SoundFORMS: Manipulating Sound Through Touch

Aubrey Colter
MIT CSAIL
32 Vassar St.
Cambridge, MA 02139 USA
ajcolter@mit.edu

Patlapa Davivongsa
MassArt Dynamic Media Institute
621 Huntington Ave.
Boston, MA 02115 USA
pdavivongsa@massart.edu

Donald Derek Haddad
MIT Media Lab
75 Amherst St.
Cambridge, MA 02142 USA
ddh@mit.edu

Halla Moore
MIT CSAIL
32 Vassar St.
Cambridge, MA 02139 USA
halla@mit.edu

Brian Tice
MIT Media Lab
75 Amherst St.
Cambridge, MA 02142 USA
tice@mit.edu

Hiroshi Ishii
MIT Media Lab
75 Amherst St.
Cambridge, MA 02142 USA
ishii@media.mit.edu

CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2892414

Abstract
UPDATED—16 February 2016. SoundFORMS creates a new method for composers of electronic music to interact with their compositions. Through the use of a pin-based shape-shifting display, synthesized waveforms are projected in three dimensions in real time affording the ability to hear, visualize, and interact with the timbre of the notes. Two types of music composition are explored: generation of oscillator tones, and triggering of pre-recorded audio samples. The synthesized oscillating tones have three timbres: sine, sawtooth and square wave. The pre-recorded audio samples are drum tracks. Through the use of a gestural vocabulary, the user can directly touch and modify synthesized waveforms.

Author Keywords
Music composition; Music synthesis; Shape-changing interface; Tangible user interface.

ACM Classification Keywords
H.5.1 Audio Input, Output; H.5.2 User Interfaces, Auditory (non-speech) feedback; H.5.5 Sound and Music Computing, Signal analysis, synthesis, and processing.

Introduction

In 1897, Thaddeus Cahill invented a device called the Telharmonium that transmitted musical tones across electronic wire using the method of additive synthesis [9]. In 1965, Dr. Robert Moog outlined a modern electronic synthesizer design in his paper, Voltage-Controlled Electronic Music Modules [5]. Moog's inventions became mainstream with their inclusion at the Monterey International Pop Festival in June 1967 [4]. The idea of electronic musical synthesis as a means to replace live musicians continues to evolve. Work on SoundFORMS augments this previous work with the aim of pushing musical synthesis into a more tangible format, where composers can physically shape the sounds they create.

In general, music synthesizers generate electronic tones or notes via oscillations with frequencies that correspond to pitches on the musical scale. A note's timbre, or audio quality, can be varied by changing the shape of the oscillating signal. With SoundFORMS, these oscillations are generated when the user interacts with a three-dimensional shape-shifting display [1]. Musical sequencers take these synthesized notes and organize them in time. SoundFORMS allows the composer to physically interact with the shape of the oscillatory frequencies. The user can sculpt the actual wave which affects the timbre of the note. Specifically, the user can invoke sine, saw, and square wave variations. In addition, the user can trigger predefined drum beat patterns in a drum sequencer section. Collaboration is possible with multiple composers working in tandem; one operating the wave synthesizing section, and the other programming the drum machine section.

Related Work

Many variations of the music synthesizer exist, including three dimensional synthesizers. One contemporary example is software that accompanies compositions written by the group George and Jonathan [8]. Their application runs in a web browser and visualizes their compositions in a rotatable, three dimensional map. This method has great impact but is limited to a screen.

Other sequencers attempt to visualize the sound with intangible pixels or lights. The Pocket Operator by Teenage Engineering visualizes actions performed by the user, rather than the sound [3]. Though interesting, it does not address the concept of physically representing sound.

Yamaha's TENORI-ON allows users to interact with the sound via a matrix of LED switches, which convey the created sound wave patterns in two dimensions [6].

A project called PocoPoco moves musical synthesis into three dimensions with a solenoid-based sequencer grid that can be programmed in real time [2]. The solenoid pins raise at user-programmed intervals. To extend a note, the user can grab and pinch the pin. Their approach is beautiful but does not allow for visualization and manipulation of the sound waves in space.

SoundFORMS builds upon this prior work and makes use of Shape Displays [1]. These displays can render three-dimensional content in a physical manner, allowing users to interact with the information. By leveraging the Shape Display and touch detection, SoundFORMS truly migrates music synthesis and sequencing into three dimensions.

Concept

Using SoundFORMS, a user can program sound sequences, visualize sequence pulses, and modify sounds with touch. A user can compose music or drum beat patterns by triggering the control pins for the synthesizer or the drum machine.

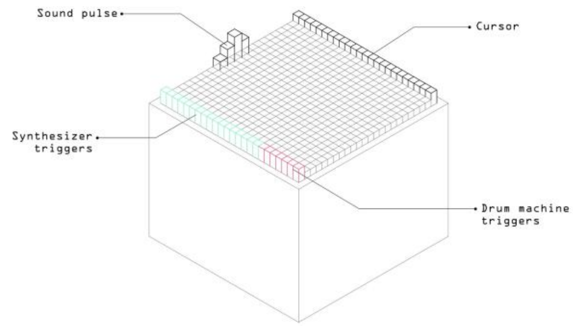


Figure 1: The Shape Display's row of pins closest to the user acts as buttons that trigger sound waves and drum beats. The user hears the sounds and sees the waveforms in real time.

