Spectrogram: An Exploration of the Fourier Transform

Jessica Li **Benjamin Scharf** **Nathaniel Wong**

Northwestern University, Evanston, IL USA
{jessicali1.2018, bscharf, nwky} @u.northwestern.edu

ABSTRACT

The concept of sound as a pressure wave is one that, on its face, is relatively easy to understand. Comprehending how a seemingly homogenous sound is actually made up of a wide range of frequencies, however, is more challenging. We propose an exhibit that allows learners to build knowledge through tangible and collaborative interactions—a research-backed approach that allows exhibit visitors to understand how an individual sound can be decomposed into frequency and amplitude pieces. The proposed exhibit focuses on the outcomes of user-created sound rather than the mathematics behind the transform, focusing on enabling learners to understand sound concepts through their own exploration. We find that test visitors enjoy the open-ended nature of the exhibit, and explore and interact with their peers as they develop their understanding of sound and wave decomposition.

INTRODUCTION

Museums are moving away from information panels and static exhibits and introducing new kinds of interaction that support groups of visitors and allow for collaborative learning [3, 11]. Traditionally, many studies have focused on tangible interaction systems to allow users to interact with physical objects. This fosters the exploratory, hands-on learning that museums desire [19]. The creation of sound via voice or movement is another form of interactivity that further allows for exploration, and the use of the body makes for a visceral experience. Especially when paired with other senses such as visual, the experience of sound is a form of multimodal interaction that is effective in fostering understanding in visitors [21].

This project provides visualization of sound waves in real-time, creating an immersive and interactive audio and visual environment. We take advantage of the inherent integration of the human senses [5, 7] through multimodal information processing to help visitors build intuition and understand that physical phenomena such as sound are ultimately composed of waves. Visitors also have the opportunity to explore other properties of the sounds they make and hear.

Learning Goals

The proposed exhibit is designed to aid visitors in their exploration of the concept of the Fourier transform and introduce them to some background knowledge and terminology about waves. It will set up opportunities to learn these concepts so that interaction with the main part of exhibit will be as fruitful as possible. In sum, our proposed exhibit has the following learning goals:

1. Visitors might appreciate that all the sounds that we hear are composed of longitudinal waves. This is true even for ultrasound, which is a sound of frequencies higher than can be heard by the human ear.
2. Visitors might discover that there are many sounds, including musical instruments, produce oscillating waves that produce fundamental frequencies (1st harmonic) and overtones (higher harmonics). Overtones are a fundamental phenomenon in nature and are what gives musical instruments their unique, rich sounds.
3. Visitors might learn that all sounds can be decomposed into a series of fundamental sound waves, or sinusoidal waves. The amplitudes of these sinusoidal waves can be represented on a spectrum to represent a unique sound at any given point in time. Each sinusoidal wave will have a unique frequency and a given amplitude, and can be mathematically summed to reconstruct the original waveform.
4. Visitors might collaboratively explore how sound spectra vary according to different types of sounds or music, or voices. They will be able to choose from a selection of audio samples that include phenomena in everyday life, musical notes played by various instruments, and compare how the sounds are represented on a Fourier transform, including at different pitches.

RELATED WORK

Our work builds on existing literature pertaining to the intersections of audio representation, interactive exhibits, and education. These can be grouped into three related but separate areas: *understanding sound, making sound tangible,* and *immersive audio exhibits.*

Understanding Sound

A study on fifth grade students' understanding of ideas about sound revealed that students had significantly better understanding when using an exploratory, hands-on, and application-based learning process than when using a traditional textbook and demonstration approach [2]. Exploratory and hands-on learning fit well in the informal learning environments of museums. An example of an exhibit focused on the underlying principles of sound is given in [8]. The authors present an interactive exhibit where visitors use Sphero Robotic Balls to explore the effects of audio controls such as delay, sustain, and reverb.

A more scientific presentation for understanding sound can be found online at the Chrome Music Lab [17]. It shows the frequencies that make up a sound that the user chooses from a set of pre-recorded instruments or sounds, or the user's microphone input. The results are displayed as a spectrogram that moves from left to right with time.

Making Sound Tangible

In [14], the authors create an interactive exhibit about medieval music, focusing on a pure audio system without any visual feedback. The exhibit is centered around being able to interact with and play with authentic replicas of medieval instruments, allowing visitors to learn about medieval music by interacting with physical instruments.

