Reducing the Effect of Imperfect Microphone and Speaker in Audio Feedback Systems

Lilac Atassi

Leonardo Music Journal, Volume 29, 2019, pp. 37-39 (Article)

Published by The MIT Press

REDUCING THE EFFECT OF IMPERFECT MICROPHONE AND SPEAKER IN AUDIO FEEDBACK SYSTEMS

Lilac Atassi

ABSTRACT An audio feedback system that iteratively uses a room as a sound filter can be an artistic medium generating fascinating sounds. In this system, the room is not the only component acting as a filter. The sound system component, i.e. the speaker and microphone, also can have a sizeable impact on the sound in each iteration. To make sure the relative influence of the room on the sound is revealed and not masked by the audio system, the author proposes using a common calibration method at the end of each iteration. The mathematical model of the system is used to explain the reasoning behind the use of this method. Following this procedure, the author conducted an experiment that shows sound interaction with the room over time being captured in the artwork.

An audio feedback system can be realized using a microphone and a speaker. Alvin Lucier's *I Am Sitting in a Room* [1] is a seminal artwork that uses an audio feedback system. Lucier first records his voice when reading a short passage. The recorded sound then is played back for several iterations. In each iteration, the microphone is used to record the sound for the next iteration. In effect, in each iteration the sound goes through the audio feedback system. In later iterations, the sound texture and quality changes to the point that it is impossible to tell the original sound was a recorded voice.

Lucier and Ashley [2] explain that the room acts as a sound filter every time the tape is played back. A second sound filter in this system is the audio system (the combination of the speaker and microphone). In this article, I argue that the effect of the audio system filter can be larger than that of the room sound filter. To not leave the relative filters' strength to chance, one can estimate and reduce the effect of the audio system. Therefore, the room has a stronger footprint on the sound in each iteration.

Uncalibrated Audio System Problem

When a sound goes through the speaker and the microphone in an audio feedback system, some frequencies get boosted and some attenuate. This is an undesirable property of the hardware. In order to reduce this undesired effect, we need first to measure it. In this section, I first discuss the method I used to do so. Second, using a mathematical model, I discuss how the measured frequency response can be compared to the effect of the room.

Lilac Atassi (researcher), Department of Music, California Institute of the Arts, Valencia, CA 91355, U.S.A. Email: lilacatassi@alum.calarts.edu.

See www.mitpressjournals.org/toc/lmj/29 for supplemental files associated with this issue.

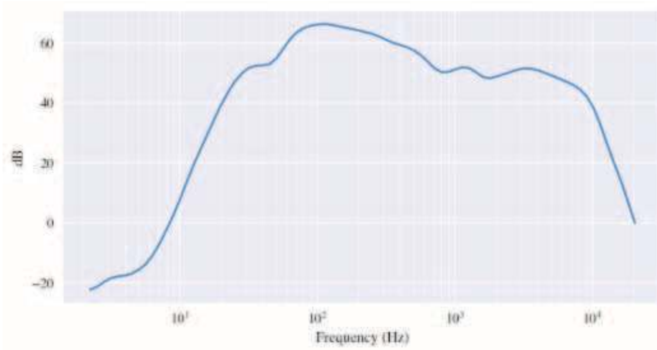

Fig. 1. The frequency response of the audio system. (© Lilac Atassi)

Using the free software package Room EQ Wizard (REW) [3], I recorded the frequency response of the microphone and speaker. REW sweeps a fixed amplitude sinusoidal tone from 2 Hz to 30,000 Hz over 6.0 seconds. In my experiments, the microphone is placed above the speaker and both are pointed in the same direction. Ideally, this recording should be done inside an anechoic room, but, as I did not have access to an anechoic chamber, I used a large number of sound-absorbing panels to reduce the effect of the room. I repeated the process of measuring the frequency response of the microphone and speaker in four rooms with varying dimensions and also at two different locations inside each room. The difference between the frequency response measurements was small. Therefore, the room effect in the frequency response measurement was negligible.

