MECHANICAL SOUND SYNTHESIS: AND THE NEW APPLICATION OF FORCE-FEEDBACK TELEOPERATION OF ACOUSTIC MUSICAL INSTRUMENTS

Edgar Berdahl and Julius O. Smith III
Center for Computer Research in
Music and Acoustics (CCRMA)
Stanford University, Stanford, CA, USA

Günter Niemeyer
Telerobotics Laboratory
Stanford University
Stanford, CA, USA

ABSTRACT

In Mechanical Sound Synthesis, real mechanical devices are employed to create sound. Users can interact directly with the variables of the sound synthesis, making interactions more intuitive to both users and audience. We focus on real-time feedback control for Mechanical Sound Synthesis and provide a classification scheme using the reality-virtuality continuum. We discover an apparently novel paradigm, which is described as augmented virtuality for real-time feedback control. Exploring this paradigm, we present preliminary results from a system enabling a user to teleoperate acoustic percussion instruments with the aid of force feedback. Mechanical looping of the teleoperation trajectories and their transformations enables the synthesis of lifelike sounds with superhuman characteristics that are nevertheless produced by mechanical devices.

1. INTRODUCTION

In 1963, Max Mathews wrote that any perceivable sound could be produced using digital sound synthesis on a computer [1]. One could argue that this, combined with the fact that computers can be programmed to execute so many different programs, created the computer music field, which develops further as new computer-related technologies advance. Consider a composer faced with the task of creating a new piece. A relevant question is, “How many pieces are possible?” If a piece is three minutes long, monophonic, and sampled with four bits at 20kHz, then about

$$2^4 \cdot (20000\,\text{samples/s})(60\,\text{s/min})(3\,\text{min}) \approx 6 \cdot 10^7 \qquad (1)$$

distinct pieces are possible. Of course, the majority of these pieces are perceptually equivalent and/or sound like noise, so one could argue that the number of perceptually distinct three-minute pieces is much smaller, but nevertheless, a truly astounding number of three-minute pieces are possible. For instance, listening day and night to the pieces described by (1) would require 329 person-years! Because so many pieces are possible, the mapping from the individual sound samples of the uncompressed audio representation to the perceived sound is extraordinarily complicated, so composers very rarely specify the bare sound samples themselves.

In fact, one could argue that the whole point of digital sound synthesis algorithms has been to reduce the space of possible pieces to make composing and performing more manageable [2]. We argue that Mechanical Sound Synthesis also reduces the space of possible pieces, and we believe that it does so in an intuitive manner due to its visceral, physical nature.

2. MECHANICAL SOUND SYNTHESIS

For the purpose of this paper, we provide a preliminary definition of Mechanical Sound Synthesis:

> *Any method for creating sound is an example of Mechanical Sound Synthesis if the device actuating the sound does not exclusively incorporate loudspeakers, unless the loudspeakers create sound in an unconventional manner.*

For example, a loudspeaker setting piano strings into motion is an example of Mechanical Sound Synthesis, whereas a loudspeaker driver coupled to a simple, large radiating board is not because it is actually just a composite loudspeaker that radiates sound according to the digital (or even analog) sound synthesis paradigm.

We believe that Mechanical Sound Synthesis has a number of advantages due to its physical nature:

- *The mechanical variables of the sound synthesis algorithm can be made directly tangible to users.* In other words some of the variables (i.e. states), which describe the sound during synthesis, are mechanical. Hence, these mechanical variables can be manipulated as if they were real objects. As a consequence, sonic interactions are physical in nature and can be especially intuitive to both the audience and users.
- *The synthesized sound is collocated with the action responsible for creating it. Depending on the implementation, the sound may even resemble the objects that create it.* Consequentially, it can be easier for the audience to understand what actions in a performance are associated with which sounds. Furthermore, in group performances, it can be easier for users to know who is responsible for which sounds.
- *In contexts with active mechanical elements such as robots, physical-seeming and lifelike sounds can result even from superhuman interactions.* Robots can perform actions with repeatability, accuracy, and speed that may be unparalleled by human users, creating physical and lifelike sounds that are nevertheless new and that were not possible before.

