Virtual Reconstruction of Spacial Reverberation
Graduate Project Report

Team Signal Wizards

Eric Farmer   Will Carroll
edf63@msstate.edu   woc17@msstate.edu
Leader   Member

December 2, 2018

1   Abstract

The objective of this project is to create a system that artificially reconstructs the reverberative properties of an environment. By capturing the impulse response of the designated space, this system allows the user to simulate the space's acoustics for various inputs. This ability is useful for sound engineers to test audio performance in different environments without the difficulty of a real-world experiment. Additionally, this system could reduce configuration time for public address (PA) systems by allowing the user to identify anomalies in the space's frequency response. This application is beyond the scope of this project, but serves as an example of the potential applications this experiment. For system performance analysis, the output of the artificial space is compared to the captured output of the real-world space for controlled inputs.

2   Introduction

The objective of this project is to design a system to artificially reconstruct the spacial presence of an environment using the environment's impulse response. This system has applications in virtual reality [1] where realistic acoustic feedback is necessary for improving the user experience. Additionally, the system could be applied in live audio performances where the audio engineer must consider the frequency response of the environment. According to [2], proper reverb is crucial to the improved subjective quality of a sound. Designing a system to artificially reconstruct an environment requires capturing the impulse response and artificially reconstructing the sound.

Capturing the Impulse Response of a Space

The impulse response of the environment is critical to the process of reconstructing its spacial presence. Capturing the impulse response is a two-part process — generate an acoustical impulse and measure the response of the space. An acoustical impulse needs to be a nearly instantaneous burst of sound. There are specialized devices to generate an acoustical impulse, but these devices are costly and usually impractical for the identified applications of this system. Based on research presented by the Audio Engineering Society [3], popping a balloon can serve as a replacement for acoustical impulse generation devices in most situations. Other research suggested a person clapping his hands in the space to create the impulse. The balloon pop method is used in this project because of its low cost and ease of repeating performance, which would be difficult if clapping hands.

Simulating Audio Responses of a Space

3 Methodology & Theory

Using Octave, the spacial presence of five locations is artificially reconstructed. Each environment has an infinite impulse response (IIR) that defines the reverberation that occurs in the space. Since this impulse response is infinite in length and cannot be captured by a digital signal processing system, the objective of this experiment is to determine a FIR filter that approximates the impulse response of the IIR filter. An FIR approximation of each space is obtained by inputting a impulse in the space and recording the output using a microphone. Because of the sampling rate and duration of sampling, the FIR may become large compared to usual filters. While a large filter may not be practical for real-time implementation, there are applications in post-processing effects and room analysis that do not require real-time speed. However, a stretch goal is to minimze the filter size without an appreciative loss in quality, as determined by the qualitative analysis, so that the filters may be used in applications such as live performances or virtual reality simulations. The experiment is conducted by measuring the room impulse response (RIR), inputting an audio sample into the room while recording the output, and comparing to a synthesized version of the output from Octave. To provide a variety of reverberation scenarios, the RIR is captured from the following locations:

- Eric Farmer's Living room

- Simrall Hall Anechoic Chamber

- Simrall Hall Stairwell

- MSU SSRC Foyer

- MSU Old Intramural Fields.

The following is an equipment list used for this project:

- Balloons: Nexci Premium Quality 12in. Helium or Air Balloons

- Microphone: dbx RTA-M Reference Microphone

- Speaker: Edifier R1280T Powered Bookshelf Speaker

- Mixer/Preamp: Peavey XR$^®$300 Powered Mixer

- Recording Software: Audacity$^®$2.3.0

- Simulation Software: GNU Octave v4.2.2

The first step to measuring the RIR is to determine the noise floor of the room and the microphone. This is accomplished by observing the output of the room when there is no input. Using identical balloons filled to the same circumference, the balloon pop simulates the impulse input into each space. To prevent uneven spacial effects, the balloon and microphone are colocated in the center of the space. An operator pops the balloon above the microphone using a sharp needle while standing an arm's length away. The microphone records the response of the room. The operator is not colocated with the microphone and the balloon to prevent distortions in the reverbation. The microphone records the impulse response until the room returns to a threshold set by the root mean

square (RMS) of the noise floor. This process determines a large FIR filter for the space and is repeated three times for each location.