The authors of [20] present a device that allows audio engineers with visual impairments to feel the amplitude of a sound, information that sighted engineers would get through seeing waveforms on a screen. The system with tangible interaction allows users to understand the properties of sound through physical manipulation and feedback. The authors in [22] developed an emoti-chair that uses haptic feedback throughout the user's body to deliver a rich, sensory experience for people with auditory impairments, exploring the idea of using tactile modality to experience music.

Immersive Audio and Visual Exhibits

Audio and visual sensory inputs can be combined to create immersive and engaging environments. The work presented in [4] proposes a new form of interaction between people and information by augmenting the physical environment through a dynamic soundscape, enhancing the impact of artistic installations and entertainment events.

The authors in [21] present a method for augmenting the visual environment in a personalized way with audio, and show that combining audio and visuals is effective in fostering understanding in users. These environments are also conducive to co-creation and participation: in [1], the authors present an exhibit that allows visitors to collaboratively create a complex sonic environment through the use of auditory icons.

BACKGROUND

We want to provide visitors with a more scientific and detailed, yet still engaging, understanding of sound. Combining the scientific information conveyed by Chrome Music Lab's Spectrogram [17] and tangible interactivity of Gonzalez et al.'s SonicExploratorium [8], we present a tangible exhibit for understanding the scientific background of sound waves.

The visitor is presented with two large microphone stands in front of a screen, inviting the visitor to stand in front of it and speak or play other noises into it. The life-size microphones are intended to attract the attention of visitors as it serves as a cultural form that it can be interacted with in an auditory way. This aspect will also help encourage visitors to engage in a familiar activity that has been re-imagined to support a novel exploratory and learning experience [16].

Aside from interaction between users, collaboration is identified as crucial to exhibit design, the benefits of which are noted by Heller et al. in [10]. This will be explored by providing invitations for visitors to work with their counterpart on the microphone, for example producing harmonized musical notes, or matching notes on a different octave. Extrinsic motivation is key to shaping visitors' behavior, and there will be cues to support their interaction with the exhibit, by way of instructions or a set of tasks that can be achieved. However, intrinsically motivated visitors can also try to explore sounds at their own pace not included in the explicit "design" of the exhibit, for example including clapping their hands or playing music from their mobile device.

A large screen displays a spectrogram of the users' voice, with time on the horizontal axis and frequency displayed on the vertical axis. Different colors are used to highlight the intensities frequencies representing the left and right microphone. The design will place a conscious effort to highlight the tight coupling between the visitors' use of the microphone and what is shown on the display, emphasizing the complexity of sound and that their limitations are only that of their voices and the sounds produced by objects around them.

On one side of the screen, visitors can learn what the lines on the screen mean, specifically referencing terms frequencies, harmonics, and additive synthesis. On the other side, they can learn about the Fourier transform and its applications, as well as the fact the Fourier analysis allows us to decompose all sounds into a sum of sine waves at different frequencies. By providing both of these functionalities alongside one another, we hope visitors will gain a stronger understanding of how these different applications of the Fourier transform fit together. Piecing the steps of the transform together in this process will not only allow visitors to have fun, but will also allow them to learn about representations of sound implicitly through their own experimentation.

Research done by when teaching understanding about sound propagation [16] showed that tangible interaction is key in understanding physics concepts like sound. In particular, researchers used an example of the effect of sound on a physical particles of dust to allow their subjects to "map object-like properties onto sound waves." The creation of sound waves is an interaction that our exhibit supports, so we do plan on taking this prior research into consideration when building our final exhibit. One particularly valuable technique this research done by Wittman, Steinberg, and Redish offered was a kind of test-and-check process. Participants in the experiment were asked to predict how a sound wave would impact the fire on a candle, and then were able to compare that understanding with the reality. After allowing users to gain some basic familiarity with our system, we could prompt them to try and guess (or maybe sketch out) how our system would react when given a particular kind of sound. By allowing visitors to interact more deeply by requiring them to think and compare, it is possible that our exhibit will enable even greater learning.

An article written by Charles Linder as a part of a physics education journal also covers this topic of promoting understanding of sound. Linder focuses on this area in particular, as his research revealed that even physics graduate students had inaccurate or incomplete understandings about the concept of sound. In a passage of valuable insight, Linder notes that, while he does not want to "deny the importance of mathematics and problem-solving in physics teaching… teaching without a significant component of explicit conceptual exploration would seem to result in students' approaching their tutorial exercises as sets of exercises in mathematics rather than a means of coming to understand physics" [13]. It is this logic that led us to focus on the outcomes of viewing sound as a waveform rather than the precise process with which the sound can be decomposed. We focus on allowing learners to understand a process via their own modification of input rather than providing the explicit mathematical expressions that result in the transform's output.

