Spectral and Frequency Analysis Techniques For Acoustic Space Characterization

Joseph Boesso

Spring, 2016

Thesis submitted in partial completion of Senior Honors Capstone requirements for the DePaul University Honors Program

Dr. Eric Landahl, Physics

Dr. Jafar Saniie, Electrical Engineering (IIT)

Table of Contents

I-Introduction **4**

II-Technical Contents **5**

2.1- Impulse Response Approximation

2.1.1- The Exponential Sine Sweep *5*

2.1.2- Inverse Filter Convolution *7*

2.1.3- Room Impulse Response Approximation *8*

2.2- Time Frequency Analysis **9**

2.2.1- The Short-Time Fourier Transform *10*

2.2.2- The Gabor Transform *11*

III- Methods **12**

3.1 Experimental Approach **12**

3.1.1- Matlab Definitions *12*

3.1.2- Conventional Laptop Measurement 14

3.1.3- Coincident Stereo Microphone Approach *14*

IV- Results/Discussion **15**

4.1 Presentation of Results **15**

4.1.1- Impulse Response Technique Experiment 17

4.1.2- Joint Spectrum Technique Experiment 20

4.2 Discussion of Results **21**

4.2.1- Implications of Impulse Response Technique 21

4.2.2- Implications of Joint Spectrum Technique 22

V- Conclusions **24**

VI- References **26**

VII- Appendix **27**

Abstract – For this thesis, two proprietary estimation techniques were developed, implemented, and tested for acoustic space characterization. The first technique developed was a customized impulse response calculation tool, which could furthermore be implemented in both inverse filtering applications, as well as a reverberation-modeling filter. The second estimation technique involved a joint spectral analysis of a frequency-swept input. This spectral analysis was based on short time Fourier Transform, and gave the relative energies of the response in both the frequency and time domains. This technique could be extended to analyze the response of the room, and make logical deductions on where resonances could occur within the space.

These techniques were developed and simulated using a Matlab environment, and the testing of the technique took place in common household areas. For both estimation techniques it was assumed that no additive noise, unwanted voices, or positioning contributed to the source. In the Impulse Response measurement technique, the testing was all done in mono. In the acoustic space joint estimation, the signal was captured in stereo. The IR calculation supported known behaviors of household rooms, and thus was determined a successful estimation method. In the joint estimation technique, results displayed harmonic resonances at specific frequencies, which could subsequently be attenuated either via signal processing or acoustic treatment. The results of both techniques can be integrated into a variety of environments where a static receiver is considered.

I- Introduction

For Scientists and Engineers alike, it has been a chief concern to try to quantify how signals behave in different environments. This environment could be a physical space such as a room or a theatre, a microscopic object, or even a galaxy. In environments such as these, it can be desirable for the environment to have certain response characteristics- depending on what its' use is. For instance, it could be desirable for a concert hall to have a reverberant response at high frequencies, and a damped response at low frequencies. In a classroom, it could be beneficial to have a dry, even response across the room and amplification from the head of the classroom to the back. In the past it has been possible to use frequency response estimation techniques, along with acoustic analysis, to characterize these acoustic spaces. With this estimation calculation, the user can- through testing- determine whether or not the desired criteria were met or not. However, the response of a room can be thought of as having two different metrics by which it can be quantified: the spectrum of frequencies over time, and the energy of the total signal in time. Though the time and frequency domain representations share their benefits, it is not possible to see exactly what frequencies contributed to a particular instance of the energy signal, and it is not possible to see how much energy belonged to each frequency in the spectrum. For acoustic spaces it was desirable to develop a way that the response could be quantified as a joint spectrum, which would allow the observation of energy concentrations at resonant frequencies, and attenuations in others. Thus, the first task at hand for this project was to find an appropriate method by which the joint spectrum could be utilized to observe the frequency and time responses of an acoustic space.

While computing this estimation technique to obtain the joint spectrum of the acoustic space, it is also possible to use the results in a single-dimension to create an impulse response. The test signal for the space characterization was selected to behave in a way such that, when the recorded response is convolved with a folded version of its' ideal signal, an impulse response approximation is observed. Using this impulse response, and performing an inverse operation, it could then be possible to subtract the environmental effects of the room from the receiver end. This was the second objective of this study, and could further attribute to the power of the estimation technique employed.

II- Technical Contents

2.1 Impulse Response Approximation

In electrical engineering, systems can be thought of as having a transfer function. Transfer functions are a mathematical tool that engineers can use to communicate the behavior of systems in the presence of an input. Furthermore, once a transfer function is obtained numerically, it is possible to perform analytic techniques to alter the behavior of the transfer function given the current description that the function provides. However, measuring a transfer function directly is a difficult and time-consuming process. Thus, most transfer function measurements in the experimental sense involve the use of impulse response estimations. An impulse response is effectively the inverse Fourier transform of the transfer function. Because of this property, the transfer function can be directly obtained from the impulse response. In this section, an experimental method by which the impulse response could be approximated is proposed.

2.1.1-The Exponential Sine Sweep

Classically, an approximate frequency response of a system can be calculated through the measurement of the output of a system in response to an input of varying frequencies. These frequencies can be discretely selected, or swept continuously through a spectrum specified by the user. In a classical sense the measured output would be observed as a linear, time invariant function of time. What this means is that the system produces an output that has both a linear relationship to the input, and has an input/output pair that remains the same when delayed in time. The challenge with experimental measurements of acoustic spaces is that rooms seldom behave as linear, time-invariant systems. Thus, it is quite difficult to obtain a system model of the rooms' frequency response in the classical sense.