In the remainder of this section, we introduce performance models for Mechanical Sound Synthesis, which classify performance according to style of the mechanical force interaction between the user and the sound synthesis. Some of the models may already be familiar to the reader, and others may not be familiar. An outline of the models is given in Figure 1 to help provide some further orientation during the discussion.

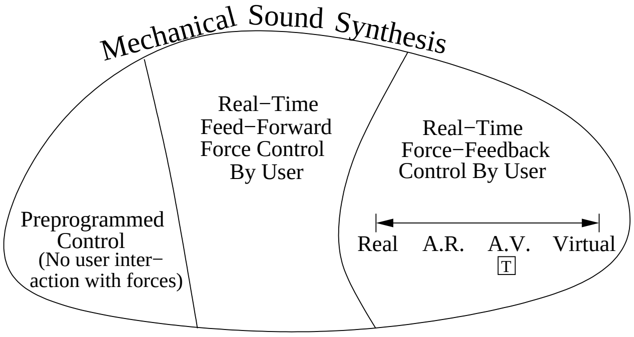

Figure 1: *Diagram representing suggested models for Mechanical Sound Synthesis.*

2.1. Preprogrammed Control (No user interaction with forces)

In the simplest model, there is no user, or there is no mechanical interaction between the user and the sound synthesis. Instead, mechnical devices such as robots may be pre-programmed to create sound by exerting forces on sound objects in the real world. For example, robots may play acoustic musical instruments according to a predetermined score, they may play semi-randomly, or they may improvise depending on what other users are playing. There is an established history of robotic musical instruments, so we will not discuss them further here [5].

2.2. Real-Time Feed-Forward Force Control By User

Mechanical Sound Synthesis can provide a user with an interface enabling him or her to control real sound objects in real-time. This model is known as "feed-forward control" when the user does not receive any force feedback. Recent examples include the Telematic Drum Circle, which can be played remotely over the Internet [6], and Pat Metheny's Orchestrion, in which robotic musical instruments are triggered by the MIDI output from an otherwise standard electric guitar [7].

2.3. Real-Time Force-Feedback Control By User

We finally consider the model representing the first author's main area of research, in which the user receives force feedback and has fine mechanical control over sound objects for real-time performance. We introduce the reality-virtuality continuum [8] to help classify subcategories of this model, as illustrated along the arrow within the *Real-Time Force-Feedback Control By User* section of Figure 1 (right) and exploded in Figure 2.

2.3.1. Real Environment

The real environment lies at the left-hand end of the continuum (see Figure 2, left). It describes the paradigm in which the user interacts with passive sound objects, such as vibrating strings, percussion instruments, hand claps, etc., which do not possess energy sources. Therefore, the objects are passive, so they can be traditional acoustic musical instruments.

2.3.2. Virtual Environment

The right-hand end of the reality-virtuality continuum (see Figure 2, right) describes the paradigm in which the user interacts with virtual sound objects and is completely immersed in a virtual environment [4]. For high fidelity musical interaction, the user is provided with visual, auditory, and force feedback as shown in Figure 3. The Association pour la Création et la Recherche sur les Outils d'Expression pioneered the earliest research in immersive force-feedback interaction with virtual sounding objects [9].

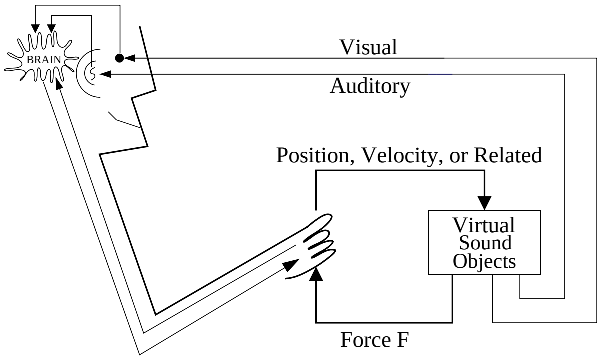

Figure 3: *User interacting with virtual sound objects with auditory, visual, and force feedback.*

