

Implementing an Audio Processing System to Simulate Realistic Distance with Sound

Andy Xu*, Christopher Hariman, Ann Yang, Ankita Inamti, Dr. Martin Jaroszewicz

Monta Vista High School, 21840 McClellan Rd, Cupertino, CA 95014

KEYWORDS. Virtual Reality, Sound Synthesis, Head Related Transfer Function, Spectrogram, Signal Flow

BRIEFS. An audio signal processing system was developed with Python to simulate patterns revealed through analysis of sound recordings on the time and spectral domains.

ABSTRACT. In order to have a more immersive auditory virtual reality experience, a method to better reproduce a sense of distance is required. We recorded sounds at varying distances to observe changes in the intensity and spectral content of moving sounds. By obtaining data from spectrograms, applying different filters, and implementing sound synthesis techniques in the time domain, we created a system that virtually emulates how we perceive those changes. We were able to simulate a sense of distance with synthesized sounds.

INTRODUCTION.

In the last 10 years, the industry has made many technological advancements in Virtual Reality (VR) visuals, but sound in VR is often less appreciated. According to Pavel Zahorik from the University of Wisconsin-Madison [1], most research creating virtual sound environments has focused on direction localization, though in order to have a more immersive experience, a method to reproduce a sense of distance and movement is required. Current methods of positioning sound in space are based on the Head Related Transfer Function (HRTF), which transforms sounds to recreate the filtering that would occur naturally due to the listener's body (ears, shoulders, etc.). However, we noticed that HRTF's measurements only account for equidistant points with varying azimuth (the horizontal angle) and elevation on a spherical plane. We aimed to make VR more realistic by identifying techniques that would allow the listener to perceive distance away from a sound source in a binaural system. In addition, we wanted to simulate a moving sound so that the listener could hear its movement across the spherical coordinates (radius, elevation, and azimuth). To achieve this, we researched psychoacoustics to understand how humans perceive sound. Through conducting our own experiments, we identified relationships between sound and distance. Because we hypothesized that a sound would most likely undergo volume and timbre changes, we researched a variety of filters to see which could best emulate frequency changes of sounds at increasing distances. First, we looked into low and high pass filters, as we predicted that as sounds get further away, either low or higher frequencies would be attenuated. As suggested, these filters would allow lower or higher frequencies to pass and attenuate the rest. To emulate observed frequency peaks and troughs, we also investigated comb filters and the phaser effect. Comb filters output a linear combination of the source

and a delay of the original signal, causing notches in the spectrum (attenuated frequencies) creating a destructive interference. A phaser effect is achieved with multiple allpass filters, whose output preserves all the original's signal energy at all frequencies. Allpass filters can delay specific frequencies at differing amounts, also achieving similar frequency peaks through manipulating the phase relationships between frequencies. Therefore, our research aimed to identify appropriate filters and other sound manipulation techniques that could simulate audio changes at varying distances in a program to improve the VR experience.

MATERIALS AND METHODS.

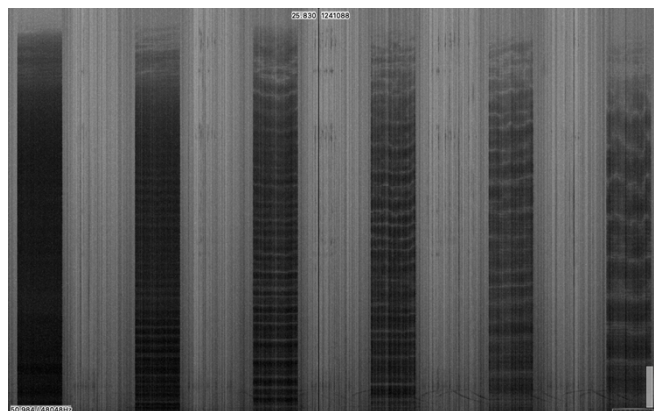


Figure 1. Spectrogram of white noise played from 1, 10, 20, 30, 40, and 50 feet away from a hyper cardioid shotgun microphone.

We analyzed two types of sounds played back with a JBL - LSR 305 speaker in an outdoor environment. These included natural and everyday sounds, and specific sounds that were synthesized for the purpose of our research. We synthesized white noise and a sine sweep using a digital signal processing module for Python. White noise has equal intensity across all frequencies giving a constant power spectral density. The sine sweep plays linearly all frequencies in the humans' audible range at equal levels. We played white noise for 3 seconds with an interval of 5 seconds in between each iteration through a monitor speaker, recorded with a Sennheiser shotgun microphone with a hyper cardioid pattern and a 13db noise level. The first iteration had the microphone 1 foot away from the speaker, and in the 5 seconds between each playing, the microphone was moved in 10 ft increments away from the speaker on the axis. As a result, the white noise was played at distances 1, 10, 20, 30, 40, and 50 feet away from the microphone. We used the same format to gather analysis for the sine sweep. The sine sweep plays through 50Hz to 15kHz in 10 seconds with 4 seconds in between each iteration. As a result, the sine sweep

was played at distances 1, 10, 20, 30, and 40 feet away from the microphone. We analyzed the waveforms and spectra of the two sounds using Sonic Visualizer to identify intensity and spectral changes respectively (Figure 1).

