Parallel Second-order Filter-based Equalizer Design for Digital Room Correction

Final project for EE522: Immersive Audio Processing

1. Introduction

Room Acoustics matters. Producers in studio need to monitor every subtle detail in music. Every ordinary people wants excellent sound quality for their playback system to enjoy movies, music, TV broadcast, etc. However, if the room where the sound system is installed gives undesired acoustical effect on the audio, these wishes could never come true, no matter how much money you invest in your electronic playback system. Bad room acoustics is a common problem in almost every realistic room. That is why we need extra calibration technologies to make up for the acoustical effect, or to put it more specifically, the frequency response of the room.

Digital Room Correction is quite a new area of study. In the past, analog filters are used to correct the frequency response of rooms. However, due to the processing limitation of analog filters, it is not practical to use them in real rooms, which are typically modeled as relatively high-order systems. With the improvement of the computational capacity of CPUs and DSPs, it is possible to apply digital filters to real rooms. Typically, the correction is implemented by a digital filter (equalizer) which is applied to a loud speaker of the sound system in a room. At a listening location, the impulse response of the room, along with the processing of the equalizer will give an overall acoustic response, which is close to the predetermined target curve, so that the unfavorable effects of the room's response can be ameliorated.

In this project, a Parallel second-order-based equalizer design method is thoroughly analyzed and implemented for loudspeaker-room correction problem. This methodology is introduced in section 2. More implementation details are shown in section 3. Results for the equalizer design and a listening experiment are shown in section 4. Some discussions about future research are shown in section 5.

2. Methodology

A naïve way of designing the equalizer is to simply reverse the frequency response of the room. But in reality, filter inversing may cause many undesired results, such as non-causality, instability, etc.

Alternatively, there are several other ways to design the equalizer filter. For example, FIR filter, IIR filter, parametric filter, frequency-warped filter (FIR and IIR), Kautz filter, etc. [1,2,3]. Among these filters, Kautz filter has the most robust equalization compared to others [3]. This project will use a designing method called Fixed-Pole Parallel Second-Order Filters, which yields same equalization as Kautz filter for the same order but will be 33% more computationally efficient than Kautz filter [4]. The outline of this method is as follows:

2.1 Theory of Parallel Second-order Filter

Any transfer function of form $H(z) = \frac{B(z)}{A(z)}$ can be rewritten in the form of

$$H(z) = \sum_{i=1}^{P} c_i \frac{1}{1 - p_i z^{-1}} + \sum_{m=0}^{M} b_m z^{-m} \tag{1}$$

where $p_i's$ are poles. If the system has a real-valued impulse response, then the poles will appear in conjugate pairs. To design the filter, we can fix these poles according to logarithmic frequency resolution [5]. This will be illustrated in a more detailed manner in the following sections.

We can combine the complex pole pairs to a common denominator. This results in second order sections with real-valued coefficients, which can be implemented more efficiently. The transfer function becomes

$$H(z) = \sum_{k=1}^{K} \frac{d_{k,0} + d_{k,1} z^{-1}}{1 + a_{k,1} z^{-1} + a_{k,2} z^{-2}} + \sum_{m=0}^{M} b_m z^{-m} \tag{2}$$

Figure 1 shows the structure of the parallel second-order filter. Since the poles of the IIR filter are determined, remaining free parameters to be solved are $d_{k,0}, d_{k,1}, b_m$. It is simpler to find the coefficients in time domain. The impulse response of the parallel signal is given by

$$h(n) = \sum_{k=1}^{K} d_{k,0} u_k(n) + d_{k,1} u_k(n-1) + \sum_{m=0}^{M} b_m \delta(n-m) \tag{3}$$

where $u_k(n)$ is the impulse response of the transfer function $1/1 + a_{k,1} z^{-1} + a_{k,2} z^{-2}$. Rewriting (3) in matrix form:

$$\boldsymbol{h} = \boldsymbol{Mp} \tag{4}$$

where $\boldsymbol{p} = [d_{1,0}, d_{1,1}, \dots, d_{K,0}, d_{K,1}, b_0 \dots b_m]^T$, $\boldsymbol{M}$ contains the modeling signals $u_k(n)$ and its delayed counterparts $u_k(n-1)$ and unit impulse $\delta(n)$ and its delayed counterparts, $\boldsymbol{h} = [h(0), h(1) \dots h(N)]^T$, which is the target impulse response.

(4) could be solved by Mean Squared Error technique, yielding the optimal parameters