To circumvent the issue of obtaining a linear system function from an inherently nonlinear entity, an input had to be selected such that the linear portion of the response could be decomposed completely from the nonlinear portion of the rooms' response. The best way in which a signal could be decomposed such that these portions of the response could be separated was to select a specially behaved input signal. This input signal was a sinusoidal signal with a frequency swept in time such that, when used as an input to a system, would produce an output that could be analyzed to retrieve both nonlinear and linear responses of the system. This well-behaved signal was known as an Exponential Sine Sweep (ESS) signal [2]. The ESS is a special class of a sinusoidal signal, and behaved similarly to a chirp signal. Chirp signals are also sinusoidal functions, but they linearly increase or decrease in frequency over time. This linear change in frequency produced a sound comparable to a birds' chirp. Much like the chirp, the ESS also can either increase or decrease in frequency. Unlike the chirp, however, the ESS had an exponential relationship in frequency change over time. Mathematically, the ESS signal could be observed as the following equation:

$$x(t) = \sin\left(\frac{2\omega_1 T}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \left[e^{\frac{t}{T}\ln\frac{\omega_2}{\omega_1}} - 1 \right] \right) \qquad (1)$$

Where the parameters w_1 and w_2 corresponded to the starting and stopping frequency of the ESS, and the parameter T corresponded to the duration of the sweep (in seconds). The ESS signal was useful for the application of room impulse responses because its' exponential frequency change gave rise to properties which allowed this separation of the linear response from the nonlinear response of the room. As rooms are largely nonlinear and imperfect, it was then desirable to use (1) as an input signal to the system, as additional mathematics could be applied to the measured response to obtain an approximately linear model of the room.

2.1.2-Inverse Filter Convolution

With the input specified as an ESS signal, an operation had to be selected such that when used between the input signal and measured output signal, a linear model of the response was returned. The mathematical operation that was selected to satisfy the desired response was known as a convolution operation.

Classically, a convolution can be observed as the following equation,

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g(t-\tau) d\tau \qquad (2)$$

The convolution equation in (2) can be interpreted intuitively as a measure of the amount of overlap between one function *f(t)* and another function *g(t)* shifted by a certain amount τ. If the two functions overlap greatly, the area is greater and thus the convolution at that point is a larger number. The opposite logic is true for the case where there is little overlap between the two functions. In signal analysis, convolutions are employed to approximate the response of an input signal in the presence of a system. Practically any system therefore can be thought of as a convolution between the input and the modeled system response function. With this in mind, it turned out that the convolution could also be used to approximate the systems' impulse response in respect to the measured output and the ESS input of (1).

The authors of [2] displayed that by flipping the ESS signal about the *y*-axis and using (2) between the flipped ESS input and the measured output, a resulting signal *h(n)* was observed.

$$h(n) = x(-n) * y(n) \qquad (3)$$

In (3), the *x(-n)* variable corresponded to the flipped version of the ESS input. This operation was colloquially referred to as the inverse filter convolution, as the system function was calculated as a convolution between the input and output, rather than the being calculated as a convolution of the input and system function. This operation was made possible because of the properties that the equation in (3) has in the frequency domain. In the frequency domain, when an input signal is reversed in time (made negative), the frequency variable becomes a reciprocal of what its initial value was.

$$x(-n) \Leftrightarrow X\left(\frac{1}{F}\right) \qquad (4)$$

under the property of (4), it was possible to derive what the system function using the measured output and the input. Because the system function was assumed to be of the Finite Impulse Response (FIR) type, the relationship of (4) would yield to a reciprocal of the input, meaning that the overall system function could be expressed as the following,

$$\begin{aligned} &X(F)H(F) = Y(F) \\ &X\left(\frac{1}{F}\right)X(F)H(F) = Y(F)X\left(\frac{1}{F}\right) \qquad (5) \\ &H(F) = X\left(\frac{1}{F}\right)Y(F) \end{aligned}$$

the derivation of (5) illustrated that under the FIR assumption of *X*, it was possible to calculate the system function of *X* via the multiplication of *X(F)* and *Y(F)* in the frequency domain. Converting into the time domain yielded the operation in (3). With the convolution operation specified, the final step of the system design for the impulse response generation was to interpret the result of (3) into a linear time-invariant system.

2.1.3-Room Impulse Response Approximation

The operation of the inverse filter convolution of (3) led to a resulting output *h(n),* which consisted of a sequence of different shaped impulse responses. These impulse responses in the result corresponded to the nonlinear (or reflective) portions of the rooms' response, and the linear portion of the room response. Each response peak corresponded to a different response harmonic of the room, and the overall amplitude of the impulsive signal sufficiently determined whether the impulsive signal was the nonlinear or linear response. Once analyzing the signal, the linear response could be determined as the impulsive signal with the largest overall size in comparison to the other impulsive signals.

To practically obtain the linear portion, a window function was specified to select the region of interest and append this region to a separate function, which was then interpreted as the approximate linear impulse response of the system.

$$\hat{h}(n) = h(n)w(n), \quad w(n) = \begin{cases} 1, n \text{ in linear portion} \\ 0 \text{ otherwise} \end{cases} \quad (6)$$

By windowing the total response, the result in (6) led to an approximation of the linear impulse response of the room. This linear response could therefore be implemented for both filtering and inverse filtering applications.

2.2 Time Frequency Analysis

To gain another dimension in the characterization of the response of a room, a technique known as time frequency (T/F) analysis was implemented on the same ESS test described in the previous section. With time frequency analysis, it was possible to observe both the frequency response and transient response aspects of the acoustic space jointly. T/F analysis in essence is a statistical measure of both the frequency spectrum and energy spectrum in a plot that is referred to as a spectrogram. From the spectrogram, it could then be possible to view the joint distribution of the signal in both frequency and time domains.

2.2.1-The Short Time Fourier Transform

There are many different methodologies available for the purpose of solving for time frequency analysis analytically. For both speed and practical considerations, the method that most optimally fit the task of deriving the joint spectrum of the rooms' response lied in a short-time Fourier transform. The short-time Fourier transform operated identically to a Fourier transform, however rather than computing the Fourier transform of the entire spectrum, a small chunk of the signal of interest is transformed. This is referred to as a short-time Fourier transform because the "windowing" of the signal corresponded to selecting intervals of time by which the Fourier transform was then computed. However, because the signals being used were digital entities, the discrete Fourier transform (DFT) was used in place of the classical analytic Fourier transform. Thus, the discrete STFT was expressed as the following summation.

$$X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n} \qquad (7)$$

The equation of (7) was similar to the classical DFT sum, however a new index *m* is included to signify the variable that shifts the newly added window function to select the portion of the signal to be transformed. The arbitrary function $w[n\text{-}m]$ is the "windowing" function that selects the chunks of time in the input signal $x[n]$. The window function is typically chosen to be a fractional size of the input signal such that when applied, the window function displays a small portion of $x[n]$ and rejects the values of x outside the windows' defined region. It can also be instructive to note that a window of length n and magnitude 1 yields the traditional DFT of $x[n]$.