When the user triggers a pin, the corresponding note is played through speakers, and a representation of a sound wave is created in the pins in front of the trigger, as shown in Figure 1. The wave marches toward the opposite end of the display, and loops back to the pins closest to the user to represent a continuous note. This visualization shows the user how different waves interact.

A user can interact with the waveforms on the display and sculpt sound in real time. We designed a limited vocabulary of hand gestures that facilitates this manipulation, as shown in Figures 2 and 3. The user performs the gestures above the pins, which are captured by a Microsoft Kinect. Though waveform manipulation is not original, it is traditionally computer-based, requiring a mouse and keyboard. Our approach allows the composer to eliminate the computer screen and peripherals while composing, opening new modes for music creation and performance. Analogous to composing music at a piano, our device uses hand-based manipulation with both visual and tactile affordances for the user.

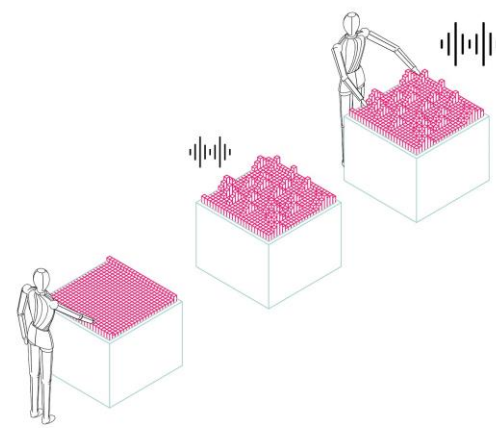
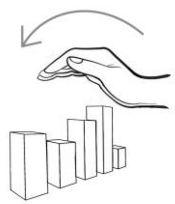


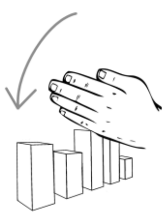
Figure 2: The user can create waveforms and manipulate them with the predefined gesture vocabulary.

Multiple users can collaborate and compose music at the same time at the same display.

Figure 3: Gesture vocabulary a user can perform to change the shape of the sound wave.



a) Sine Wave: arc the hand as if tracing over a hill.



b) Sawtooth Wave: perform a chop motion that is perpendicular to the wave.

c) Square Wave: chop the wave perpendicularly with both hands.

Interaction

When starting, the Shape Display presents a blank canvas, with the exception of the control pins, to the user. The control pins are elevated to indicate functionality. Two pins are assigned to each of seven scale notes with one spacer pin between each pair that corresponds to a spacer column on the display. The musical scale can be changed based on the user's preference. Four pins are assigned to different pre-recorded drum sounds.

The synthesizer columns form a wave, the shape and frequency of which corresponds to the actual sound wave of the note being played. The drum column pins march up the display, away from the user, in a pattern that matches the beat being played in the pre-recorded audio. The sound wave or beat pattern loops back to the pins in the column at the front of the display, closest to the user, until the user turns off the column's sound.

A user can physically manipulate the waves produced by the Shape Display with our gesture vocabulary, illustrated in Figure 3. The gestures allow users to change the shape of the waveform, which in turn changes the sound's timbre. We developed four gestures that allow the user to alter the sound with their hands:

1. To sculpt a wave into a sine wave, the user performs a motion that looks like they are running their hand over a hill, gently gliding the hand along the tops of the pins.
2. To transform a wave into a sawtooth wave, the user chops the wave with a single, vertical hand.
3. To change a wave into a square wave, both hands are held vertically, slightly apart, and the user brings them down toward the table.
4. To terminate a sound, the user presses the pins down toward the table with their palm.

The first three gestures can only be applied to sound waves. The gesture to stop the sound can be applied to both sound waves and drum beats.

Technical Overview

Our system is built on top of the inFORM: CooperFORM, a 24x24 pin square dynamic Shape Display [1]. Our code is written in Javascript and interfaces with CooperFORM through openFrameworks.

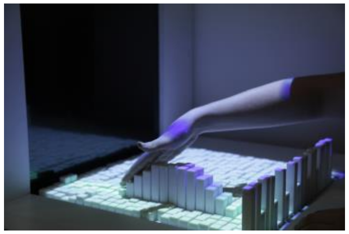
We use the web audio API [7] to generate the different synthesizer sounds and drumbeats.

User input to the system is detected by a Kinect mounted above the Shape Display. This interaction is bubbled up to Javascript handlers for the various gestures.

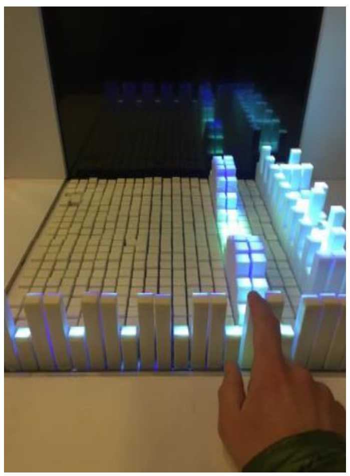
We project light onto the pins in the columns that are in use. Based on the pin height, retrieved from the Shape Display, we create a simple bitmap image. If the column is playing, the tall pins in the wave are highlighted blue, and the short pins are highlighted green.

