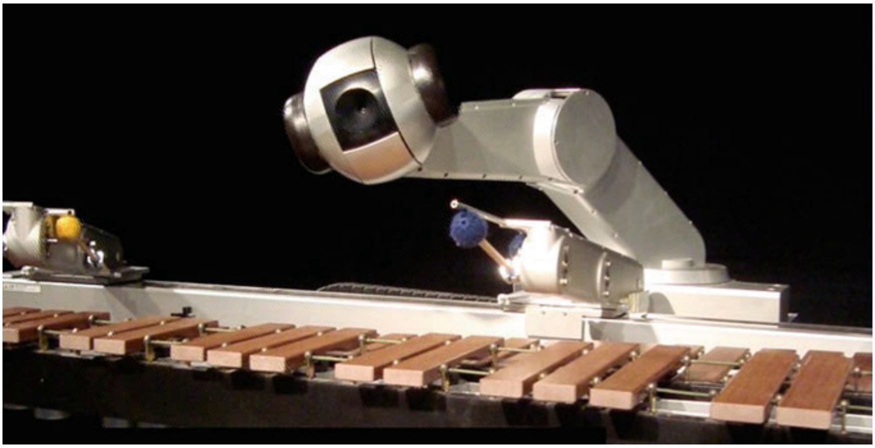

Fig. 1.1 *Shimon* of Georgia Tech Center for music technology

Table 1.1 Percussive robots

Developer	Robotic musician
Logo's Foundation	**Vibi**—An automated vibraphone with individual dampers for each bar
	Troms—Seven single drums each outfitted with a set of strikers positioned at various locations on the drumhead
	Vacca—Series of cowbells equipped with different style hammers driven by solenoids [12, 13]
Expressive Machines Musical Instruments (EMMI)	**MADI**—*Multi-mallet automatic drumming instrument* consists of 15 solenoid strikers positioned around a single snare drum allowing the system to take advantage of different timbres available at specific locations on the drum [14]
Trimpin	**Automated idiophones**—Pitched percussion outfitted with individual solenoids that drive strikers that can be interchanged and adapted to achieve different timbres [15]
Karmetik	**Notomotion**—A drum equipped with several rotary and pull solenoids with a rotating mounting structure allowing for a wide range of timbres
	Raina—A rainstick attached to a drive train that slowly rotates [16]
Georgia Tech Center For Music Technology (GTCMT)	**Haile**—Anthropomorphized robotic percussionist equipped with a linear motor and solenoid for striking a drum [3]

Table 1.2 String playing robots

Developer	Robotic musician
Logo's Foundation	**Hurdy**—An automated bowed bass instrument that uses several motors to bow the strings
	Aeio—An automated cello in which the strings are excited by using two electromagnets driven by a two phase signal on opposite sides of the string [12, 13]
EMMI	**PAM**—PAM: Poly-tangent Automatic multi-Monochord [14]
Trimpin	**Krantkontrol**—An installation of 12 guitar-like instruments with a plucking mechanism and solenoids for fretting
	"If VI was IX"—An installation containing hundreds of guitars with pitch and plucking acuators [15]
Baginsky	**Aglaopheme**—A slide guitar with solenoids for plucking and fretting and a motor used to alter the position of the bridge [18]
Victoria University of Wellington and New Zealand School of Music	**MechBass**—A four string modular robotic bass guitar player. Each string is plucked using a stepper motor and a solenoid attached to a carriage moves along the string via a belt drive for fretting [19]

speeds to pluck strings and relies on a moving bridge assembly that rides on pulleys driven by a DC motor.

A different, more anthropomorphic approach is taken by Chadefaux et al., which utilizes motion capture of human harp playing to design of a robotic harp playing finger. Silicone in various shape and stiffness is placed on the robotic fingertips in order to replicate different vibrational characteristics [20]. An anthropomorphic violin bowing arm was developed by Shibuya et al. The shoulder, elbow, and wrist were modeled to create expressive playing by controlling the bow's pressure, position, and rate of movement [21, 21, 22]. Similarly, Toyota's robotic violin player replicates human form and motion.[6] See additional string playing robots in Table 1.2.