2.2.2-The Gabor Transform

With the STFT defined in (7) it was necessary to select a window function to best estimate the joint spectrum of the rooms' response in both the time and frequency domains. While there are many window functions to choose from, it was satisfactory to select a window that gave an equal resolution of both the time and frequency domains. The overall way in which the resolution in time and frequency can be assessed is via a metric known colloquially as the uncertainty principle of T/F analysis. This uncertainty principle took a form similar to that of the Heisenberg uncertainty principle in quantum mechanics.

$$\Delta t \cdot \Delta f \geq 1 \qquad (8)$$

The uncertainty principle of (8) showed that the product of the bandwidth in time and frequency, Δt and Δf respectively, could not be smaller than unity. This meant that for joint spectrum analysis technique, the product of the bandwidth and the frequency spectrum resolutions were inherently dependent on one another.

Keeping (8) in mind, the window function that gave equal, optimal resolutions in both time and frequency was a Gaussian window function. The window function was known as a Gaussian function because of its' similar appearance in shape to a traditional Gaussian function.

$$w(n) = e^{-\frac{1}{2}\left(\frac{n-(N-1)/2}{\sigma(N-1)/2}\right)^2} \qquad (9)$$

The window function in (9) could be specified for the overall length of the window N, and the variance (width) of the window σ. The mean of the window was automatically set to be at the middle of the specified length. There were tradeoffs inherent to the length of the window and the variance, and these tradeoffs were empirically determined to provide the best fit for the spectrum.

Using a Gaussian window in the STFT definition, it was then possible to define a new function that could be compactly expressed as a special case of the STFT, known as the Gabor transform.

$$G_x(t,f) = \int_{-\infty}^{\infty} e^{-\pi\sigma(\tau-t)^2} e^{-i2\pi f(\tau)} x(\tau) d\tau \qquad (10)$$

The Gabor transform could be discretely implemented via the application the Gaussian window (9) in the STFT generalized equation of (7). When practically applying the Gabor transform, a second parameter had to be specified to smooth the jump from the sequentially windowed spectra. This smoothing parameter *F* was applied as a universal subtraction from the initial jump index *m* in window function of (7).

Under this Gabor transform, it was then possible to realize the joint spectrum of any signal. To assure the spectrum estimations were accurate, different classical signals could be implemented to assure the technique was properly specified.

III- Methods

3.1 Experimental Approach

Both the impulse response calculation and the spectrum estimation techniques were implemented using the computational software Matlab. Thus, prior to any experimentation both the IR calculation and the spectrum estimation techniques had to be developed in a Matlab environment. The contents of this section will explain in detail how each step of the two programs were developed and implemented to make meaningful measurements.

3.1.1-Matlab Definitions

The input signal that would be used to approximate both the IR response and the joint spectrum was the ESS signal that was specified under (1). This signal was generated using the function specified in (A) of the appendix. This function required a length of the sweep in time, a starting frequency and a stop frequency. From these parameters, the equation (1) was discretely calculated given a sampling rate of 44.1 kHz. The ESS signal generated using the function generated in Matlab produces a sinusoidal signal that behaved as the designed intended.

From the generation of the input signal, two separate scripts were then developed so that the IR calculation and the Time/Frequency spectrum estimation could be computed. The RIR impulse script first operated under the assumption that the input to the measurement was an ESS signal. Following that, the script would take a measured output signal, and perform the calculations specified by (3) to obtain the approximation of the Room Impulse response. The sampling rate was set to 44.1 kHz, however for enhanced speed performance a reduced sampling rate of 16 kHz could be used to perform the operation. The script made use of built-in Matlab functions to approximate both the convolution operation as well as the time-reversed input signal. Once the convolution result was obtained, the script then windowed the function under an empirical investigation to obtain the approximate linear response of the signal. This resulted in the linear impulse response function that was desired.

The final Matlab script computed the Gabor Transform of the same measured output signal for the spectrum estimation. The Gabor Transform script first specified the windows' length, variance, and jump distance. Following this definition, the script calculated the Gaussian window with the desired parameters, and appended it to an original variable and a dummy variable. Following the calculation of the window, the script then received an input signal whose spectrum was to be analyzed. With this input signal, the script multiplied the input signal by the window function and saved the result. With this windowed signal, the script then used the function `fft` to calculate the spectrum of the windowed chunk. After saving the result, the script then shifted the window by the specified amount and performed the same multiplication and `fft` operation. This procedure was repeated until the end of the window reached the last sample of the input signal. Following these computations, it was then possible to observe the spectrum of the signal as a function of time and frequency. This plot was realized using the `surf` Matlab command.

3.1.2-Conventional Laptop Approach

Once all Matlab codes were assured to be working properly using well-behaved controlled signals, it was then possible to begin initial experimentations on the extent to which both the IR calculation and the T/F spectrum analysis technique worked. These techniques were therefore first tested using inexpensive, commonplace testing apparatus. The ESS signal was first broadcasted from the Matlab environment to an external audio device using the built-in `play` function in Matlab. With this, the ESS signal could be broadcast in any point in space. To receive the audio signal, Matlab also used its' `audiorecorder` and `record` commands to use the built-in laptop microphone as the receiver of the generated audio signal. These commands recorded a mono audio signal from Matlab, and stored the result as an object of a special type. This object could then be translated into a data array using a conversion command in Matlab. To approximate a typical situation, the ESS sweep test was done in a living room setting. The source speaker was placed in a position that modeled where a person using the microphone for communication purposes may speak. The program would first initialize the recording object, and then listen to the response as the ESS was played in Matlab. With this measured output signal, it was possible to apply both the RIR calculation and spectrum estimation techniques to the response.

3.1.3-Coincident Stereo Microphone Approach

The second experimental approach was used for a higher resolution spectrum analysis. As the impulse response wasn't as critical to solve in the high-resolution case, this apparatus was reserved for joint spectrum analysis. The apparatus used for this system included two professional grade small diaphragm condenser microphones arranged in a coincident pair. The coincident pair simulated human hearing by placing the two microphone diaphragms as close as possible, and perpendicular to each other.