$$\boldsymbol{p_{opt}} = (\boldsymbol{M^H M})^{-1} \boldsymbol{M^H h} \tag{5}$$

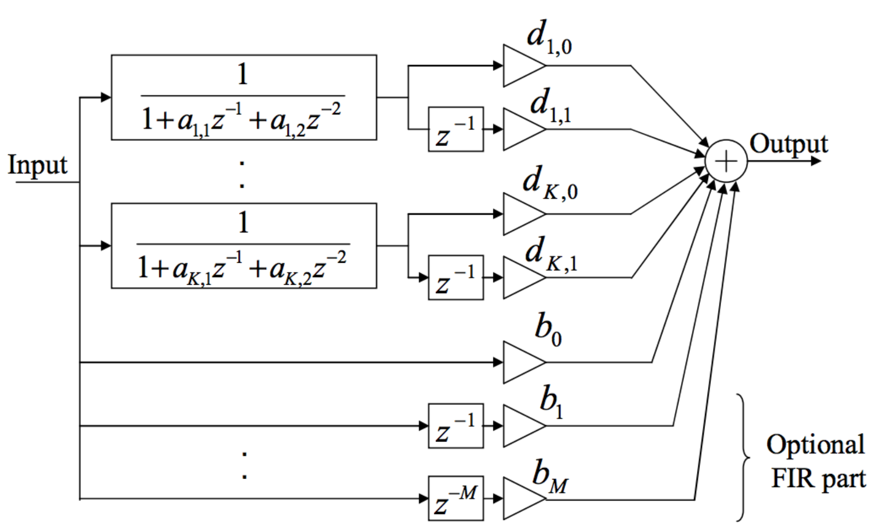

Figure 1 Structure of parallel second-order filter

2.2 Direct Equalizer Design

Now, we can adopt a direct equalizer design procedure to avoid filter inversion. Suppose the impulse response of the equalizer is $h_{eq}(n)$, the input of the equalizer is the room response, denoted by $h_s(n)$, and we desire that the output $h(n)$ match the target response $h_t(n)$. $h(n)$ is computed as

$$h(n) = h_{eq}(n) * h_s(n)$$
$$= \sum_{k=1}^{K} d_{k,0} u_k(n) * h_s(n) + d_{k,1} u_k(n-1) * h_s(n)$$
$$+ \sum_{m=0}^{M} b_m \delta(n-m) * h_s(n) = \sum_{k=1}^{K} d_{k,0} s_k(n) + d_{k,1} s_k(n-1)$$
$$+ \sum_{m=0}^{M} b_m h_s(n-m) \quad (6)$$

where the signal $s_k(n) = u_k(n) * h_s(n)$ is the room response filtered by transfer function $1/1 + a_{k,1}z^{-1} + a_{k,2}z^{-2}$. The above equation has same structure as (3). Similarly, rewriting (6) in matrix form:

$$\boldsymbol{h} = \boldsymbol{M_{eq}} \boldsymbol{p} \quad (7)$$

where $\boldsymbol{M_{eq}}$ contains the new modeling signals $s_k(n)$, its delayed counterparts $s_k(n-1)$, the room response $h_s(n)$ and its delayed counterpart. Solving for optimal $\boldsymbol{p_{opt}}$, yielding

$$\boldsymbol{p_{opt}} = (\boldsymbol{M_{eq}}^{\boldsymbol{H}} \boldsymbol{M_{eq}})^{-\boldsymbol{1}} \boldsymbol{M_{eq}}^{\boldsymbol{H}} \boldsymbol{h_t} \quad (8)$$

From the above illustration, we know that by setting a fixed, predetermined pole set, what we are doing is to solve for the weights parameters for the second-order filter sections. And this could be solved in a closed, algebraical way. Therefore, this design method is very simple and robust. All we need to set manually is the pole set.

2.3 Pole Positioning Strategy

In reference [5] and [6], the author discussed and compared the details of different poles positioning strategies. In this project, a predetermined set of poles will be used. This is advantageous because we can get adaptive frequency resolution by choosing pole. For example, by fixing the poles in a logarithmic frequency scale, the design will have a logarithmic frequency resolution.

For example, we have decided to set pole sequence $\{f_k\}, k = 1,2 \ldots K$. Then the normalized frequencies of the poles in radian, θ_k, can be computed as follows:

$$\theta_k = \frac{2\pi f_k}{f_s} \quad (9a)$$

where f_s denotes sampling frequency. The poles can be computed as follow:

$$p_k = e^{-\frac{-\Delta\theta_k}{2}} e^{\pm j\theta_k} \quad (9b)$$

where $\Delta\theta_k$ denotes the bandwidth of the kth second-order section, which came from following formulas:

$$\Delta\theta_k = \frac{\theta_{k+1} - \theta_{k-1}}{2} \qquad for\ k = 2,3 \dots K-1$$
$$\Delta\theta_1 = \theta_2 - \theta_1$$
$$\Delta\theta_K = \theta_K - \theta_{K-1} \qquad (10)$$

Formula (9b) sets the radii of the pole p_k in such a way that the transfer functions of the parallel parts pass approximately at their -3dB points.

In this project, I set 12 frequencies as the pole frequencies in Logarithmic scale in low frequency range (30 – 200Hz) in order to have higher resolution at this range and detect the dips caused by Normal Modes. And we place 13 frequencies as the pole frequencies in Logarithmic scale in high frequency range (250 – 18000Hz). This leads to a total order of 50 for the parallel filter.

The advantage of this pole positioning is that it simplifies design process. All we have to do is to solve the Mean Square Error equations for the weight coefficients and we don't have to do prior IIR filter design. Besides, it does not require pre-smoothing of the transfer function curve. By this pole positioning strategy, we can solve the MSE equations directly from the original impulse response and the target. More details about this kind of positioning can be found in [5] and [6].

3. Implementation

The diagram of the equalizer system is shown in Figure 2.