Figure 1 shows the frequency response of the speaker and microphone. In the frequency range of interest, between 40 Hz to 15 KHz, the fluctuation is over 10 dB. Therefore, with adjustment of the amplification for zero gain at the average amplitude of the frequency response, some frequencies will have a positive gain and some will have a negative gain.

A simple mathematical model can help us to compare the factors that modify the spectrum of the sound in each iteration. Let $a(t)$ denote the frequency spectrum of the sound recorded at a particular moment in iteration t. We want to model how the frequency spectrum is modified from iteration t to iteration $t + 1$, $\Delta a(t + 1) = a(t + 1) - a(t)$. By assuming this is a linear system, that is $a(t + 1) = H \cdot a(t)$, H is the transfer-function matrix, we can rearrange the difference equation as $\Delta a(t + 1) = (H - I) \cdot a(t)$, where I is the identity matrix. That means some elements of a (frequency spectrum) converge to zero and some others diverge. The convergence and divergence rates are proportional to the number of iterations.

If H is a diagonal matrix, the neighboring frequencies have no influence on each other. In practice, due to the nonlinearity of the audio system [4], the power in the frequency domain leaks to neighboring frequencies. But to simplify the discussion, I assume the nonlinearity effect is negligible, and therefore H is diagonal. H can be written as the sum of two matrices, $H = S + R$. S and R describe the effect of the sound system and the room respectively. Hence, we are assuming only the room and the audio system modify the sound between two iterations in our simplified model.

Figure 1 shows that the largest gain magnitude of S is about 10 dB. That means some frequencies are amplified to over 3 times the amplitude of other frequencies. This can be contrasted with the typical amplitude gain caused by the room acoustic. The strongest amplitude gain caused by reverberation will happen when the first sound reflection constructively interferes with the feed-forward sound. Assuming there is no sound energy loss after the reflection, the amplitude is at most twice the original value. Therefore, it is clear that the effect of the uncalibrated audio system can mask the effect of the room acoustic.

The previous model explains how the sound changes from one iteration to another. All delay is acoustical delay due to the size of the room; there is no significant delay in the electronic system.

Equalization

It is possible to mitigate the effect of the imperfect frequency response of the audio system by calibrating the system and equalization. The equalization filter h in the frequency domain is the inverse of the frequency response of the audio system, given that for a flat frequency response $DFT(c) \cdot DFT(h) = 1$, where $DFT(c)$ is the vector shown in Fig. 1. Applying the equalization filter to the recorded sound from one iteration yields the sound that should be played in the following iteration. The audio signal can be convolved with the equalization filter in the time domain or via the frequency domain. (These methods are equivalent.) For this project, I decided to convolve the audio signal with the equalization in the frequency domain, as the filter in the frequency domain is readily available. $y_{k+1} = IDFT\ (DFT(x_k) \cdot DFT(h))$, where x_k is the recorded sound at iteration k, y_{k+1} is the sound to be played at iteration $k + 1$ and $IDFT$ is the Inverse Discrete Fourier Transform. Note that as explained, $DFT(h)$ is simply the inverse of the audio system's frequency response; therefore, in the frequency domain, multiplying this filter by the frequency response yields a flat response.

A common approach is to use the overlap-add method to efficiently convolve the equalization filter h (the impulse response) with the audio signal x using Discrete Fourier Transform (DFT) [5]. Also note this equalization does not require execution in real time and is done by the developed software [6] only between each iteration of the process.

After applying the equalization filter and recording the frequency response using REW, the fluctuation in the frequency response of the audio system was less than 2 dB in the 40 Hz to 10 KHz range. Therefore, the effect of the audio system in each iteration will be smaller than the effect of the room.