These microphones were connected to a professional audio interface, and the samples were recorded using audio software. The recorded audio was then imported into Matlab, and both channels were converted into left and right side responses and processed separately. In place of the limited frequency response of the Bluetooth speaker, a hi-fidelity stereo system was utilized to broadcast the ESS signal. In a practical sense this experimental setup corresponded to assessing the joint spectrum of an acoustic space used as a listening room. This scenario was of paramount interest, as the acoustics of a listening room are desired to be as flat as possible. This technique could then assess the extent at which an even frequency response was observed in the room.

IV- Results/Discussions

4.1 Presentation of Results

Using both experimental techniques proved to be a straightforward affair. The first procedural step was to assure the proper behavior of the ESS signal generated using the code specified in the previous section. An example of an ESS signal swept from 5 Hz to 15 kHz can be observed below.

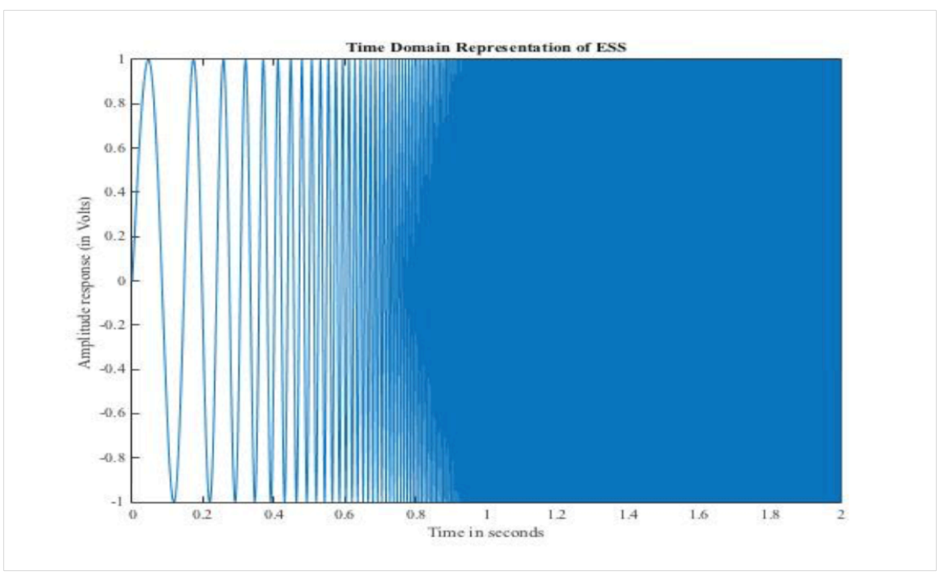

Figure 1: A 2 second ESS wave sampled at 44.1 kHz swept from 5 Hz to 15 kHz.

The power spectrum of the generated signal was also assessed to observe how the power of the ESS signal was distributed across the frequency. The power spectrum was calculated from a simple DFT of the input signal.

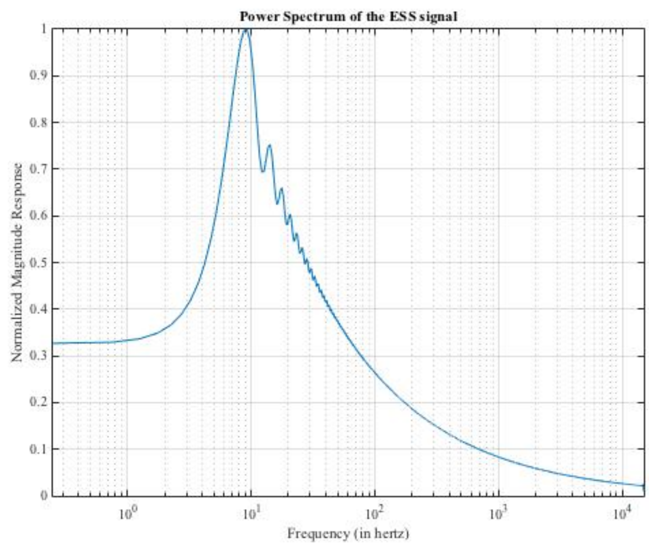

Figure 2: A plot of the power spectrum of the same ESS input signal specified in Figure 1. The overall magnitude of the power was normalized to 1, and the frequency axis (x-axis) was plotted logarithmically.

Having assured that the input signal was well behaved and thus appropriate to use for the experimental procedure, it was necessary to assure the proper operation of spectrum analysis script. The spectrum analyzer was tested using a number of classical input signals. One such example used a chirp signal with a modulation parameter in combination with a cosine wave. Mathematically, the signal could be expressed as the following,

$$x(t) = \sin(2\pi(1000t + \frac{3500 - 1000}{6}t^2) + \cos(2000\pi t) \qquad (10)$$

where the arguments in the sine and cosine terms corresponded to the relative frequency at which both waves were specified at. According to (10), the chirp should start at a frequency of 1000 Hz and end at a frequency of 3500 Hz. The cosine wave should be a constant 1000 Hz. Under these conditions, the analysis technique yielded the following joint spectrum,

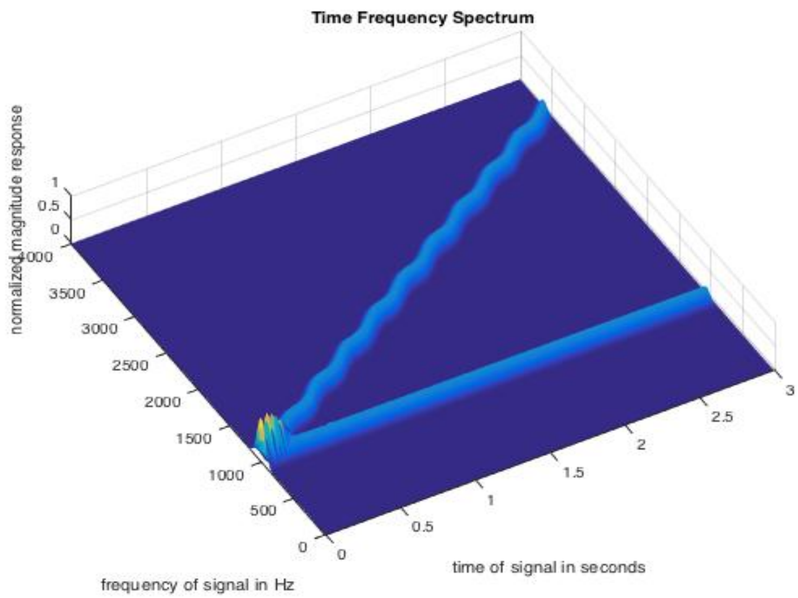

Figure 3: The joint Time/Frequency spectrum of the input signal corresponding to (10). The image shows clearly the desired response of both the modulated chirp and cosine signals.