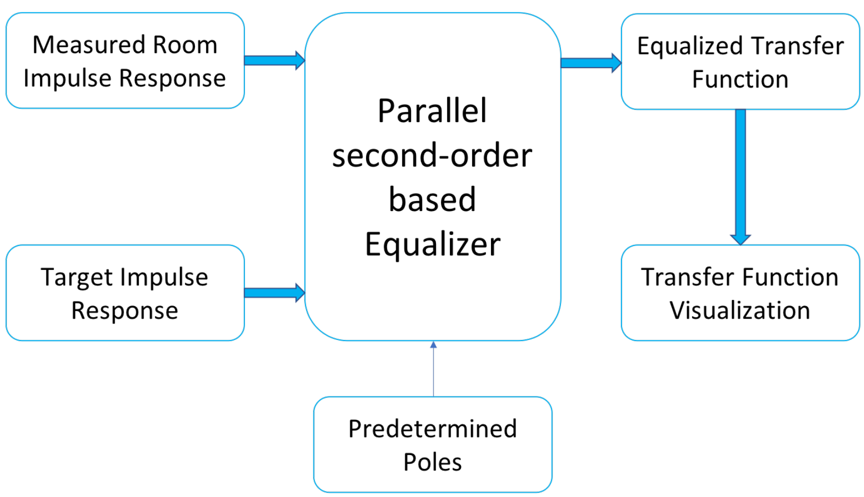

Figure 2 System Diagram

As is discussed in the previous section, the equalizer takes inputs as measured room impulse response and target impulse response, and generate an equalized transfer function, which is close to the target frequency response in frequency domain. Then the equalized transfer function is further post-processed for visual evaluation. For the equalizer part, we already know from section 2 that all we need to care about is determining the poles, which has been implemented in the way described previously. Hence, in the implementation of the whole system, what are left are Measuring room impulse response, Setting desired target impulse response and Visualizing the equalized transfer function.

3.1 Room Impulse Response Data Acquisition

In this project, two rooms are used as the objects to be corrected.

The first one comes from the Aachen Impulse Response (AIR) database provided by Stefan Kühl et al. from RWTH Aachen University [7]. In their dataset, properties of multiple rooms, including meeting room, lecture room, office, a big church, etc. are recorded. Considering the reverberation time should be appropriate for the listening test in this project, I used the office as the target room to be corrected. The basic information about this room is shown in Table 1

Room dimensions	5.00m×6.40m×2.90m
h_L	1.2m
h_M	1.2m
h_{LM}	1.0m
Wall surface	Glass windows, concrete
Floor cover	Carpet
Furniture	Wooden desk, shelves, chairs
Reverberation Time (RT_{60})	0.37 s

Table 1 Basic information of the office

where h_L denotes loudspeaker height, h_M denotes microphone height, h_{LM} denotes the distance between loudspeaker and microphone. The measurement sound source and device information are shown in Table 2.

Loudspeaker	Genelec 8130
Sound source	Maximum Length Sequence
Microphone	Beyerdynamic MM1(Omnidirectional)

Table 2 Sound source and device information for measurement for office

The other room used in this project is from a sound studio of the Laboratory of Music Acoustics Technology (LabMAT) at the Department of Music Studies of the University of Athens [8]. The studio consists of three main spaces, the Mixing Suite, the Control Room, and the Live Room. Here I am using the Live Room as our room to be corrected. The live room have some rotating acoustic panels with rail-system installed, as shown in Figure 3. According to the dataset website [8], the room size is 397 m^3 and the reverberation time is 0.59s. Sound source and device information is shown in Table 3.

Figure 3 The live room

The locations of loudspeaker and microphone are shown in Figure 4. In our case, we are using the measurement configuration where loudspeaker is set at location “S” and the microphone is at location “R8”, which gives an appropriate reverberation time.

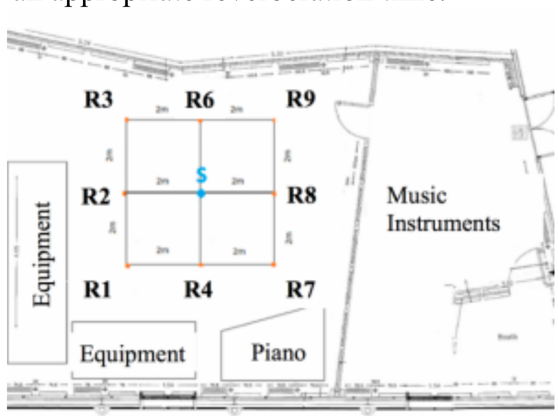

Figure 4 Locations of loudspeaker and microphone for live room

Loudspeaker	Brüel & Kjær Type 4292-L
Sound source	Swept Sine; 22Hz - 22kHz; 15 Seconds
Microphone	Behringer ECM-8000 (Omnidirectional)

Table 3 Sound source and device information for live room

The purpose of using this live room is to show that our equalizer remains powerful regarding some extreme acoustical environments, as will be shown in the next section.

3.2 Target Response design

The parallel second-order based equalizer method use MSE technique to match the input impulse response to desired target impulse response directly. So theoretically you can set the target curve with arbitrary magnitude. In this project, three types of target curves are used. They are shown in Figure 5, 6 and 7. To design the target frequency response with any shape, I used the `fdesign.arbmag` method. Through this method, you can design an IIR or FIR filter with any fixed frequency vector and magnitude vector. This method will return the numerator and denominator coefficients b and a. You can check with `fvtool` method to make sure that the target curves satisfy your expectation.