1.3.1.3 Wind Instruments

One of the main challenges for wind-instrument playing robots is to produce the necessary energy fluctuations that would make the instrument resonate at desired frequencies and volumes. Some designers address this challenge using non-humanoid designs and function. Roger Dannenberg's *McBlare* is a non-humanoid robotic bag-

[6] http://www.toyota-global.com/innovation/partner_robot/.

Table 1.3 Wind playing robots

Developer	Robotic musician
Logo's Foundation	**Ob**—An automated oboe in which the reed is replaced with a acoustic impedance converter that models a real reed in a human mouth cavity [12, 13]
National ICT Australia	**Robot Clarinet**—A robotic clarinet player that controls for blowing pressure, lip force, and lip position to affect pitch, sound level, and spectrum [23]

pipe player that is based on an air compressor that provides air and controls the power to robotic "fingers" which open and close the sound holes. While the speed of McBlare's control mechanisms surpasses that of expert bagpipe performers it cannot equal the subtle human skill of adapting finger movements to the characteristics of the reed [24]. Another example of a non-anthropomorphic robotic wind instrument is the Logo's Foundation's *Autosax* [12], which uses "compression drivers and acoustic impedance converters that feed the drive signal to the instrument via a capillary."

A different approach for wind instrument playing robots focuses on replicating and understanding human physiology and anatomy. Ferrand and Vergez designed a blowing machine that explores the nonlinear dynamics natural human airflow by modeling the human mouth [25]. Their system's performance was affected by the servo valve's friction, cut-off frequency, and mounting stiffness of the artificial lips. The Takanishi Group at Waseda University developed a humanoid flute-playing robot as well as a sax-playing robot [26, 27]. Their systems include multiple degree-of-freedom (DOF) actuators for controlling finger dexterity, lip and tongue control, and lung control using a valve controlled air pump. Additional wind playing robots are listed in Table 1.3.

1.3.2 Augmented and Novel Instruments

While the examples described in previous sections demonstrate how musical robots are designed to play traditional instruments based on how humans play those instruments, some robotic researchers have been exploring new ways to play traditional instruments. Such methods can be useful for acquiring additional sonic variety and playability. Trimpin's *Contraption Instant Prepared Piano* is an example for such instrument augmentation. It enables mechanical bowing, plucking, and vibrating of piano strings to increase the harmonic and timbral range of the instrument. In many augmented instruments the human performer and robotic mechanisms share the control of the sound producing characteristics. This can lead to new musical experiences that combine the traits inherent to the interaction of two individual musicians and the interaction between performer and instrument. Richard Logan-Greene's systems,

Table 1.4 Augmented and novel instruments

Developer	Robotic musician
Logo's Foundation	**Dripper**—A rain machine which controls the precise size and frequency of each drip [12, 13]
Augmented Instruments Lab (Queen Mary University)	**Magnetic Resonator Piano**—An acoustic grand piano augmented with electromagnets inside the instrument to create vibrations in the strings [28]
	Digital Bagpipe—An electronic bagpipe chanter interface that is equipped with infrared sensing to record finger positions for later analysis [29]
Karmetik	**ESitar**—A hyperinstrument that uses sensors to measure human gesture during sitar performance and interface directly with a computer [30]

Brainstem Cello and *Actionclarinet*,[7] can provide such shared interactive control as the instruments are outfitted with sensors that respond to the human performer dampening the strings or manipulating pitch in real-time.

There are also examples of utterly new instruments in which automated mechanical functions are necessary to create a performance. Trimpin's "Conloninpurple" installation is made up of marimba-like bars coupled with aluminum resonating tubes and solenoid triggers. Each key is dispersed throughout a room to create a massive instrument capable of "3-dimensional" sound. Logan-Greene's *Submersible* is a mechanical instrument that raises and lowers pipes in and out of water to modify their pitch and timbre (Fig. 1.2) [31]. Additional augmented instruments are listed in Table 1.4.

1.4 Musical Intelligence

The ability to produce sound is only one important aspect of robotic musicianship. Designing cognitive models and algorithms that enable machines to behave like skilled musicians is another priority. Intelligent musical machines must embody the traits necessary for applications encompassing performance, interaction with other musicians, and improvisation. The ultimate goal of robotic musicianship is to create robots that can demonstrate musicality, expression, and artistry while stimulating innovation and newfound creativity in other interacting musicians. In order to achieve artistic sophistication, a robotic musician must have the ability to extract information relevant to the music or performance and apply this information to its underlying musical decision processes.