The result of Figure 3 confirmed that the joint analysis technique behaved experimentally as it was desired to theoretically. With the operation of the joint spectrum analyzer confirmed, it was possible to perform experimental evidence to assess the response of the room using both analysis methods highlighted.

4.1.1-Impulse Response Technique Experiment

The first testing took place using the basic mono laptop microphone apparatus along with the Bluetooth speaker. The testing took place in a common household area that was free from external additive noise. It was important to note that both the receiving and sending apparatus were limited in the extent of their frequency response, and evenness in frequency response. To circumvent this issue, further equalization by way of filtering could be applied to normalize the frequency response to unity across the spectrum of interest. However, for the sake of brevity these imperfections were ignored, as the ESS measurement technique assured a linear response of the room. With this in mind, the following impulse response was approximated via the methodology described for the RIR calculation.

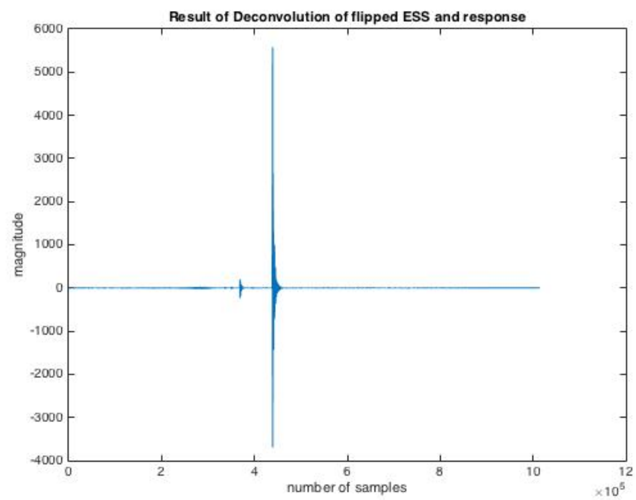

Figure 4: The resulting train of impulse responses as predicted in (3). The magnitude response is purely qualitative and has no significance to the response.

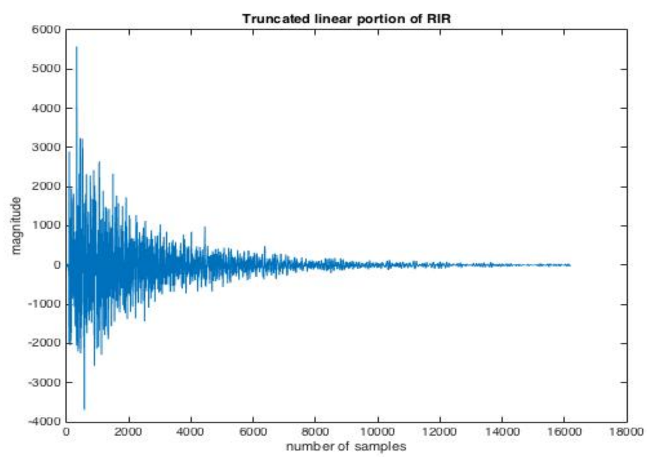

Figure 5: The windowed impulse response to approximate the linear portion of the large rooms' acoustic response. The entire system operated on a sampling rate of 16 kHz.

The room impulse response could further be investigated for a sanity check. In acoustics, it is commonly known that most rooms exhibit something known as comb filtering. Comb filtering is a type of acoustic response that has harmonic points of cancellation across its' frequency response. Therefore, the approximated impulse response in Figure 5 would most likely have a magnitude response that exhibited a comb-filtered trend.

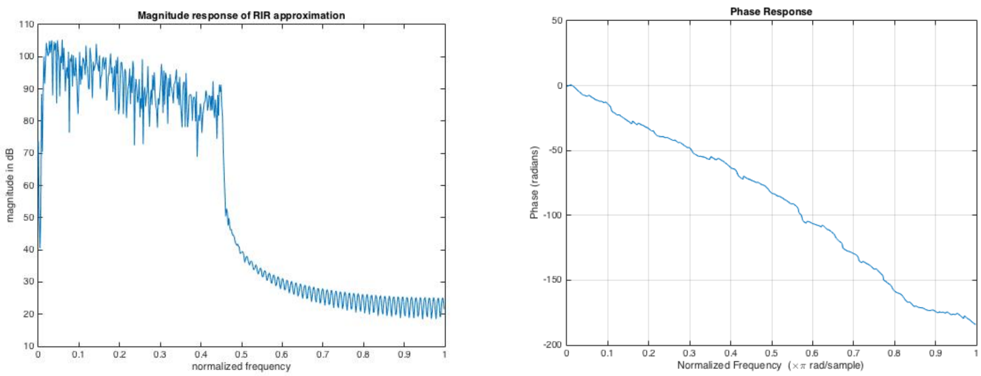

Figure 6: The magnitude and phase responses of the IR approximated in Figures 4 and 5. The magnitude was set to an arbitrary scale, and thus was qualitative. The frequency axis was normalized to half of the sampling frequency; therefore the frequency 1 corresponded to a frequency of 8 kHz .

The resulting filter shape of Figure 6 served as a benchmark by which the impulse response could be qualitatively confirmed as accurate.

For the smaller room, the impulse response and frequency response characteristics were calculated under the same assumptions and experimental procedures of the previous experiment. Under the identical testing conditions, the following results were measured.

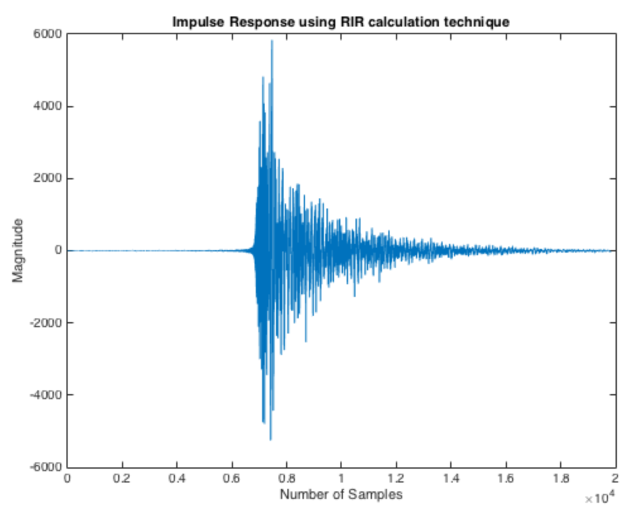

Figure 7: The windowed impulse response to approximate the linear portion of the small rooms' acoustic response. The entire system operated on a sampling rate of 16 kHz.