2.3.3. Augmented Reality

Other paradigms implied by the reality-virtuality continuum (see Figure 2) have implications for Mechanical Sound Synthesis. For example, consider *augmented reality (A.R. in Figure 1, right)* for graphical user interfaces, in which an image of the real environment is augmented with virtual labels or highlights for providing additional information [8]. Analogously, an augmented reality audio headset has been designed that resembles a stereophonic hearing aid that superimposes audio from the local, real environment with audio from virtual environments, such as audio from remote people or informational services [10].

In the context of Mechanical Sound Synthesis, we argue that the augmented reality paradigm must reflect an environment containing real sounding objects, whose behavior is augmented by virtual objects. For instance, a user can manipulate a traditional guitar string, while a virtual bow employs feedback control to electromagnetically sustain the real string's vibration [11]. Similarly, a user can interact with a xylophone bar while its dynamics are changed by a virtual spring, mass, or damper implemented by feedback control [12]. The virtual environment affects the user's interactions, but the user interacts only *in*directly with the virtual objects.

2.3.4. Augmented Virtuality

Consider also *augmented virtuality (A.V.)*, where a virtual environment is augmented with elements or avatars from the real environment. This paradigm lies closer to but not at the virtual end of the reality-virtuality continuum as depicted in Figure 2 (right). For the force-feedback model, this paradigm implies force-feedback interaction with a virtual world, where some elements correspond to avatars of real objects. These avatars could represent other users or real sounding objects, such as acoustic musical instruments. We believe that this paradigm is in fact new, which is why we study it further in the remainder of this paper.

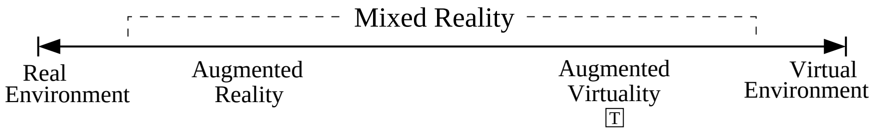

Figure 2: *Reality-Virtuality Continuum.*

3. AN EXAMPLE OF AUGMENTED VIRTUALITY INTERACTION FOR MECHANICAL SOUND SYNTHESIS: FORCE-FEEDBACK TELEOPERATION OF MUSICAL INSTRUMENTS

3.1. Introduction

Consider a user interacting with a virtual environment that incorporates force-feedback. For instance, the user could be manipulating a robotic arm that provides force-feedback. Imagine that by engaging in different parts of the virtual environment, the user could meet and interact with avatars of real musical instruments. Because the avatars would correspond to real sounding objects, they would effectively augment the virtual environment in which the user would be immersed. Teleoperation is an example of this, as indicated by the boxed *T* in Figures 1 and 2.

3.2. Teleoperation

In a typical application of teleoperation, the user manipulates a *master* robotic arm, which is linked by way of feedback control to a *slave* robotic arm. The user controls the slave arm remotely to complete a task. For instance, the user could control the slave to perform minimally invasive surgery inside a patient or to carry out work in a dangerous environment, such as inside a nuclear reactor. For instance, let $\mathbf{x_m}$ and $\mathbf{x_s}$ be the vector positions of the ends of the master and slave robotic arms, respectively. Then the time derivatives $\dot{\mathbf{x}}_\mathbf{m}$ and $\dot{\mathbf{x}}_\mathbf{s}$ represent the vector velocities of the master and slave arms, respectively. A virtual spring with stiffness k and a virtual damper with damping constant R are typically employed to link together the master and slave. In other words, the vector force on the slave robotic arm

$$\mathbf{F_s} = k(\mathbf{x_m} - \mathbf{x_s}) + R(\dot{\mathbf{x}}_\mathbf{m} - \dot{\mathbf{x}}_\mathbf{s}). \tag{2}$$

