Sensitivity of Multichannel Room Equalization to Listener Position
Sunil Bharitkar∗ Student Member, IEEE, Philip Hilmes, and Chris Kyriakakis Member, IEEE

Abstract—Traditionally, room response equalization is performed to improve sound quality at a given listener in applications ranging from automobile, home-theater, movie theater, multimedia education in classrooms. However, room responses vary with source and listener positions. Hence, in a multiple listener environment, equalization may be performed through spatial averaging of magnitude responses at locations of interest (e.g., in movie theater equalization). However, the performance of averaging based equalization, at the listeners, may be affected when listener positions change, or due to mismatch between microphone and listener positions (i.e., displacement effects). In this paper, we present a statistical approach to map displacement effects to a magnitude response averaging equalization performance metric. The results indicate that, for the analyzed listener configurations, the zone of equalization depends on, (i) distance of microphones/listeners from a source, (ii) the listener arrangement, and (iii) the source signal spectral composition.

I. 1. INTRODUCTION

A typical room is an acoustic enclosure that can be modelled as a linear system whose behavior at a particular listening position is characterized by an impulse response. The impulse response yields a complete description of the changes a sound signal undergoes when it travels from a source to a receiver (microphone/listener). The signal at the receiver consists of direct path components, discrete reflections that arrive a few milliseconds after the direct sound, as well as a reverberant field component. In addition, it is well established that room responses change with source and receiver locations in a room [1], [2]. In other words, a room response can be uniquely defined by a set of spatial co-ordinates \( l_i = (x_i, y_i, z_i) \). This assumes that the source is at origin and the receiver \( i \) is at the spatial co-ordinates, \( x_i, y_i \) and \( z_i \), relative to a source in the room.

Due to variations in room responses with listener positions relative to a source, in a multiple listener environment, room equalization should be performed for all listeners present in the room. With a good multiple listener equalization technique, all listeners in a given environment will experience high quality sound. Furthermore, the equalization technique should be robust to variations in listener head movements or microphone listener mismatch (e.g., the microphone used for measuring the room responses could be spaced from the actual listener head position during playback). Specifically, the equalized response should not vary significantly in the vicinity of the listeners.

One method for providing simultaneous multiple listener equalization is by measuring the room responses with microphones at all possible listener positions, averaging the magnitude response measurements, and inverting the result. The microphones are generally positioned, during measurements, at the expected center of a listener head. Although this equalization is aimed at achieving uniform frequency response coverage for all listeners, its performance is often limited due to, (i) mismatch between microphone measurement location and actual location for the center of the listener head, or (ii) variations in listener locations (e.g., head movements).

In this paper, we propose a statistical approach using modal equations for evaluating the robustness of equalization based on magnitude response averaging, due to the introduction of variations in room responses (generated either through (i) or (ii)), for different listener arrangements relative to a fixed sound source. Modal equations are used since they accurately model the sound pressure fluctuations (i.e., standing wave modes) at low frequencies which are much harder to equalize. In Section 2, we introduce necessary background used in the development of the proposed robustness analysis. Section 3 is dedicated to the development of the robustness analysis for spatial average based equalization. In Section 4, we present results based on simulations for different listener arrangements relative to a fixed sound source. Section 5 concludes the paper.

II. 2. ROOM ACOUSTICS FOR SIMPLE SOURCES

The Green’s function derived from the wave theory for sound fields in an enclosure is given by [1], [3]

\[
p_{\omega, \text{avg}} = \frac{1}{N} \sum_{l=1}^{N} |p_{\omega}(q_l)|
\]

\[
p_{\omega}(q_l) = \frac{jQ\omega\rho_0}{\pi} \sum_{n} \frac{p_n(q_l)p_n(q_o)}{K_n(k^2 - k_n^2)}
\]

\[
n = (n_x, n_y, n_z); k = \omega/345; q_o = (x_i, y_i, z_i)
\]

\[
k_n = \pi[(\frac{n_x}{L_x})^2 + (\frac{n_y}{L_y})^2 + (\frac{n_z}{L_z})^2]^{1/2}
\]

\[
\int \int V \int p_n(q_l)p_m(q_o) dV = K_n; (n = m) = 0 (n \neq m)
\]

where the eigenfunctions \( p_n(q_l) \) can be assumed to be orthogonal to each other under certain conditions, and the point source being at \( q_o \).

The authors are with the Immersive Audio Laboratory, Integrated Media Systems Center, University of Southern California, 3740 McClintock Avenue, Los Angeles, CA 90089-2564.

Abstract—Traditionally, room response equalization is performed to improve sound quality at a given listener in applications ranging from automobile, home-theater, movie theater, multimedia education in classrooms. However, room responses vary with source and listener positions. Hence, in a multiple listener environment, equalization may be performed through spatial averaging of magnitude responses at locations of interest (e.g., in movie theater equalization). However, the performance of averaging based equalization, at the listeners, may be affected when listener positions change, or due to mismatch between microphone and listener positions (i.e., displacement effects). In this paper, we present a statistical approach to map displacement effects to a magnitude response averaging equalization performance metric. The results indicate that, for the analyzed listener configurations, the zone of equalization depends on, (i) distance of microphones/listeners from a source, (ii) the listener arrangement, and (iii) the source signal spectral composition.