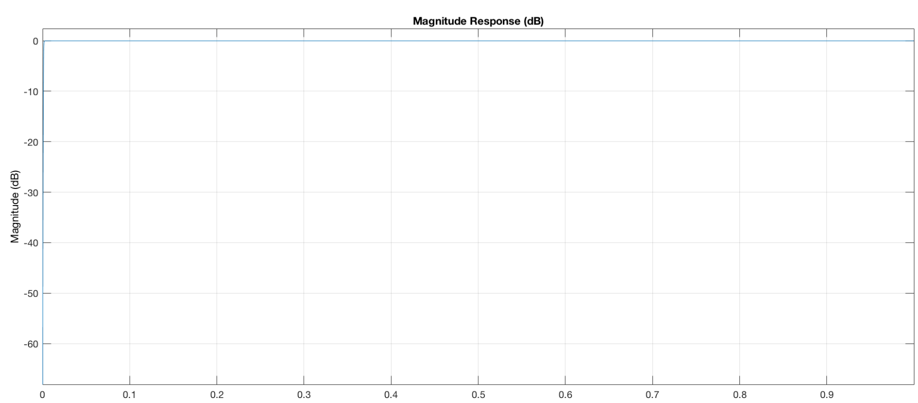

Figure 5 Target curve with flat response over frequency range of 30Hz to 20kHz

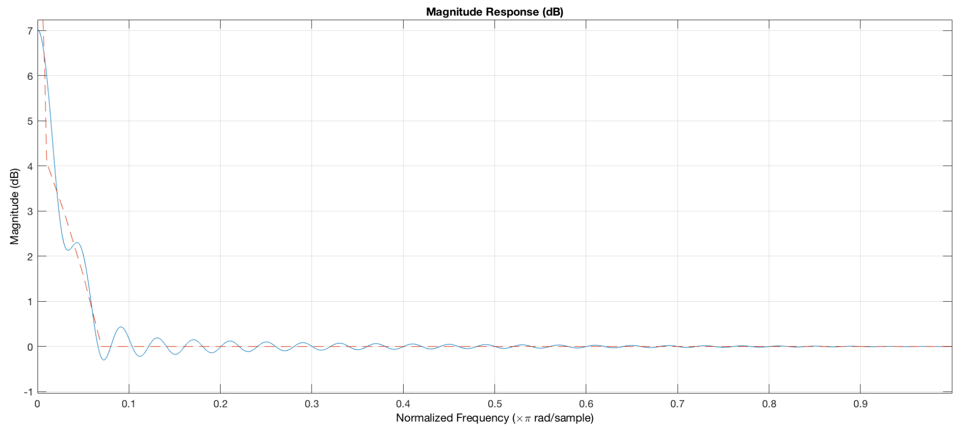

Figure 6 Target curve with low frequency boosted over range of 30Hz to 1kHz

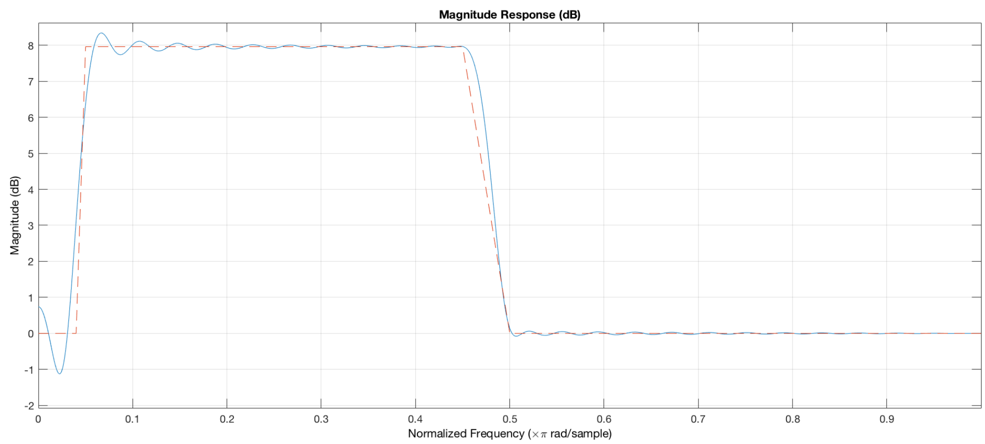

Figure 7 Target curve with middle frequency boosted over range of 1kHz to 10kHz

3.3 Equalized Transfer Function Visualization

Now we have come to the final part of the whole loudspeaker-room correction system. After previous steps, we have got the output, i.e. the equalized frequency response. However, direct analysis of measured and equalized room frequency responses can often lead to unclear results, especially with respect to the perceived effect that the room's reverberation properties will have on any listener [9]. A typical operation to extract simplified while useful information from these responses is to apply fractional-octave smoothing. This smoothing operation is advantageous for evaluating the useful spectral properties of the room, while suppressing the undesirable properties that might affect our observation. Audio engineering practice and later results have shown that such a processing artifact is acceptable for the case of acoustic/electro-acoustic/audio system responses, since at one hand allows a clearer visual interpretation of the measured data and at the other hand is in agreement to the sensitivity pattern of human auditory mechanism [10].