[7] http://zownts.com/.

Fig. 1.2 *Submersible* by Richard Logan-Greene

1.4.1 Sensing and Perception

1.4.1.1 Machine Listening

Providing machines with the ability to sense and reason about their environments is an important objective of robotics research. For robotic musicians, the emphasis on sensing and analysis of environmental input focuses on the auditory domain, although visual sensing is important for musicianship as well. Extracting meaningful information from an audio signal is referred to as *machine listening* [2]. There are numerous features, in different levels of complexity and perceptual relevance, that can describe the characteristics of a musical signal. Lower-level features typically describe physical characteristics of the signal while higher-level descriptors, though often more difficult to compute, tend to correlate more strongly with human musical perception, including elements such as mood, emotion, and genre. These higher-level descriptors are usually classified by using a combination of low-level (amplitude and spectral centroid) and mid-level (tempo, chord, and timbre) features (see Table 1.5).

Determining which attributes of the sound should be used for robotic perception relies on the desired functionality of the robot and its interaction with human musicians. For example, when a robotic musician designer intends for the system to interact with a drummer, audio analysis techniques that focus on rhythmic and timing qualities would be preferred over an analysis of pitch and harmony. In fact, for many of the drumming robots described above, audio processing algorithms have indeed focused on aspects related to timing. A first step in rhythmic analysis is to detect musical events such as individual notes. This process is referred to as *onset detection* and can be implemented using a variety of methods based on the derivative

Table 1.5 Relevant 'machine listening' terms

Loudness—A low level feature that correlates with the human perception of a signal's amplitude. Different methods of measurement exist for addressing perceptions at different frequencies
Mel Frequency Cepstral Coefficients (MFCCs)—A compact representation of the spectral envelope that is computed from the cepstral representation of a signal using a mel-frequency scale
Pitch—A perceptual feature of sound that correlates with frequency. However, the relationship is non-linear in that a perceived interval (such as an octave) at higher frequencies is separated by a larger number of hertz than at lower frequencies
Root Mean Square (RMS)—A low level feature that is used to describe the magnitude of an audio signal
Spectral Centroid—A low level feature describing the center of gravity within the frequency domain
Spectral Flux—A low level feature describing the change in spectral shape and computed by calculating the average difference between consecutive spectral frames. It is often used for onset detection because peaks arise at pitch changes and the beginning of new notes
Timbre—A characteristic of sound that is often described as a sound's *color* or *texture*. It allows us to differentiate between two sound sources that have equal pitch and loudness. Many definitions of timbre exist, but each has been criticized by researchers and there is currently no consensus on any one particular definition. It is easier to describe timbre as what it is not: pitch and loudness. Timbre is a multi-dimensional property meaning (unlike pitch and loudness) it cannot be ordered on a scale from low to high. Typically, audio features that describe the spectral shape of a signal (such as centroid, flux, and MFCCs) are used to describe the timbre of a sound
Tonality—A predominant system used in Western music that is determined by the relationships among a song's *key*, *chords*, and *notes*. It provides context to harmonic structure by describing the specific functions of individual notes and chords
Tonalness—A signal property describing the amount of *tonal* and *noisy* components in a signal. Signals that exhibit high periodicity are expected to be perceived as tonal and those with inherently low periodicity (such as white noise) are perceived as noisy

of energy with respect to time, spectral flux, and phase deviation [32]. Instruments which exhibit longer transients (such as cello or flute) and rich playing techniques (such as vibrato and glissandi) can pose challenges for these forms of detection. It is also a subject of debate whether the system should detect all physical onsets or all humanly perceptible onsets. Some psychoacoustically informed methods support the natural human perceptions in different portions of the frequency spectrum [33, 34].

An analysis of onsets over time can provide additional musically meaningful information related to note density, repetition, and beat. The notion of beat, which can correspond to the rate at which humans tap their feet to the music, is characteristic to many musical genres. Beat is often computed by finding the fundamental periodicity of perceived onsets using autocorrelation [35]. The *Haile* robot, for example, continuously estimates the beat of the input rhythms and uses the information to inform its response [11]. Barton's *HARMI* (Human and Robot Musical Improvisation) software