I. 1. INTRODUCTION

A typical room is an acoustic enclosure that can be modelled as a linear system whose behavior at a particular listening position is characterized by an impulse response. The impulse response yields a complete description of the changes a sound signal undergoes when it travels from a source to a receiver (microphone/listener). The signal at the receiver consists of direct path components, discrete reflections that arrive a few milliseconds after the direct sound, as well as a reverberant field component. In addition, it is well established that room responses change with source and receiver locations in a room [1], [2]. In other words, a room response can be uniquely defined by a set of spatial co-ordinates \( l_i = (x_i, y_i, z_i) \). This assumes that the source is at origin and the receiver \( i \) is at the spatial co-ordinates, \( x_i, y_i \) and \( z_i \), relative to a source in the room.

Due to variations in room responses with listener positions relative to a source, in a multiple listener environment, room equalization should be performed for all listeners present in the room. With a good multiple listener equalization technique, all listeners in a given environment will experience high quality sound. Furthermore, the equalization technique should be robust to variations in listener head movements or microphone listener mismatch (e.g., the microphone used for measuring the room responses could be spaced from the actual listener head position during playback). Specifically, the equalized response should not vary significantly in the vicinity of the listeners.

One method for providing simultaneous multiple listener equalization is by measuring the room responses with microphones at all possible listener positions, averaging the magnitude response measurements, and inverting the result. The microphones are generally positioned, during measurements, at the expected center of a listener head. Although this equalization is aimed at achieving uniform frequency response coverage for all listeners, its performance is often limited due to, (i) mismatch between microphone measurement location and actual location for the center of the listener head, or (ii) variations in listener locations (e.g., head movements).

In this paper, we propose a statistical approach using modal equations for evaluating the robustness of equalization based on magnitude response averaging, due to the introduction of variations in room responses (generated either through (i) or (ii)), for different listener arrangements relative to a fixed sound source. Modal equations are used since they accurately model the sound pressure fluctuations (i.e., standing wave modes) at low frequencies which are much harder to equalize. In Section 2, we introduce necessary background used in the development of the proposed robustness analysis. Section 3 is dedicated to the development of the robustness analysis for spatial average based equalization. In Section 4, we present results based on simulations for different listener arrangements relative to a fixed sound source. Section 5 concludes the paper.

II. 2. ROOM ACOUSTICS FOR SIMPLE SOURCES

The Green’s function derived from the wave theory for sound fields in an enclosure is given by [1], [3]

\[
p_{\omega, \text{avg}} = \frac{1}{N} \sum_{l=1}^{N} |p_{\omega}(q_l)|
\]

\[
p_{\omega}(q_l) = \frac{jQ\omega\rho_0}{\pi} \sum_{n} \frac{p_n(q_l)p_n(q_o)}{K_n(k^2 - k_n^2)}
\]

\[
n = (n_x, n_y, n_z); k = \omega/345; q_o = (x_i, y_i, z_i)
\]

\[
k_n = \pi[(\frac{n_x}{L_x})^2 + (\frac{n_y}{L_y})^2 + (\frac{n_z}{L_z})^2]^{1/2}
\]

\[
\int \int V \int p_n(q_l)p_m(q_o) dV = K_n; (n = m) = 0 (n \neq m)
\]

where the eigenfunctions \( p_n(q_l) \) can be assumed to be orthogonal to each other under certain conditions, and the point source being at \( q_o \).

The authors are with the Immersive Audio Laboratory, Integrated Media Systems Center, University of Southern California, 3740 McClintock Avenue, Los Angeles, CA 90089-2564.
For a rectangular enclosure with dimensions \((L_x, L_y, L_z)\), 
\( q_o = (0, 0, 0) \), the eigenfunctions and \( K_n \) in (1) are
\[
p_n(q) = \cos \left( \frac{n_x \pi x}{L_x} \right) \cos \left( \frac{n_y \pi y}{L_y} \right) \cos \left( \frac{n_z \pi z}{L_z} \right)
\]
\[
K_n = \frac{1}{L_x L_y L_z} \frac{\cos^2 \left( \frac{n_x \pi x}{L_x} \right) \cos^2 \left( \frac{n_y \pi y}{L_y} \right) \cos^2 \left( \frac{n_z \pi z}{L_z} \right)}{8}
\]  

(2)

III. 3. Robustness Analysis of Equalization using Magnitude Response Spatial Averaging

A performance function, \( W^{(i)}(\epsilon) \), that is used for analyzing the robustness of spatial average equalization, to room response variations is given as
\[
W^{(i)}(\epsilon) = E\left[ |p_\omega(\nu^{(i)}_e) p_\omega^{-1}(q_i) - p_\omega(q_i) p_\omega^{-1}(q_i)|^2 \right] \tag{3}
\]

where, \( p_\omega(\nu^{(i)}_e) \) is the pressure at locations in the \( \epsilon \) neighborhood of position \( i \) having pressure \( p_\omega(q_i) \) (\( \epsilon \)-neighbourhood is defined as all points at a distance of \( \epsilon \) from location \( i \)), and \( E\{\cdot\} \) denotes the expectation operator.