In this project, 1/3 octave smoothing to the spectral magnitude is applied, as shown in Figure 8. It was found that after smoothing, the late reflection energy is largely suppressed, clarity is increased, spectral mean and deviation is lowered, especially at mid and high frequencies [9]. Mathematically, this is done through convolving the original magnitude of the transfer function with a window having low-pass property in frequency domain. In this project, we are using a Hanning window, which has the length of 1/3 octave. For discrete signal sequences, this process can be described as formula (11)

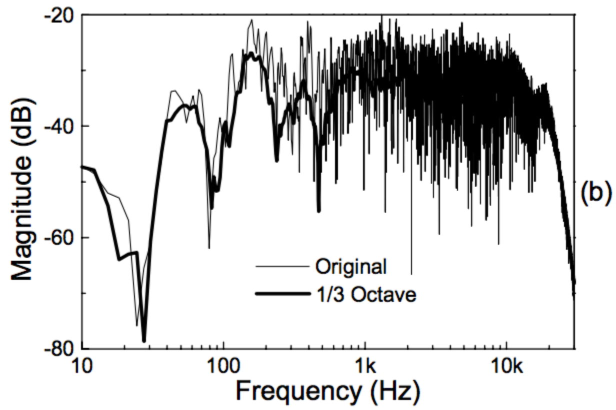

Figure 8 Effect of 1/3 octave smoothing

$$|H_{sm}(k)| = |H(k)| * W_{sm}(k) = \sum_{i=0}^{N-1} |H((k-i) \bmod N)| \times W_{sm}(i) \qquad (11)$$

4. Results and Discussion

4.1 The equalized frequency response

For the Office, the unequalized frequency response, equalized frequency response and the equalizer frequency response, and their smoothed counterparts are shown in Figure 9.

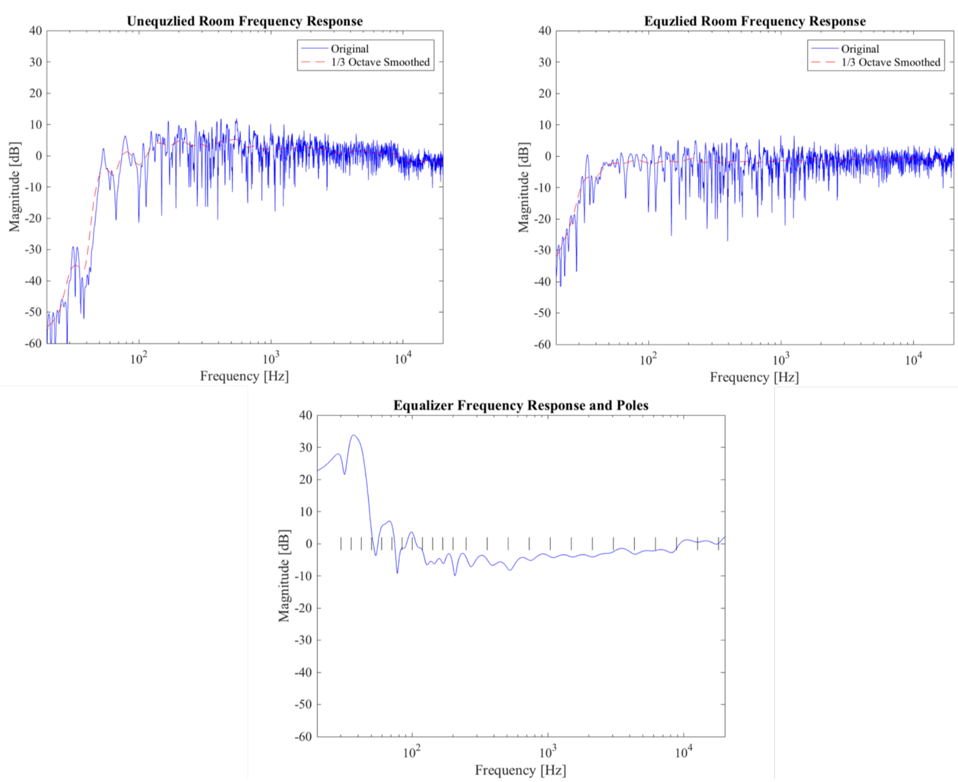

Figure 9 Frequency response correction for office

From Figure 9, we can see that the equalizer successfully corrects the frequency response curve to a nearly-flat curve, which satisfies our expectation. Figure 10 and 11 show the calibration cases for bass boost and mid-frequency boost target curve. In the same way, the equalizer works well.

