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

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

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, I propose using a common calibration method at the end of each iteration. The mathematical model of the system is studied to discuss the reason behind using this method. Following this procedure, I conducted an experiment that shows the sound interaction with the room over time is 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] [1] is a seminal artwork that uses an audio feedback system. He 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. At the 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 [1] 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 paper, 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 the sound goes through the speaker and the microphone, 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.

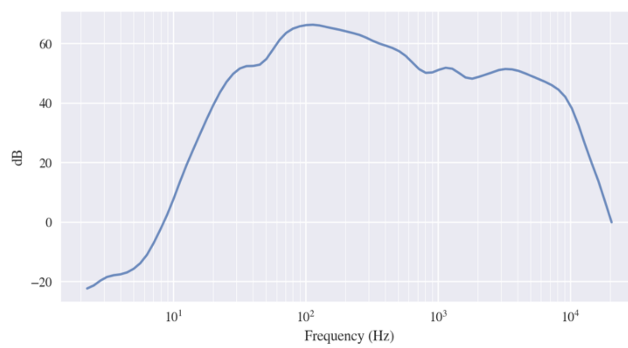
Using the free software package Room EQ Wizard (REW)[2], I recorded the frequency response of the microphone and speaker. REW sweeps a fixed amplitude sinusoidal tone from 2 Hz to 30,000 Hz over 6.0 seconds. In my experiments, the microphone is placed above the speaker and both are pointing to the same direction. Ideally, this recording should be done inside an anechoic room. 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

Fig. 1. The 55-foot diameter wooden dome where I ran the experiment in. The speaker and microphone were placed at the center. Soundproofing foam tiles are placed between the speaker and the floor, and between the microphone and the speaker.

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 2 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, by adjusting the amplification to have zero gain at the average amplitude of the frequency response, some frequencies will have a positive gain and some will have a negative gain.

Fig. 2. The frequency response of the audio system.



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 [2], the power in the frequency domain

[1] http://www.lovely.com/titles/cd1013.html

[2] https://www.roomeqwizard.com

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 2 shows 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 of the delay is acoustical delay due to the size of the room and there is no significant delay in the electronic system.

Equalization

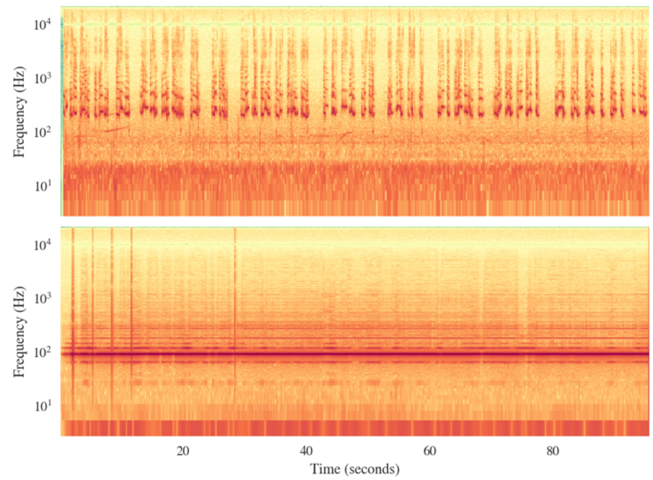
It is possible to mitigate the effect of the imperfect frequency response of the audio system by calibrating the system and equalization. The equalization filter h in the frequency domain is the inverse of the frequency response of the audio system. That is derived from the fact that for a flat frequency response $DFT(c) \cdot DFT(h) = 1$, where $DFT\ (c)$ is the vector shown in Figure 2. Then we can apply the equalization filter to the recorded sound from one iteration to get 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 will be played at iteration k+1, and $IDFT$ is the Inverse Discrete Fourier Transform. Note as explained, $DFT(h)$ is simply the inverse of of the audio system's frequency response, therefore in the frequency domain multiplying this filter by the frequency response we get 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) [3]. Also note this equalization does not need to be done in real-time, and it is done by the developed software[3] 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 distortion of the sound by the audio system in each iteration is going to 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.

Fig. 3. The spectrogram after 12 iterations in the experiment run inside a wooden dome.



Experiment

Using a condenser microphone, Bluebird-SL, and a loudspeaker, Yamaha HS7, I ran an experiment inside a wooden dome following the same procedure as in *I am sitting in a room*. The initial sound used in this experiment was my own voice. See Figure 1 for the image of the room. The recorded sound at the end of each iteration first went through the equalization filter, and then the gain was adjusted to have the maximum gain across all frequencies equal to one. This process is automated by the custom-made computer code.

Fig. 4. The spectrum of original sound (blue), after the first (dashed red) and 12th (dashed green) iterations in the experiment run inside the wooden dome.

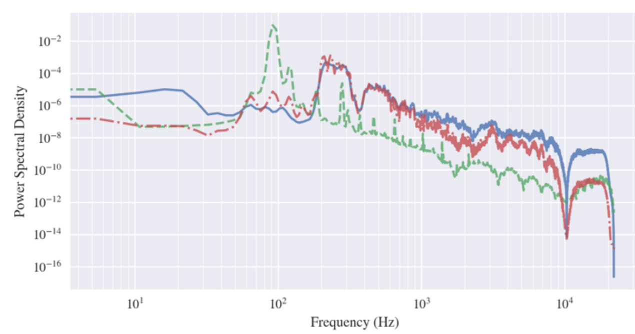


Figure 3 shows the spectrogram of the recorded sound during the first and 12th iterations, in the top and bottom panels respectively. It is apparent that some resonances are present. Figure 4 depicts the spectra of the initial, first, and 12th recorded sound. Notice in the 12th iteration sound, the distance between the three main peaks is about 40 Hz. These peaks are not present in the spectrum of the initial sound.

[3] The filter code in Python can be found at https://github.com/lilac-atassi

The large difference between the frequency response shown in Figure 2 and the spectrum in Figure 4 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.

Conclusion

Many artworks have explored the usage of a room as an audio filter. The microphone and speaker in an audio feedback system act as secondary audio filters. In this paper, I showed that with 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 experiment that I conducted inside a wooden dome supports this claim.

References and Notes

1. A. Lucier, and R. Ashley, *Music 109: Notes on Experimental Music*. (Wesleyan University Press, 2014.)

2. Wolfgang Klippel, “Tutorial: Loudspeaker Nonlinearities Causes, Parameters, Symptoms,” *J. Audio Eng. Soc 54 (10):* pp. 907–939. (2006)

3. Monson H. Hayes, *Schaum´s Outline of Digital Signal Processing. 1st ed.* (New York, NY, USA: McGraw-Hill, Inc., 1998).