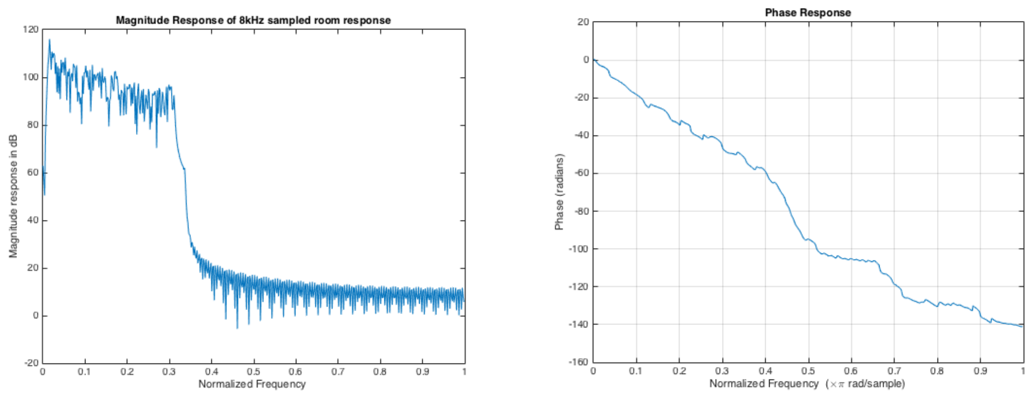

Figure 8: The magnitude and phase responses of the IR approximated in Figures 7. The actual value of the magnitude response was purely qualitative, as it was a function of the magnitude of the recorded input. The frequency axis was normalized to half of the sampling frequency; therefore the frequency 1 corresponded to a frequency of 8 kHz .

4.1.2-Joint Spectrum Technique Experiment

The second experiment utilized the set of stereo microphones arranged in a coincidental pair. The objective of this experiment, unlike the previous example, was to observe the joint spectrum of the response for use in identifying acoustic properties unable to be seen in either the time or frequency analysis alone. Under the prescribed test setup, the following T/F spectra were observed for both the left and right channels of the stereo pair.

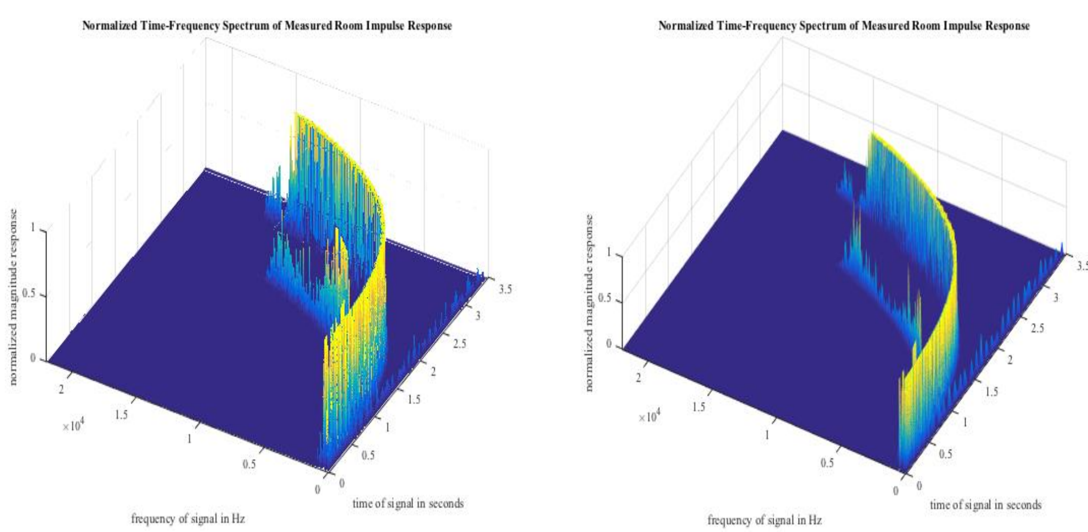

Figure 9: The joint spectrum of the stereo room response to the ESS signal. The spectrum plot on the left corresponded to the left channels' response. The spectrum plot on the left corresponded right channels' response.

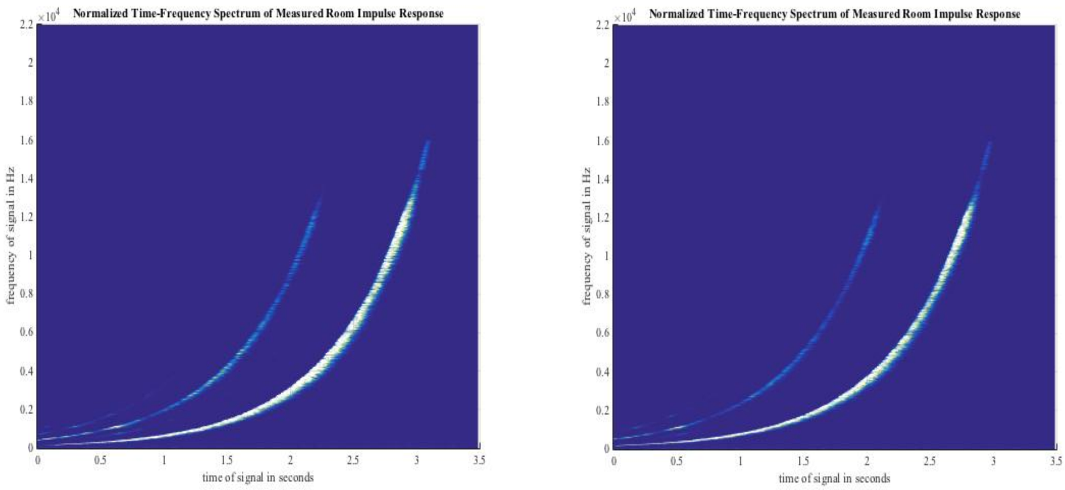

Figure 10: The joint spectrum of the stereo room response to the ESS signal. The order of the plots corresponded to the left and right channels, but the axis was moved to emphasize the artificial harmonics generated.