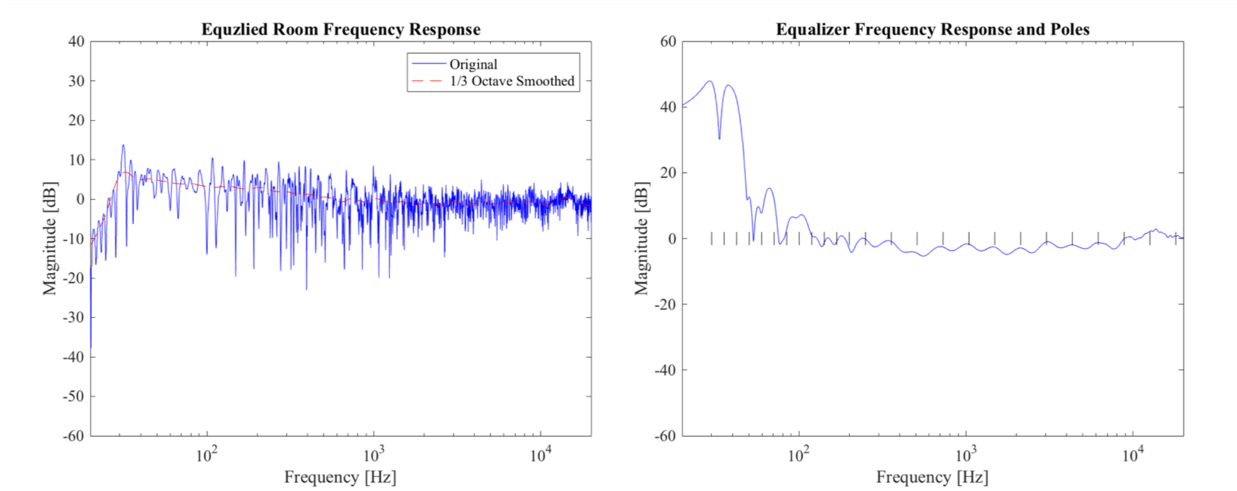

Figure 10 Bass boost equalizer for office

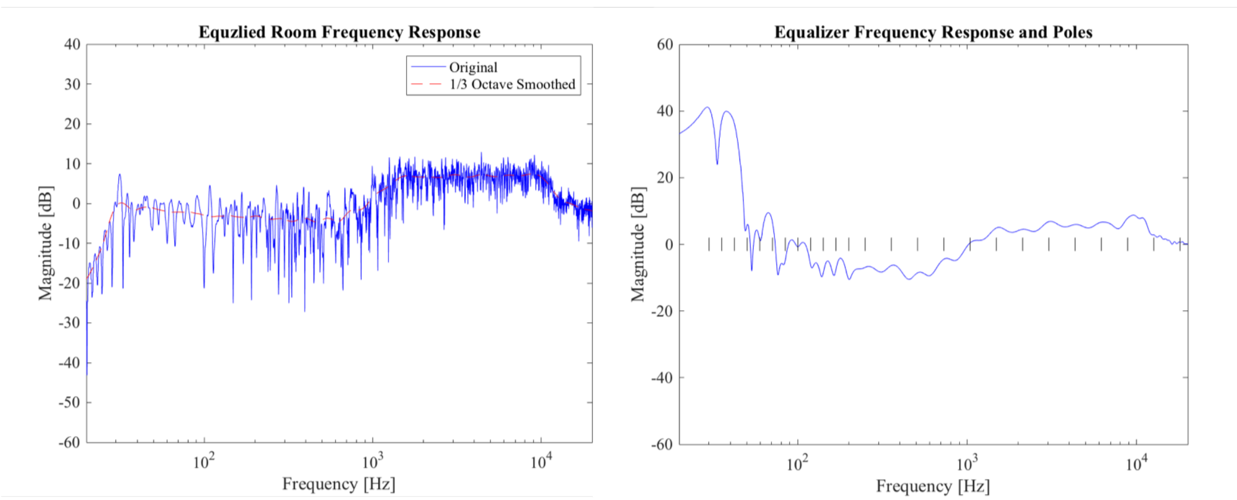

Figure 11 Mid-frequency boost equalizer for office

To verify the processing power generalization ability, the equalizer is applied to a special acoustical room, the live room described in section 3.1. As shown in Figure 12, the original frequency response of this room is far from the flat target curve. It proves the equalizer's ability to correct the frequency response in extreme cases.

4.2 Listening preference test

Now we see that the equalizer is quite powerful for the frequency responses correction. But how will they work for real audios? In this part, I used two music genre, Classic Music and Electro-Dance Music (EDM). They are convolved with the unequalized room responses, equalized room responses with flat target, equalized room responses with bass or mid-frequency boosted target of the office and the live room, respectively. Then 21 people are asked to give their preference to these resulting audios.

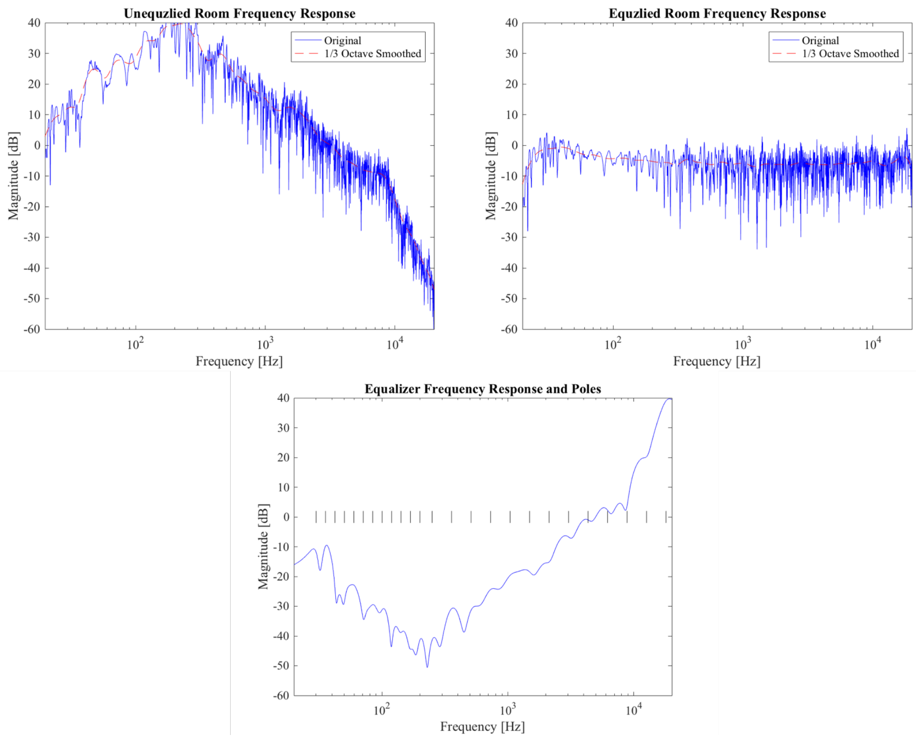

Figure 12 Frequency response correction for live room

The result for the live room, using two types of music is shown in Figure 13.