DATA ANALYSIS.

We designed the signal flow and developed a system in Python that simulates decreasing sound intensity and frequency attenuation with a low pass filter (Figure 2). With a given distance and time, a sound source could undergo amplitude changes, a low pass filter, and a phaser effect to reflect natural changes.

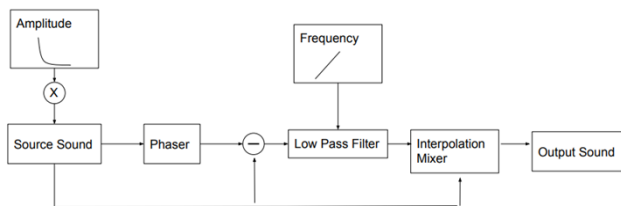


Figure 2. Diagrams of Signal Flow of sound going through our python system

RESULTS AND DISCUSSION.

After plotting our data, we discovered that as the distance between a sound source and a microphone/listener increased, the amplitude of the detected sound decreased nonlinearly. After further researching the physics of sound, we learned that humans' perception of sound decreases logarithmically, explained by the Inverse Square Law, as distance increases. This law states that intensities of light, sound, and other natural phenomena are inversely proportional to the distance from the source squared. Volume of sound is measured in decibels, or dB, which represents logarithmic intensity ratios. We found that as distance increased, the residuals from the regression also increased.

To visualize our data, we mapped our audio recordings in spectrograms, which display frequencies on the y-axis, time on the x-axis, and degrees of amplitude by color (a brighter color indicates more energy). Our spectrograms revealed higher frequencies to be less prevalent compared to lower ones at increasing distances. In the spectrogram above for white noise, the longer distance iterations clearly show attenuated higher frequencies, shown as the white horizontal stripes (Figure 1). At shorter distances, the spectrum shows equal intensities across all frequencies like what we would expect. We found that the highest prevalent frequency in the white noise iterations decreased linearly.

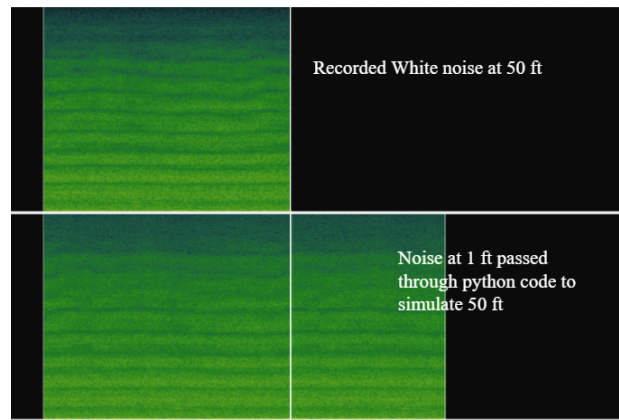


Figure 3. Spectrograms of recorded white noise at 50 feet (top) and filtered noise recorded at 1 foot (bottom)

The white noise spectrograms for longer distances also revealed certain ranges of attenuated frequencies, or holes, in the otherwise equally bright spectrum. At increasing distances, these dampened ranges in the spectrum became wider. In the sine sweep spectrograms, the line representing a linear run through 50Hz-15kHz was not straight, and the sweep appeared to take longer to go through certain ranges of frequencies. This unevenness also demonstrated the ranges of attenuated frequencies.

The two spectrograms in Figure 3 compare the recorded white noise and the programmed white noise. Filters allowed us to replicate similar ranges of attenuated frequencies, shown by the darker lines in the green frequency spectrum.

The findings from our experiment matches with previously published papers. For example, Ignacio Spiouasas's study on the influence of a sound's spectrum on human distance perception also concluded that "[sounds] are judged to be more distant as their high-frequency content decreases relative to the low-frequency content" [2]. With the found relationship between sound and distance, we were able to successfully create an audio processing system, making use of low pass filters and a phaser, that can simulate realistic sound characteristics (Figure 3).

ACKNOWLEDGMENTS.

We'd like to give special thanks to Dr. Martin Jaroszewicz from COSMOS UC Irvine for his invaluable guidance throughout our research.

REFERENCES

1. Zahorik, P. Auditory Display of Sound Source Distance. *Proc. Int. Conf. on Auditory Display*. **2002**.
2. Spiouasas, I.; Etchemendy, E. P.; Eguia, C. M.; Calcagno, R. E.; Abregú, E.; Vergara, O. R. Sound Spectrum Influences Auditory Distance Perception of Sound SOURCES Located in a Room Environment. *Frontiers in Psychology*. **2017**, 8.