In this context of multimedia learning where visitors to the exhibit provide sound that is played back and combined with visual representations, we also must consider how all of this information will be consumed by the learner. Research conducted by Mayer, Heiser, and Lonn indicates that, when working in multimedia contexts, it is important to consider situations where including an excess of information results in excessive cognitive load, and in turn reduced learning. In particular, the inclusion of "irrelevant video clips and narration segments to a coherent multimedia explanation can hurt student understanding of the explanation" [15]. This is a big factor in why our proposed exhibit is relatively minimal in its depiction of sound. While we could choose to include labels throughout and textual explanations in our actual exhibit, doing so would likely distract from the important action and response interaction we focus on. Allowing users to play around with the system and see how their interactions have an effect rather than explicitly explaining effects is one way we attempt to avoid this multimedia overload.

Other research—while ultimately focused on teaching understanding of properties of sound different than those our exhibit focuses—indicated that the introduction of computer simulation was effective in helping students deepen their initial understandings of how sound worked. Authors Gunhaart and Srisawasdi found through their own experiment with teaching through computational models found "reconceptualization, around four physics concepts of sound wave properties including reflection, refraction, diffraction, and interference" [9]. Despite differences in what our exhibit aims to convey, we believe that the common ground in concepts of sound wave properties means that our exhibit will be an effective tool in teaching core sound concepts.

We base our museum exhibit on this existing body of research as we aim to produce an effective learning experience for the complex concept of sound. We view our multimedia exhibit as a culmination of research and our own knowledge and experience learning how sound works, and how it can be mathematically and visually described. We are confident that the tangible interactions afforded by the microphones, auditory, and visual feedback in conjunction with collaborative opportunities the exhibit affords will be effective in helping our visitors understand sound, its visible and tangible forms, and its decompositions.

IMPLEMENTATION

Our exhibit design consists of a physical microphone, different instruments and sound-making-objects for users to try, and a screen that will display the spectrogram from the microphone input.

Users will be able to explore the frequency spectra of sounds that they make through the microphones, most intuitively by using their voice, including speaking, singing, and whistling.

At any time, up to two groups of visitors will be able to speak into the microphone and immediately view a Fourier transform of their voice on the screen. They will be able to interact with each other, such as mimicking voice pitch, comparing spectrograms, or producing harmonies with other people.

Visitors can also choose to pick up objects on the table that represent different sounds and musical instruments. They can then compare what they hear to what is displayed on the screen, mimic the sounds they hear with their voice, or even combine different sounds and voices to create a cacophony of noises. The can also produce other sounds like clapping or stomping, or trying an array of musical instruments available to them (such as a harmonica, whistle, ukulele, shaker, triangle and other percussion instruments).

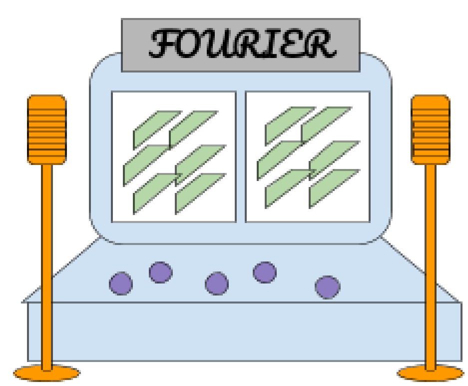

Figure 1. Exhibit layout containing screens, microphones, and table surface containing objects.

For the spectrogram display, we adopted code from Chrome Music Lab's web implementation as it has engaging animations, especially compared to other open source implementations. An example of our modified spectrogram display is shown below in Figure 2.

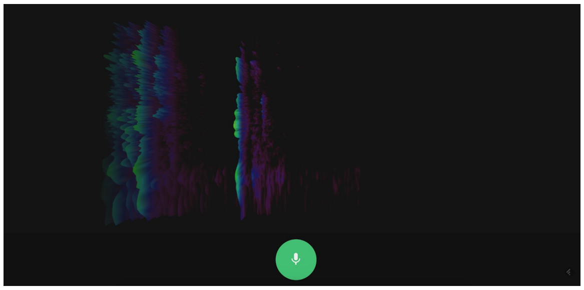

Figure 2. Spectrogram Display using color to represent frequency.