The final system is illustrated by the photographs in Figure 4.

Figure 4: Photographs of a working SoundFORMS prototype.



a) A user shapes the sound wave.



b) A user triggers a sound wave while drum beats also play.

Current Limitations

Various limitations exist with SoundFORMS. First is the inherent issue of pin resolution. The square inFORM display has a total of 576 pins. This issue could be resolved by migrating to a Shape Display with higher resolution, though such a display does not currently exist.

Another limiting factor is the availability of high resolution pin-height detection. The resolution is currently limited to 8 bits, or 256 possible pin heights.

The inFORM display itself is bulky and not modular. Its current form factor is best suited for a semi-permanent installation in a fixed location.

The system also lacks precise tempo control. Further work needs to be conducted to fine tune the actuation and timing of the pins.

The device is limited by the current narrow gestural vocabulary. As more gestures are implemented in the software, the device has the potential to be more useful and more enjoyable.

Finally, the sound quality could be increased by gathering samples and integrating MIDI instrument sounds.

Application

SoundFORMS can be used in live concerts by electronic musicians and DJs. They can make synthesized music and SoundFORMS will create a beautiful visualization for their audience. Collaboration is also possible; multiple composers could operate each section independently.

SoundFORMS can be used to demonstrate a synthesizer for music education, teaching concepts of frequency, timbre, and superposition.

Finally, SoundFORMS can be used as a stand-alone instrument in recording studios and concerts.

Future Work

A natural extension of the current implementation of SoundFORMS is expanding the gesture vocabulary for wave sculpting. We also consider employing a Leap Motion, instead of the Kinect, for more accurate hand sensing.

SoundFORMS aims to provide the user with an experience touching sound. However, in the current implementation, the user can touch the pins, but a Kinect is used to recognize the gestures performed, and the display responds accordingly. We are limited by the Shape Display’s construction and the pins’ motor accuracy in sensing touch. A future iteration of SoundFORMS, potentially on a different Shape Display, will fully integrate touching the pins as the primary means of interaction.

We envision creating a second composition mode with still pins. In this mode, the composer adjusts the pin heights to build and shape still waveforms, which can affect the timbre. The Shape Display saves the user’s placement of pins, and the sounds can be played back in the original mode.

Another addition is to implement additive synthesis of waveforms in the second composition mode. Beat patterns can be built by using the principle of superposition with adjacent waveforms. By examining a cross-section of the display, the user can see how beats and waves align.

Due to the plastic cases of the Shape Display's pins, acoustic sounds also occur as the pins move up and down. The sound is a light clacking and varies with the speed and pattern of the pins' movement. We hope to explore the interaction of acoustic and synthetic sounds, perhaps finding a new genre of music that particularly suits the Shape Display.

We see a future iteration of SoundFORMS as a protoboard for audio mapping. Much like how a breadboard is used to prototype electronic circuits, soundFORMS could be used as a three dimensional musical mapping prototype surface to explore new types of human interaction with sound.

Conclusion

SoundFORMS is a sound synthesizer and sequencer that allows users to interact with a physical representation of a sound wave in real time, giving them the opportunity to touch sound. The only sequencer and synthesizer of its kind, SoundFORMS promises to add value to the music-making community.

Acknowledgments

We thank the Tangible Interfaces Group at the MIT Media Lab, especially Xiao Xiao, Ken Nakagaki, Viirj Kan, and Luke Vink.

References

1. Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, Hiroshi Ishii. 2013. inFORM: Dynamic Physical Affordances and Constraints through Shape and Object Actuation. In *UIST'13 Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, 417-426.
2. Yuya Kikukawa, et al. A design process of musical interface "PocoPoco": An interactive artwork case study. 2013. In *International Journal of Asia Digital Art & Design*, vol. 17, 26-35.
3. Peter Kirn. How TE's $59 Drum Machine Sounds – And How The Pocket Operators Work. 2015. Retrieved December 8, 2015 from http://createdigitalmusic.com/2015/01/heres-tes-59-pocket-operator-rhythm-sounds-work/.
4. Monterey International Pop Festival. Retrieved on December 8, 2015 from http://montereyinternationalpopfestival.com/.
5. R. A. Moog. 1965. *Journal of the Audio Engineering Society*, Vol. 13, No. 3, pp. 200–206.
6. Yu Nishibori, Toshio Iwai. TENORI-ON. 2006. In Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME06), 172-175.
7. Web Audio API by Mozilla. Retrieved on December 5, 2015 from https://developer.mozilla.org/en-US/docs/Web/API/Web_Audio_API.
8. The Website of George and Jonathan. Retrieved December 8, 2015 from http://www.georgeandjonathan.com.
9. Jay Williston. Thaddeus Cahill's Teleharmonium. Retrieved on December 5, 2015 from http://www.synthmuseum.com/magazine/0102jw.html.