To summarize, the equalization process has three steps. First, $DFT(h)$ is calculated once by inverting the frequency response of the audio system. Second, the recorded sound at the end of each iteration is converted into the frequency domain using the Discrete Fourier Transform $DFT(x_k)$ over short segments (following the overlap-add method). Then, using the Inverse Discrete Fourier Transform, the sound is converted back into the time domain and is ready to be played back for the next iteration.

Experiment

Using a condenser microphone, Bluebird-SL, and a loudspeaker, Yamaha HS7, I ran an experiment inside a 55-foot-diameter wooden dome following the same procedure as in *I Am Sitting in a Room*. The microphone was placed on the top of the speaker, and both were placed at the center of the room on the floor. Soundproofing foam tiles were placed between the speaker and the floor and between the microphone and the speaker. The initial sound used in this experiment was my own voice. The recorded sound at the end of each iteration first went through the equalization filter, and then the gain was adjusted to have maximum gain across all frequencies equal to one. This process is automated by custom computer code.

Figure 2 shows the spectrogram of the recorded sound during the first and 12th iterations, in the top and bottom panels respectively. Some resonances are apparent. Figure 3 depicts the spectra of the initial, first and 12th recorded sound. Note that in the 12th iteration sound, the distance between the three main peaks is about 40 Hz. These peaks are not present in the spectrum of the initial sound.

The large difference between the frequency response shown in Fig. 1 and the spectrum in Fig. 3 suggests that the audio system was not the dominant factor in shaping the sound, and most likely the room acoustic had the biggest impact on shaping the sound.

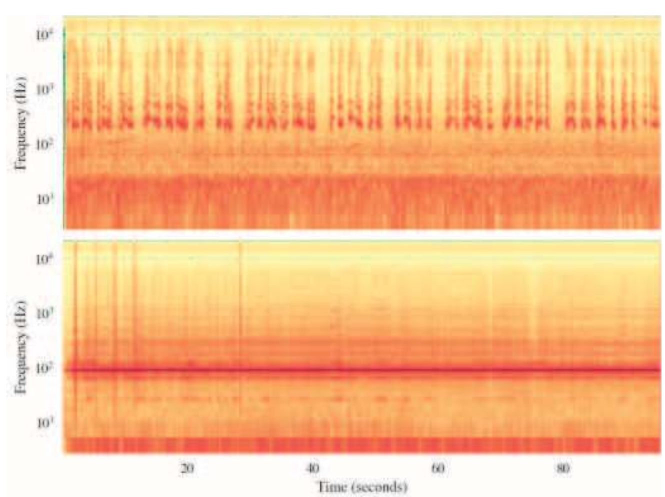

Fig. 2. The spectrogram after 12 iterations in the experiment run inside a wooden dome. (© Lilac Atassi)

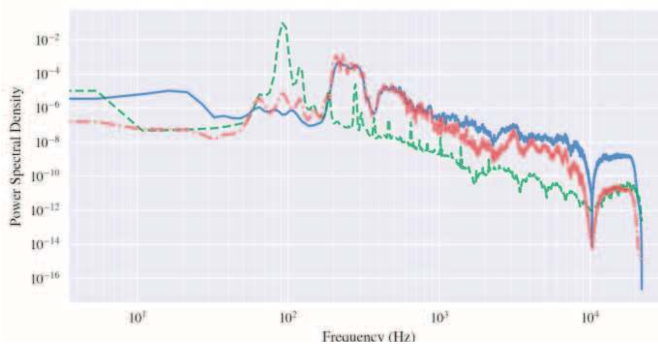

Fig. 3 The spectrum of original sound (solid line), after the first (dashed and dotted line) and 12th (dashed line) iterations in the experiment run inside the wooden dome. (© Lilac Atassi)

Conclusion

Many artworks have explored using a room as an audio filter. The microphone and speaker in an audio feedback system act as secondary audio filters. In this article, I show that, using equalization, it is possible to reduce the effect of the audio system's frequency-response on the sound. This frequency-response correction process allows the footprint of the room on the sound to be more pronounced. The result of the conducted experiment inside a wooden dome supports this claim.