Our final design consists of two of these spectrograms that are color coded, with each color representing the audio representation of the two microphones, as seen in Figure 3.

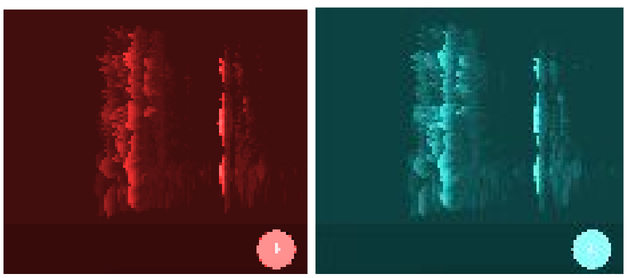

Figure 3. Side-by-Side Spectrograms corresponding to two different microphones.

Visitors will be presented with an inviting space to explore their voice through the microphones, or pickup nearby objects and musical instruments to use. A large display will show the spectrograms in real time as visitors interact with the exhibit. The screen also shows hints on what they can experiment with, like whistling or clapping their hands, to expand their realm of possibility. A mockup of this can be seen in Figure 4.

Figure 4. Final design in use with microphone, guitar, shaker, and screen

There is also an accompanying graphic that explains the concepts of the spectrogram to be placed separate from the screen so that users can read more in-depth about what they are seeing. This graphic can be found in the Appendix. A link to our demo can be found at https://www.youtube.com/watch?v=7c6LstFCh3c .

DISCUSSION

We performed a short user testing session with our design with a laptop, physical microphone, and various Boomwhackers of different musical pitches. Since our goal was to create an engaging and tangible exhibit that allows for collaboration and exploration, we observed user behavior as it pertains to engagement, exploration, and collaboration.

We found that users were immediately drawn to the visuals on the screen. Many noted that the 3D effect and colors of the spectrogram were very interesting. In addition, as we were working on our exhibit in the hallway one passerby noted that, "That is the coolest thing I've ever seen." Based on these observations, we felt that our point of initial engagement is interesting enough to attract visitors. Once at our exhibit, users then noticed the microphone and Boomwhackers, and all of them spent at least two minutes exploring the different effects of their actions. During our testing session we explained the principles of the spectrogram while they explored, but in our final prototype we changed this to an accompanying graphic that explains what the user is seeing. The graphic is separate from the spectrogram on the screen so that the user can read it if they want to, or disregard it if they simply want to play with the noise-making objects and spectrogram. Thus according to the principles of active prolonged engagement [12], the exhibit provides interesting visuals for initial engagement, and then allows for open-ended exploration with some gentle guidance.

We found that the exhibit is conducive to exploration. The user can interact in many ways using just their voice, and users enjoyed using the tangible forms of interaction, such as the Boomwhacker sticks to spot differences in the spectrogram. Users also did some unexpected things, such as playing songs from their phone into the microphone to view the spectrogram. Because the exhibit is open-ended and there are infinite ways to create sounds, the exhibit allows for exploration and thus individual discovery [18].

We also found that the format of the exhibit lends itself to collaboration. The microphone can pick up sounds from multiple sources at once so multiple people can use the microphone, and there are multiple tangible noise-making objects. We observed that users coordinate with each other to make certain sounds to observe. For example, two users coordinated the whacking of two Boomwhackers to see what the spectrogram would look like with both of those pitches at the same time. The exhibit is also scalable with more microphones and more noise-making objects.

CONCLUSION AND FUTURE WORK

We present a tangible exhibit that introduces the concepts of the Fourier Transform through sound. By focusing on the outcomes of viewing sound rather than the precise scientific process, we allow users to understand a process via their own explorations. The exhibit is composed of various noise-making objects, a screen which displays the results of the users' explorations in real-time, and a physical microphone which takes advantage of cultural forms to encourage users to partake in a familiar activity.

Currently, the implementation of spectrogram output from microphone input can work with any computer-microphone pair, but this is not as conducive to making direct comparisons of the spectrogram between two distinct users, because the users would have to translate what they are seeing between different screens. An easier way to make direct comparisons would be two have two separate microphone inputs on one computer screen.

We would also like to explore other methods of tangible interaction for introducing the concepts of sound and Fourier Transform. An exciting direction would be to have a set of strings whose frequencies we could tune. Then we could decompose a user's voice into his or her fundamental frequencies and change the frequency of the strings to match, creating an individualized and playable instrument for each user's voice.