After measuring the RIR and the space has returned to its noise floor, each audio input sample is played and the output of the room recorded on the microphone. To prevent spacial distortions, the audio input sample is played through a single speaker colocated under the microphone. Both the speaker and the microphone are positioned such that the focus of the output/input is upward. This process is repeated for each location.

After collecting the RIR for each space, Octave is used to convolve each audio input sample with each spacial impulse response. This provides a set of simulated outputs versus a set of real outputs for each location. Qualitative and quantitative measurements were obtained comparing the simulated and real outputs, which is discussed in the following section.

The stretch goal is performed by using Octave to decimate the filter and calculating the output for each audio input sample at each location. The decimation is performed on the set of decimation factors, $D = [2, 4, 8, 16, 32, 64, 128, 256]$. Due to the amount of samples generated by this process, only the speech samples were qualitatively measured.

4   Analysis of Results

After performing the actions described in the Methodology & Theory section, the results were analyzed quantitatively and qualitatively.

For the quantitative analysis, a mean square error (MSE) was generated to determine the difference between the simulated and the real outputs. Based on plots of the three impulse responses collected at each of the five locations, the third impulse response was selected for each location to be representative of the set. This was done to decrease the processing time and increase the code simplicity. Once the impulse response for each location was selected, the next step was to calculate the MSE between the simulated and real outputs. This was accomplished using `calculate_mse.m` from Appendix A, and a pseudocode of the process is shown in Algorithm 1.

Input: Simulated Sample; Real Sample
Output: MSE between the two samples

1 **foreach** *loc ← locations* **do**
2   sim_sample ← `normalize`(*simulated sample at loc*)
3   real_sample ← `normalize`(*real sample at loc*)
4   Apply Butterworth LPF to sim_sample & real_sample
5   indx_sim ← `find`(sim_sample > `max`(sim_sample))
6   indx_real ← `find`(real_sample > `max`(real_sample))
7   **if** indx_sim < indx_real **then**
8   │ `zeropad`(sim_sample)
9   **end**
10   **else**
11   │ `zeropad`(real_sample)
12   **end**
13   **if** `length`(sim_sample) < `length`(real_sample) **then**
14   │ `truncate`(real_sample)
15   **end**
16   **else**
17   │ `truncate`(sim_sample)
18   **end**
19   `immse`(sim_sample, real_sample)
20 **end**

Algorithm 1: Simulated vs. Real Output MSE Calculation

For the qualitative analysis, a test group of five people was selected to listen to the simulated and real outputs and give an opinion on the similarities of the samples. For the stretch goal, each person listened to the original simulated voice sample and compared it to each decimated voice sample at each location. The qualitative scale was selected to be the following:

- 1 - Nearly identical sounds; minimal distinguishable features

- 0.5 - Similar sounds; the simulated version could pass as a replacement for the real version in some situations

- 0 - Dissimilar sounds; the simulated version fails as a replacement for the real version

Digital Signal Processing

Eric Farmer (EDF63)
Will Carroll (WOC17)

The results of the qualitative and quantitative analysis are shown in Figure 1 and Figure 2. The plot in Figure 3 was designed to determine if a correlation existed between MSE and the qualitative measurements. The plot in Figure 4 shows the correlation between the decimation factor and the qualitative measurements.

	Average Response	*MSE @ Max*	*MSE @ 75%*
Anechoic Chamber	0.5	0.77%	0.86%
	0.6	1.30%	1.29%
	0.4	5.27%	7.96%
Eric's Apartment	0.6	1.39%	1.59%
	0.6	1.31%	1.46%
	0.2	3.54%	3.89%
Intramural Field	0.4	5.16%	5.10%
	0.4	2.22%	2.43%
	0.3	4.73%	5.26%
Simrall Stairwell	0.7	3.11%	3.81%
	0.7	1.96%	1.97%
	0.4	3.89%	4.93%
SSRC Foyer	0.3	1.26%	0.96%
	0.5	1.03%	0.95%
	0.3	6.00%	7.18%

Figure 1: Qualitative & Quantitative Results for Simulated vs. Real Outputs

Digital Signal Processing

Eric Farmer (EDF63)
Will Carroll (WOC17)