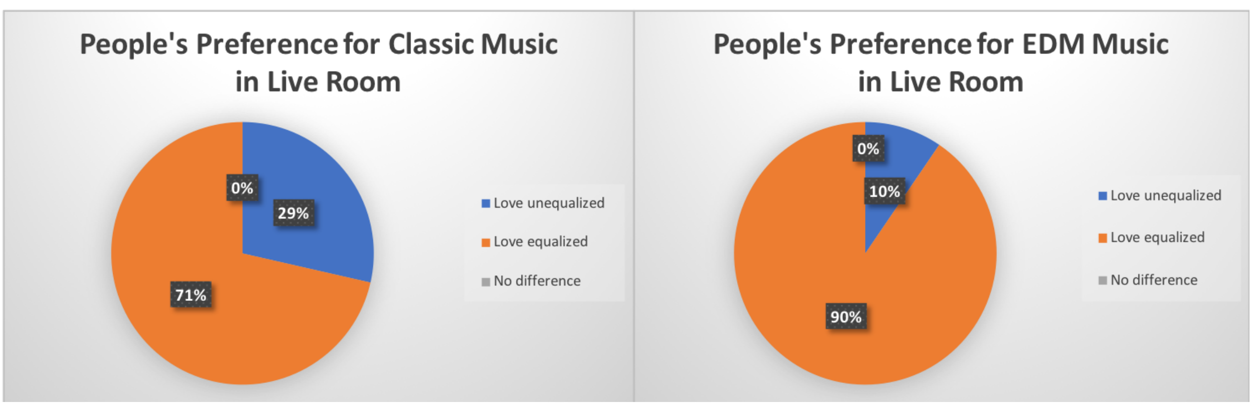

Figure 13 Listening test in live room

Since the live room has a relatively extreme acoustical property, both types sound terrible in this room. From Figure 13 we can see that people have a preference with a large margin for the audio after equalization, proving that our calibration equalizer improve people's listening experience.

For the office room, the test results using Classic and EDM Music are shown in Figure 14 and 15 respectively. First, I carried out the question on asking people whether they heard the difference of unequalized and equalized frequency response and give their preference. The result is at the left-hand side. And then I added a new target for the equalizer (treble-boost or bass-boost) and asked them to give their preference among the unequalized, flat-target equalized and the new-target equalized frequency response, the result is at right-hand side. From Figure 14, we can see that for Classic Music, most people did not perceive difference between the unequalized and equalized frequency response. This is because what the equalizer did is compensating the dips in low frequency range and did not modify much in middle and high frequency. Since Classic Music has rich frequencies in middle and high frequency, while tend to lack some low frequency components, the correction might have little effect on Classic Music.

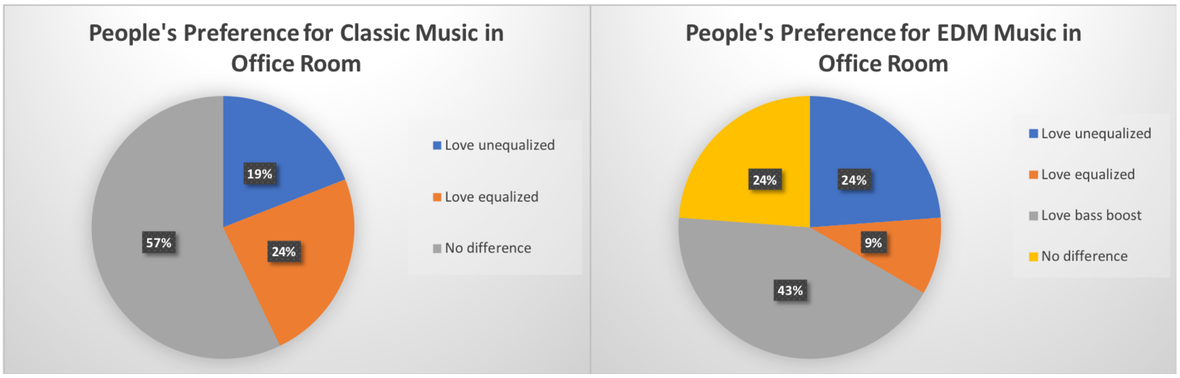

Figure 14 Listening test in office room (Classic Music)

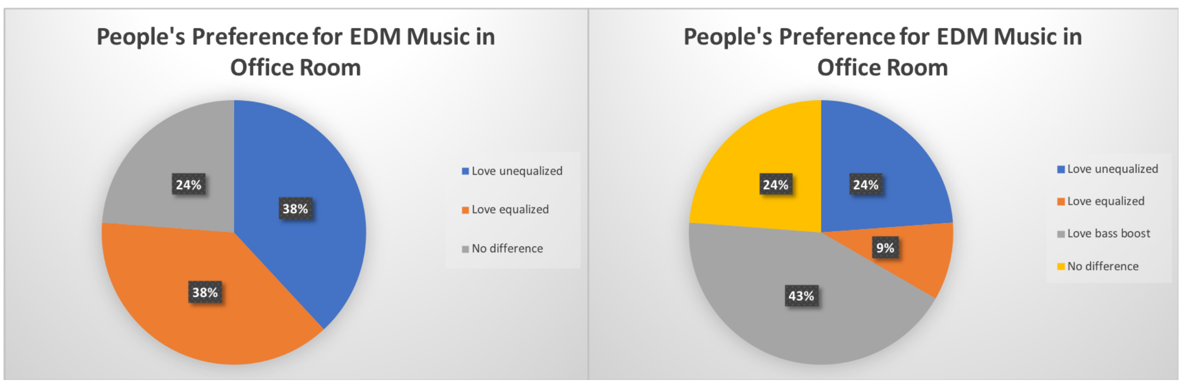

Figure 15 Listening test in office room (EDM Music)