Between the results of Figure 9 and Figure 10, it was possible to make conclusions on the acoustic response of the room.

4.2 Discussion of Results

4.2.1-Implications of Impulse Response Technique

Under the experimental procedure of the monaural experiment using the simple laptop apparatus, certain observations were made evident. With the resulting impulse response shown in Figure 5 and the magnitude response in Figure 6, it was clear that the impulse response had a somewhat flat behavior across the frequency spectrum, with slight peaks and valleys in certain frequencies. The overall magnitude response also appeared to have an approximately low pass behavior, though this behavior could have been due to the lack of energy in the high frequency bands of the test signal. The frequencies that had slight attenuations appeared to have a harmonic relationship, and thus the magnitude response appeared as a comb shape. This meant that there had been some comb filtering in frequencies at integer multiples and divisions of special frequency wave. The impulse response result of Figure 5 had three distinct portions. The initial spike in the impulse response corresponded to the direct response of the room to the impulsive signal. The portion beyond the initial spike, but not before the tailing signal (around 4000 samples) was the early reflection of the signal. The rest of the IR corresponded to the reverberant part of the rooms' response.

The phase response of the large rooms' IR had an approximately linear phase response, which could be observed by superimposing a line through the phase response plot of Figure 6, and seeing the discrepancies between the ideal line and the measured response. Linear phase corresponded to an approximately FIR filter behavior, and could be easily implemented using DSP.

The second set of results, in comparison to the larger rooms' response, did have a similar trend. The main differences in the small rooms' measured IR lied in the early reflection and reverberation portions of the IR. The IR measured in Figure 8 had an overall larger early reflection portion of the response, but a quicker reverberation tail. This physically corresponded to a smaller room, as smaller rooms are typically more reflective, and less reverberant. This could also be observed in the magnitude response of the small room, as there were more apparent hills and valleys in the magnitude response of the small room in comparison to the large room. The magnitude response of the small room also contained a faster overall attenuation from the highest frequency, and did not ring as severely at the high frequencies as the large rooms' approximated IR. Furthermore, the phase response of the smaller room was not as well behaved as the large room, thus its' consideration as an FIR linear-phase model would be more difficult to consider.

4.2.2-Implications of Joint Spectrum Technique

The joint spectrum estimation results detailed more precise results of the acoustic response of the room studied. As this room was desired to have a majority linear response with good attenuation at harmonic frequencies, the joint spectrum response was capable of assessing where potential standing waves or resonances occurred in the room. In the experiment that generated Figure 9 and 10, the ESS wave was swept from 10 to 20kHz over a time period of 13 seconds. The overall shape of both the left and right channel responses could be interpreted straightforwardly. The linear portion of the response was shown in the spectrum as the yellow (strong) portion of the spectrogram. The main point of interest in the stereo spectrogram was in the harmonic responses. These harmonic responses were the lighter shaded exponential curves above the main linear response in the spectrogram. Figure 10 illustrated that the lighter shaded curve correlated harmonically with the main yellow curve. It could be interpreted as approximately 2 times the fundamental. Figure 10 also exhibited the existence of the 3rd harmonic, though this harmonic was much smaller in magnitude than the fundamental and 2nd. These light curves corresponded to an artificial harmonic generated by the room, and was largely an undesirable effect when considering acoustic properties of a listening room. This artificial resonance tended to cutoff close to the end of the sweep, but in between the start and stop peaks of the artificial harmonic, different resonances occurred between the left and right channels. On the left channel, resonances were observed precisely around 1 kHz and 10 kHz, and loosely between 4 kHz and 6 kHz. The right channel observed an overall harmonic response that was not as great as the left side, but did observe resonances at 1kHz, 10 kHz, and more precise resonances at 4 kHz and 5 kHz. This meant that there was most likely an object, or barrier in the room that caused a standing wave at frequencies that were integer multiples of 1kHz. This corresponded to most likely having an object that was reflective at a distance at integer multiples of the wavelength of the source wave (1kHz). The resulting possible distances that could lead to these resonances were tabulated below.

Frequency (in Hz)	Length (in Feet)
1000	1.13
2000	0.565
3000	0.283
4000	0.141
5000	0.071
6000	0.035

Table 1: The series of frequencies and their corresponding wavelengths, assuming room temperature conditions

From the distances of Table 1, it could then be possible in theory to diagnose potential origins of these resonances in the room by assuming the frequencies were reflected geometrically perfectly off of objects, thus one could measure in the field the distance from the source speakers to various reflective objects that, when translated at a perfect 90 degrees, came into the field of the receivers' pickups. With these distances mapped out, it could then be possible to use absorptive material to remove these resonances from the room. This technique could then be iteratively applied to systematically shape the response of a room to a behavior that is desired by the user.

V- Conclusions

The body of this research documented two different approaches in which a common household room could be characterized using methods and techniques consistent with signal analysis. The first technique described developed a way in which an acoustic space could be statically characterized as an impulse response. This impulse response could then be used for either inverse filtering (removing effects of noise), or for a reverberation effect for music. For the inverse filtering case, one such application could be for use in hearing aid products or video broadcasting purpose. Though the testing used ESS signals of a 2 second time duration, a shorter ESS sweep with a constricted bandwidth could be implemented to model the typical pitch and time durations associated with human conversation. This ESS sweep could be performed iteratively, such that a statistical average of RIRs' could be generated. Once this RIR was generated it could then be possible to subtract the effects of the rooms' response straightforwardly. By result of the systems' definition in the frequency domain, the theoretical signal without the effect of the room could be obtained by dividing the measured output and the impulses responses' system functions. This technique could therefore produce equalization to the receiver end that would provide a cleaner signal to the user. The performed experiments displayed that the RIR could be calculated and modeled to behave as an entity consistent with DSP operation techniques.

In the joint T/F analysis technique, it was possible to further characterize the acoustic space for way of improving the acoustic response. This method succeeded in identifying where resonance frequencies occurred, as well as the existence of standing waves. This technique could apply to a user who was interested in characterizing an acoustic space for a potential audio listening area, where true frequency response and even fidelity across the spectrum is of paramount importance. In closing, the research performed served as an introduction to how these two methods could be implemented both theoretically and practically to characterize the acoustic response of a room.