Acknowledgments

While working on this project, I benefited from technical advice by Robert Wannamaker and Mosalam Ebrahimi, and I am greatly thankful.

References and Notes

1 Alvin Lucier, *I Am Sitting in a Room*: www.lovely.com/titles/cd1013.html (accessed 9 December 2018).

2 Alvin Lucier and Robert Ashley, *Music 109: Notes on Experimental Music* (Middletown, CT: Wesleyan Univ. Press, 2014).

3 Room EQ Wizard Software: www.roomeqwizard.com.

4 Wolfgang Klippel, "Tutorial: Loudspeaker Nonlinearities Causes, Parameters, Symptoms," *Journal of the Audio Engineering Society* **54**, No. 10, 907–939 (2006).

5 Monson H. Hayes, *Schaum's Outline of Digital Signal Processing*, 1st Ed. (New York: McGraw-Hill, 1998).

6 The filter code in Python can be found at www.github.com/lilac-atassi.

Manuscript received 19 October 2018.

LILAC ATASSI *is currently a PhD candidate in computer music at the University of California, San Diego. She holds a BFA in piano performance from the California Institute of the Arts (2017) and has obtained her MFA in music composition and experimental sound practices, also from CalArts (2019). Her current research interests are in the areas of interface design and experimental sound practices.*

doi:10.1162/LMJ_a_01060

3D NOTATIONS AND THE IMMERSIVE SCORE

David Kim-Boyle

ABSTRACT The author discusses his use of generative three-dimensional notations for representing musical forms. Several key works, programmed in the Max/OpenGL platform, are described in detail, and the author discusses current development with Microsoft's HoloLens. The author argues that such immersive technology promotes a physical engagement with the score in which the work is an emergent property of an open-ended play.

For over 15 years I have been exploring various nonlinear open-form musical structures afforded by generative real-time scores. Such scores often integrate complex nonlinear processes within generative techniques ranging from the use of Markov chains or other stochastic processes to determine the temporal ordering of events through to the use of data derived from timbral analysis to drive low-level structural transformations such as pitch distributions. The procedural generation of musical forms has been explored by a large number of composers, but only over the past ten years have composers begun to explore procedurally generated real-time performance scores [1]. Of particular interest in my creative and research practice are the use of three-dimensional scores in dynamic visualizations, which present performers with representations of musical form, the potentialities of which are explored through guided play [2].

Three-dimensional scores fundamentally present an effort to transcend the materiality of the printed page. While visual artists have grappled with the affordances of perspective since the fourteenth century, the applications of perspective and three-dimensional structures in musical notations have been of only relatively recent interest. The use of depth as a structural determinant is suggested in works such as *Fontana Mix* (1958), *Cartridge Music* (1960) or *Variations III* (1962) by John Cage [3–5] or in Toshi Ichiyanagi's *Music for Piano No. 7* (1961) [6], where printed transparencies are overlaid to create musical structures, and Kenneth Gaburo explores the use of superimposed text in his *Lingua II: Maledetto II* (1967–1968) [7]. In each of these works, while form emerges from material depth, the two-dimensional surface upon which the works are denoted is a fundamental constraint [8].

While I alluded to three-dimensional notations in my *Valses and Etudes* (2005, rev. 2011), for piano and computer, where pages of musical notation fly in and out of a musical display, it was not until my *point studies no. 2* (2013), for any two pitched instruments and computer, that I started to explore the musical affordances of three-dimensional notated structures with greater focus. The techniques I developed in *point studies no. 2*, including the use of stochastic processes to instantiate pitch and rhythmic structures and transfor-

David Kim-Boyle (composer), Sydney Conservatorium of Music,
The University of Sydney, 1 Conservatorium Road, Sydney, NSW 2000, Australia.
Email: david.kim-boyle@sydney.edu.au. ORCID: 0000-0001-9150-675X.

See www.mitpressjournals.org/toc/lmj/29 for supplemental files associated with this issue.