To implement Newton’s third law, the vector force on the master robotic arm

$$\mathbf{F_m} = -\mathbf{F_s}. \tag{3}$$

3.3. Using The NovInt Falcon

In our laboratory, we chose to implement teleoperation of musical instruments using NovInt Falcon robotic arms due to their attractive price of approximately US $150/each. The user could operate the master robot simply by moving the Falcon grip as shown in Figure 4 (right). Percussion instruments and drumsticks were attached to the ends of of the slave robotic arms, which were interchangeable. For instance, Figure 4 (left) shows a slave playing a shaker, while the left and right sides of Figure 5 show slaves playing a snare drum and a tambourine, respectively. An online video demonstration at
`http://www.youtube.com/watch?v=TfqG5_OanLo`
demonstrates the quality of the teleoperation [13].

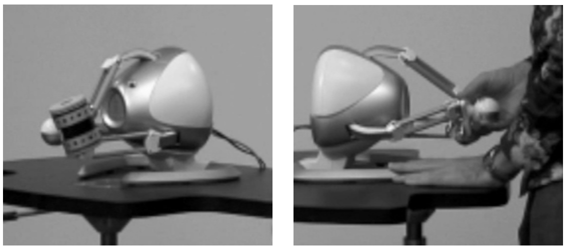

Figure 4: *Shaker attached to slave robotic arm (left) and user operating the master robotic arm (right).*

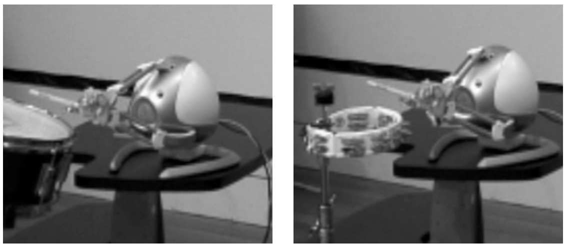

Figure 5: *Slave playing a snare drum (left) and slave playing a tambourine (right).*

3.4. Force Feedback Makes Teleoperation Easier

It is generally known that force feedback makes it easier for a human operator to control a slave. We informally observed that this was particularly true for teleoperation of the güiro musical instrument. Figure 6 shows how the slave robot was interfaced to the güiro with a drumstick. In the following video demonstration, the user first teleoperates the güiro with the help of force feedback. Then force feedback is disabled, and the user is not able to play the rhythm as accurately because it is difficult for him or her to judge whether the drumstick is pressing against the güiro at the correct force level to facilitate the scraping dynamic characteristic of güiro playing [13].
`http://www.youtube.com/watch?v=_BWWfEk3alE`

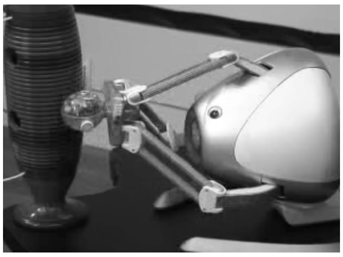

Figure 6: *Slave robot interfaced to a güiro musical instrument with a drumstick.*

3.5. Mechanical Looping

One intriguing question is, “How can one enable a master robot to control distinct parts on multiple slaves simultaneously?” One could employ artificial intelligence algorithms to cause the slave robots to behave more autonomously, but we are currently more interested in intimate control of the slave robots. We believe that this is key in realizing the advantages of Mechanical Sound Synthesis, as outlined in Section 2.

One solution involves “teaching” each slave to play its part one at a time using force-feedback teleoperation. For example, while teaching the ith slave, the master trajectory $(\mathbf{x_{m,i}(t)}, \dot{\mathbf{x}}_{\mathbf{m,i}}\mathbf{(t)})$ is recorded into a loop buffer. Then later when desired, the ith slave can be caused to play the part again by reading $(\mathbf{x_{m,i}(t)}, \dot{\mathbf{x}}_{\mathbf{m,i}}\mathbf{(t)})$ out of the buffer and applying (2) to control the ith slave. Computer music techniques for manipulating wave tables can be employed to warp and otherwise transform the trajectories [2]. For instance, by speeding up the wave table playback, the slave robots can be caused to produce lifelike sounds even when moving at superhuman speeds.