	Decimation Factor	Average Response
Anechoic Chamber	2	0.8
	4	0.7
	8	0.6
	16	0.4
	32	0.5
	64	0.4
	128	0.2
	256	0.4

Eric's Apartment	2	0.4
	4	0.1
	8	0.1
	16	0.1
	32	0
	64	0
	128	0
	256	0

Intramural Field	2	0.7
	4	0.8
	8	0.5
	16	0.3
	32	0.4
	64	0.1
	128	0.1
	256	0.4

Simrall Stairwell	2	0.9
	4	0.4
	8	0.1
	16	0.1
	32	0
	64	0
	128	0
	256	0

SSRC Foyer	2	0.5
	4	0.1
	8	0
	16	0
	32	0
	64	0
	128	0
	256	0

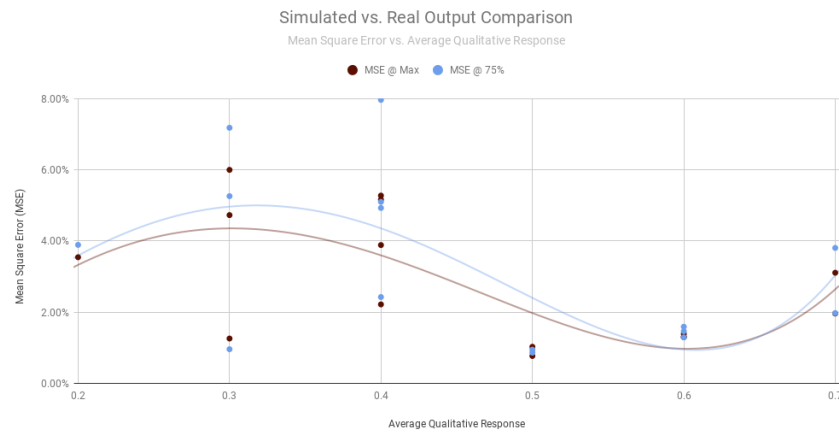Figure 2: Qualitative Results for Decimated Simulated Outputs

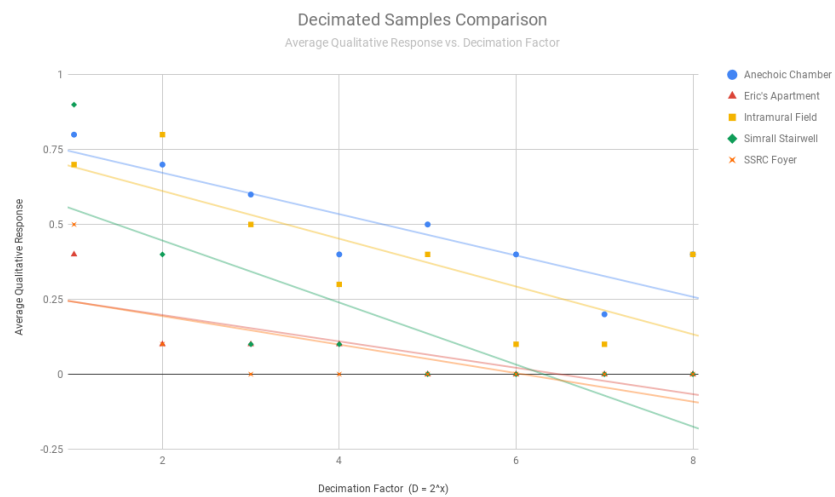Figure 3: MSE Correlation to Qualitative Response

Figure 4: Decimation Factor Correlation to Qualitative Response

5   Conclusion

TODO: Discuss things to improve: - Select the impulse response that generated the minimal MSE from the set for each location. - Discuss microphone/speaker nonlinearities in frequency response

A   Resources

The code and data used in this project is located in the following GitHub repository:

DSP_VirtualSpacialReconstruction

B   Location Photos

References

[1] M. Beig, *Scalable immersive audio for virtual environments*. PhD thesis, 2018.

[2] D. Johnson and H. Lee, "Perceptually optimised virtual room acoustics,"

[3] J. S. Abel, N. J. Bryan, P. P. Huang, M. Kolar, and B. V. Pentcheva, "Estimating room impulse responses from recorded balloon pops," in *Audio Engineering Society Convention 129*, Audio Engineering Society, 2010.