VI- References

[1] Cohen, L. (1995). *Time-frequency analysis*. Englewood Cliffs, N.J: Prentice Hall PTR.

[2] Culda, T. M., Popa, V., Stanomir, D., & Negrescu, C. (2013). Reducing time in acoustic impulse response measurements using exponential sine sweeps. *International Symposium on Signals, Circuits and Systems ISSCS2013*. Retrieved March 21, 2016.

[3] Demirli, R., & Saniie, J. (2010). Asymmetric Gaussian Chirplet model for ultrasonic echo analysis. *2010 IEEE International Ultrasonics Symposium*. Retrieved March 22, 2016.

[4] Jungmann, J. O., Mazur, R., & Mertins, A. (2014). Joint time-domain reshaping and frequency-domain equalization of room impulse responses. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Retrieved March 21, 2016.

[5] Necsulescu, D., Zhang, W., Weiss, W., & Sasiadek, J. (2009). Room Acoustics Measurement System Design Using Simulation and Experimental Studies. *IEEE Trans. Instrum. Meas. IEEE Transactions on Instrumentation and Measurement, 58*(1), 167-172. Retrieved March 22, 2016.

[6] Sarris, J., & Cambourakis, G. (n.d.). Time frequency analysis and parametric approximation of room impulse responses. *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)*. Retrieved March 21, 2016.

[7] Spa, C., Rey, A., & Hernandez, E. (2015). A GPU Implementation of an Explicit Compact FDTD Algorithm with a Digital Impedance Filter for Room Acoustics Applications. *IEEE/ACM Transactions on Audio, Speech, and Language Processing IEEE/ACM Trans. Audio Speech Lang. Process., 23*(8), 1368-1380. Retrieved March 21, 2016.

VII- Appendix

A) Matlab function to produce an Exponential Sinusoidal Wave for testing purposes

```
function [ x ] = ESS( T,f1,f2 )
%Produces the exponential sine sweep of a signal using
%parameters T=duration of sweep, w1 and w2 are the respective
%frequencies in degrees, t is a time vector
fs=44100;
x=zeros(1,(fs*T));
j=1;
for t=0:(1/fs):T-1/fs
    x(j)=sin(T*2*pi*f1*(exp((t/T)*log(f2/f1))-1)/(log(f2/f1)));
     if (T*f1*(exp((t/T)*log(f2/f1))-1)/(log(f2/f1)))==f2
         break
     end
    j=j+1;
end


% samples per second
   dt = 1/fs;                       % seconds per sample
   StopTime = T;                    % seconds
   t = (0:dt:StopTime-dt);
   N = size(t,1);

 plot(t,x)

 set(gca,'FontName','Times New Roman','FontSize',10);
   xlabel('Time in seconds');
   ylabel('Amplitude response (in Volts) ');
   title('Time Domain Representation of ESS');

   X = fftshift(fft(x));

   dF = fs/N;                        % hertz
  f = fs/2*linspace(-1,1,T*fs);             % hertz

   figure;
   semilogx(f,abs(X/max(X)));
   xlim([0 f2])
   grid ON
 set(gca,'FontName','Times New Roman','FontSize',10);
   xlabel('Frequency (in hertz)');
   ylabel('Normalized Magnitude Response');
   title('Power Spectrum of the ESS signal');

   end
```

B) Matlab script to create a the impulse response approximation

```
x=ESS(10,10,16000);
 x=x';

 impulse=conv(flipud(x),Left);
```

```
impulse=impulse';
j=1;
for index=272810:298000

        impresp(j)=impulse(index);
        j=j+1;
end
```

C) Matlab script to approximate the joint spectrum of a signal x

```
Mwin=6000*3; %Length of window (in samples), MUST BE ODD
Mjump=1000*3; %jump between deltaT window slices
Nfft=64*256; %128 point DFT
var=0.5; %variance for gaussian window
wcenter=(Mwin+1)/2; %define the midpoint of window
% recobj=audiorecorder(8000, 24, 1); % create a voice signal
index=128; %start index variable
% recordblocking(recobj,2000/16000);
% myrecording=getaudiodata(recobj);
% delta=zeros(1,10000);
% delta(500)=1;
%
modsine=(1/40*(cos((3*sin((50*(0:9999))/8000))+(2*pi*2500*(0:9999)/8000
))))+(1/40)*cos(2*pi*3000*(0:9999)/8000); % modulated term
% sine=(1/40)*cos(2*pi*1000*(0:9999)/8000); %sine term
% deltcos=(1/40)*cos(2*pi*1000*(0:9999)/8000)+delta; %sine term with
delta
%
modchirp=(1/40*(cos((0.0001*((0:9999).^2))+(3*sin((50*(0:9999))/8000))+
(2*pi*1000*(0:9999)/8000)))); %chirp with modulation

%x=ESS(2,5,15000);
figure;
x=ESS(2,5,15000);
x=x';
n=0;
gwin=zeros(1,Mwin); %initialize gaussian window
f=0:(44100)/length(x):((44100)-(44100)/length(x)); %initialize
frequency vector
while index <=length(x)
clear gwin; %reset gaussian window every iteration
gwin=zeros(length(x),1);
shift=n*Mjump;
for win=1:Mwin+1
if win+shift>=length(x)
break
end
gwin(win+shift)=exp(-.5*(((win-(Mwin/2))/(var*(Mwin/2)))^2));
end
sigtrans=gwin.*x; %window the signal chunk
sigfft=fft(sigtrans); %fft the window chunk
hideal=[ones(1,length(x)/2) zeros(1,length(x)/2)]; %ideally filter
signal to remove aliasing
spec(:,n+1)=sigfft.*hideal'; %populate spectrogram
n=n+1;
index=shift; %Update index variable according to shift
end
```

```
t=0:3.5/n:(3.5-3.5/n);
surf(t',f',(1/20)*(spec.*conj(spec)),'Edgecolor','none')
 set(gca,'FontName','Times New Roman','FontSize',10);
xlabel('time of signal in seconds')
ylabel('frequency of signal in Hz')
zlabel('normalized magnitude response');
title('Normalized Time-Frequency Spectrum of Measured Room Impulse
Response');
ylim([0 22000]);
zlim([0 1]);
caxis([0 1])
```