ACKNOWLEDGMENTS

We would like to thank Professor Michael Horn for his great ideas and endless support, and Mmachi and Izaiah for their help during this fun and educational quarter.

REFERENCES

1. Back, M. and Cohen, J., 1998, April. The sound of your stuff: designing a complex auditory display for an interactive museum exhibit. In CHI 98 Conference Summary on Human Factors in Computing Systems (pp. 239-240). ACM.
2. Barman, C.R., Barman, N.S. and Miller, J.A., 1996. Two teaching methods and students' understanding of sound. School Science and Mathematics, 96(2), pp.63-67.
3. Clark, H. and Brennan, S. Grounding in communication. In L.B. Resnick, J.M. Levine and S.D. Teasley, eds., Perspectives on socially shared cognition. American Psychological Association, Washington D.C., 1991.
4. Eckel, G., 2001. Immersive audio-augmented environments: the LISTEN project. In Proceedings fifth international conference on information visualisation (pp. 571-573). IEEE.
5. Eisenstein, S., Leyda, J. (1942): "Filme Sense". Harcourt Brace Javanovich.
6. Florian Heller, Thomas Knott, Malte Weiss, and Jan Borchers. 2009. Multi-user interaction in virtual audio spaces. In CHI '09 Extended Abstracts on Human Factors in Computing Systems. ACM.
7. Gaver, W. (1989): "The Sonic Finder: an Interface that Uses Auditory Icons". Human Computer Interaction, 4.
8. Gonzalez, B., Adams, A.T. and Latulipe, C., 2014, April. SonicExploratorium: an interactive exhibit of sonic discovery. In CHI'14 Extended Abstracts on Human Factors in Computing Systems (pp. 395-398). ACM.
9. Gunhaart, Audcharaporn, and Niwat Srisawasdi. "Effect of Integrated Compute-based Laboratory Environment On Students' Physics Conceptual Learning of Sound Wave Properties." Procedia-Social and Behavioral Sciences 46 (2012): 5750-5755.
10. Horn, M.S., 2018. Tangible Interaction and Cultural Forms: Supporting Learning in Informal Environments. Journal of the Learning Sciences, 27(4), pp.632-665.
11. Hornecker, E., Stifter, M. Learning from interactive museum installations about interaction design for public settings. Proc. of OzCHI 2006. ACM, pp. 135-142
12. Humphrey, T. and Gutwill, J.P., 2017. *Fostering active prolonged engagement: The art of creating APE exhibits*. Routledge.
13. Linder, Cedric J. "Understanding sound: so what is the problem?." Physics Education 27.5 (1992): 258.
14. Maria Karam, Carmen Branje, Gabe Nespoli, Norma Thompson, Frank A. Russo, and Deborah I. Fels. 2010. The emoti-chair: an interactive tactile music exhibit. In CHI '10 Extended Abstracts on Human Factors in Computing Systems. ACM.
15. Mayer, Richard E., Julie Heiser, and Steve Lonn. "Cognitive constraints on multimedia learning: When presenting more material results in less understanding." Journal of educational psychology 93.1 (2001): 187.
16. Michael Wittmann, Richard N. Steinberg & Edward F. Redish (2003) Understanding and affecting student reasoning about sound waves, International Journal of Science Education, 25:8, 991-1013.
17. Music Lab, "Spectrogram." Chrome Experiments. https://musiclab.chromeexperiments.com/Spectrogram
18. Oppenheimer, F., 1976. EVERYONE IS YOU OR ME/E. *Technology Review*, *78*(7), pp.30-35.
19. Price, S., Rogers, Y., Scaife, M., Stanton, D., and Neale, H. Using tangibles to promote novel forms of playful learning. Interacting with Computers 15, 2 (2003), 169–185.
20. Tanaka, A. and Parkinson, A., 2016, May. Haptic Wave: a cross-modal interface for visually impaired audio producers. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 2150-2161). ACM.
21. Terrenghi, L., Zimmermann, A. Tailored audio augmented environments for museums. In Proceedings of the 9th international conference on Intelligent user interfaces (IUI '04). ACM, New York, NY, USA, 334-336.
22. Wolf, M., Lee, E. and Borchers, J., 2007, April. Education, entertainment and authenticity: Lessons learned from designing an interactive exhibit about

medieval music. In CHI'07 Extended Abstracts on Human Factors in Computing Systems (pp. 1887-1892). ACM.

APPENDIX