3.6. Live Performance

To demonstrate the advantages of force-feedback teleoperation of musical instruments as well as Mechanical Sound Synthesis in general, we prepared a performance piece. The user manipulated a single master robot (see Figure 7) to teleoperate and teach mechanical loops to four slave robots. The piece *Edgar Robothands: When The Robots Get Loose* was performed at the Center for Computer Research in Music and Acoustics (CCRMA) at the CCRMA Winter Concert on February 18, 2010. Besides the aforementioned robots, it incorporated some traditional percussion instruments as well as the haptic drum, a drum pad that triggers Indian percussion sounds and that provides force feedback to enable the user to play superhuman drum rolls [11]. An audio rendition of the piece is available at the following URL:

http://ccrma.stanford.edu/~eberdahl/CompMusic/RL.mp3

Figure 7: *Still image from performance incorporating one master Falcon robot and four slave Falcon robots.*

4. CONCLUSIONS

We find this application compelling. In particular, we informally found the synthesized sound to be lifelike even though it was generated by robots, and even when the robots were playing superhuman parts. These superhuman parts were generated by mechanical looping, where trajectories were transformed that were previously generated by a real human user while teleoperating the instruments. We believe that this application combines the advantages of computer music with the lifelike sounds of acoustic musical instruments. As robots become even more affordable to the computer music community, we believe that Mechanical Sound Synthesis will play an increasingly important role. We look forward to re-experiencing the computer music revolution from the point of view of mechanical engineering and Mechanical Sound Synthesis.

5. REFERENCES

[1] Max Mathews, “The digital computer as a musical instrument,” *Science*, vol. 142, no. 3592, pp. 553–557, November 1963.

[2] Curtis Roads, *The Computer Music Tutorial*, MIT Press, Cambridge, MA, February 1996.

[3] J. Torkington, S.G.T. Smith, B. I. Rees, and A. Darzi, “Skill transfer from virtual reality to a real laparoscopic task,” *Surgical Endoscopy*, vol. 15, pp. 1076–1079, 2001.

[4] Julius O. Smith, *Physical Audio Signal Processing: For Virtual Musical Instruments and Audio Effects*, W3K Publishing, http://ccrma.stanford.edu/~jos/pasp/, December 2008.

[5] Ajay Kapur, “A history of robotic musical instruments,” in *Proc. Internat’l Computer Music Conf.*, Barcelona, Spain, Sept. 5-9 2005, pp. 21–28.

[6] Byeong Sam Jeon, “Telematic drum circle,” *Leonardo*, vol. 43, no. 1, pp. 11, February 2010.

[7] Pat Metheny, “Orchestrion,” Available at http://www.patmetheny.com/orchestrioninfo/, accessed April 04, 2010.

[8] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino, “Augmented reality: A class of displays on the reality-virtuality continuum,” in *Proceedings of SPIE Telemanipulator and Telepresence Technologies*, 1994, vol. 2351, pp. 282–292.

[9] Claude Cadoz, Annie Luciani, and Jean-Loup Florens, “Synthèse musicale par simulation des mécanismes instrumentaux,” *Revue d’acouqistique*, vol. 59, pp. 279–292, 1981.

[10] Miikka Tikander, “Usability issues in listening to natural sounds with an augmented reality audio headset,” *Journal of the Audio Engineering Society*, vol. 57, no. 6, pp. 430–441, June 2009.

[11] Edgar Berdahl, *Applications of Feedback Control to Musical Instrument Design*, Ph.D. thesis, Stanford University, Stanford, CA, USA, January 2010.

[12] Henri Boutin and Charles Besnainou, “Physical parameters of an oscillator changed by active control: Application to a xylophone bar,” in *Proceedings of the 11th International Conference on Digital Audio Effects*, Espoo, Finland, Sept. 1-4 2008.

[13] Edgar Berdahl, “Edgar Berdahl’s YouTube Videos,” Available at http://www.youtube.com/user/eberdahl810, accessed April 04, 2010.