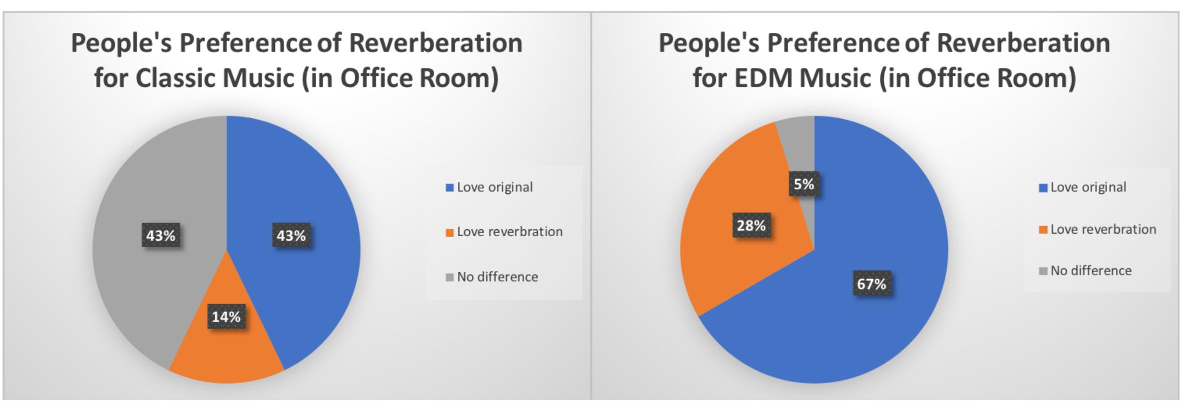

Figure 16 People's preference for reverberation of two types of music

As shown in Figure 15, we can see that for EDM Music, 20% more people can perceive the difference between unequalized frequency response and equalized frequency response. But what is hurtful is that people do not show a prominent preference for the corrected flattened frequency response. Besides, for EDM Music, people showed a significant preference for the bass-boosted frequency response, as shown in the right side of Figure 15.

At last, a question regarding people's preference for reverberation is carried out. People are asked to compare the very original audio and the audio which is equalized. The main difference here is that the equalized audio contains reverberation generated by the room. From Figure 16 we can see that reverberation is not a significant factor affecting people's experience with Classic Music, while for EDM Music, people showed a prominent preference for the audio without reverberation. Hence, we can say that reverberation is a music genre-based factor for music experience. For a room for EDM Music purpose, apart from the frequency response equalization, dereverberation techniques should also be applied to improve listening experience.

5. Future Direction

- **Dereverberation:** As shown in the previous section, reverberation can be a significant factor regarding some music genre. So, in order to further improve the Room Correction effect, additional dereverberation algorithms should be applied.
- **Car Audio Quality Improvement:** Apart from spaces like rooms, the equalizer can also be used in spaces like cars. However, in the case of car, noise can be a more significant factor. Noise cancellation algorithms can be added in the processing system.

6. References

[1] S. Bharitkar and C. Kyriakakis, "Loudspeaker and room response modeling with psychoacoustic warping, linear prediction, and parametric filters," in Proc. 108th AES Conv., Preprint No. 6982, San Francisco, CA, USA, Oct. 2006.
[2] Matti Karjalainen et al., "Frequency-Warped Signal Processing for Audio Applications", 108th AES Convention, preprint No. 5171, Paris, France, February 2000.
[3] T. Paatero and M. Karjalainen, "Kautz filters and generalized frequency resolution: Theory and audio applications," J. Audio Eng. Soc., vol. 51, 2001.
[4] Bal´azs Bank, "Perceptually Motivated Audio Equalization Using Fixed-Pole Parallel Second-Order Filters", IEEE SIGNAL PROCESSING LETTERS, VOL. 15, 2008.
[5] Balázs Bank, "Loudspeaker and room response equalization using parallel filters: comparison of pole positioning strategies", AES 51st International Conference, Helsinki, Finland, 2013 August 22–24
[6] Balázs Bank, "Audio Equalization with Fixed-Pole Parallel Filters: An Efficient Alternative to Complex Smoothing", AES 128th Convention London, UK, May 22–25, 2010.
[7] Marco Jeub, Magnus Schäfer, and Peter Vary, "A Binaural Room Impulse Response Database for the Evaluation of Dereverberation Algorithms", Proceedings of International Conference on Digital Signal Processing (DSP), Santorini, Greece, July 2009
[8] http://www.openairlib.net/auralizationdb/content/live-room-sound-studio-laboratory-university-athens
[9] Panagiotis D. Hatziantoniou and John N. Mourjopoulos, "The Application of Complex Spectral Smoothing In The Analysis of Room Acoustics", Proceedings of First National Conference Recent Advances in Mechanical Engineering, Patras, Greece, September 17-20, 2001
[10] P.Hatziantoniou and J.Mourjopoulos,"Generalized Fractional-Octave Smoothing of Audio and Acoustic Responses", Journal of the Audio Engineering Society, Vol. 48, No. 4, pp. 259 - 280, April 2000.