The performance measure (1) is defined in such a manner that when the displacement \( \epsilon \), about position \( i \) (whose response \( p_\omega(q_i) \) is originally used for determining the spatially averaged equalization filter \( p_\omega^{-1}(q_i) \), is zero, then \( W^{(i)}(\epsilon) = 0 \). Furthermore, the performance measure is computed as an average of the error between the response at the equalized location and the response at a displaced location having distance \( \epsilon \) from the equalized location. Obviously, this is not a psychoacoustic measure (i.e., it is not computed with different weights at different frequencies).

However, a future direction would be to form a psychoacoustically motivated composite measure using a weighted combination of \( W^{(i)}(\epsilon) \) over specific frequencies \( f \in [20 \text{ Hz}, 20 \text{ kHz}] \). For simplicity, in our analysis, we assume variations in responses due to displacements (or mismatch) in a horizontal plane (x-y plane). The analysis can be easily extended to include the vertical plane. Thus, simplification of (3) leads to
\[
W^{(i)}(\epsilon) = N^2 / \left( \sum_{i=1}^{N} |p_\omega(q_i)|^2 \right)^2 * \\
\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} |E\{p_\omega(\nu^{(i)}_e) p_\omega^{-1}(\nu^{(j)}_e)\}|^2 \\
- \frac{1}{N} \sum_{i=1}^{N} |E\{p_\omega^{-1}(\nu^{(i)}_e)\} p_\omega(q_i)|^2 \\
- \frac{1}{N} \sum_{i=1}^{N} |E\{p_\omega(\nu^{(i)}_e)\} p_\omega^{-1}(q_i)|^2
\]  

IV. 4. Practical Application Issues

We only need to compute the statistics associated with Terms (I), (II) and (III) (the terms within the expectations) in (4), since Term (IV) is a deterministic quantity.

Now, \( E\{p_\omega(\nu^{(i)}_e) p_\omega^{-1}(\nu^{(i)}_e)\} \) is the average over all locations along a circle of radius \( \epsilon \) from the \( i \)-th listener location. Assuming the source, all listeners, and each of the listener displacements are along the same z-plane \( (z = 0) \), then (1) in (4) can be simplified using the following equations,
\[
p_\omega(\nu^{(i)}_e) = \frac{j8Q_{\omega} p_{0}}{V} \sum_{n} \cos \left( \frac{n_x \pi x}{L_x} \right) \cos \left( \frac{n_y \pi y}{L_y} \right) \left( k^2 - k^2_{n} \right)^{\frac{-1}{2}}
\]

(5)

\[
E\{p_\omega(\nu^{(i)}_e) p_\omega^{-1}(\nu^{(i)}_e)\} = |\psi|^{2} \sum_{n,m} (1/\psi_{3}) \psi_{3}
\]

(6)

\[
\psi_{1} = \frac{8Q_{\omega} p_{0}}{V} \\
\psi_{2} = (k^2 - k^2_{n})(k^2 - k^2_{m}) \\
\psi_{3} = E\left\{ \cos \left( \frac{n_x \pi x_{i}^{(i)} y_{i}^{(i)}}{L_x} \right) \cos \left( \frac{n_y \pi y_{i}^{(i)}}{L_y} \right) \cos \left( \frac{m_x \pi x_{i}^{(i)}}{L_x} \right) \cos \left( \frac{m_y \pi y_{i}^{(i)}}{L_y} \right) \right\}
\]

(7)

Now, with \( \phi_{x}^{(i)} = x_i + \epsilon \cos \theta \) and \( \phi_{y}^{(i)} = y_i + \epsilon \sin \theta \)
\[
E\left\{ \cos \left( \frac{n_x \pi x_{i}^{(i)}}{L_x} \right) \cos \left( \frac{n_y \pi y_{i}^{(i)}}{L_y} \right) \cos \left( \frac{m_x \pi x_{i}^{(i)}}{L_x} \right) \cos \left( \frac{m_y \pi y_{i}^{(i)}}{L_y} \right) \right\} =
\]
\[
\frac{1}{2\pi} \int_{0}^{2\pi} \cos \left( \frac{n_x \pi x_{i}^{(i)} + \epsilon \cos \theta}{L_x} \right) \cos \left( \frac{n_y \pi y_{i}^{(i)} + \epsilon \sin \theta}{L_y} \right) \cos \left( \frac{m_x \pi x_{i}^{(i)} + \epsilon \cos \theta}{L_x} \right) \cos \left( \frac{m_y \pi y_{i}^{(i)} + \epsilon \sin \theta}{L_y} \right) d\theta
\]

(8)

Eq. (8) can be solved using the Matlab trapezoidal function. However, we found an approximate closed form expression to be computationally much faster. The following expressions were derived from standard trigonometric formulae; and using the first two terms in the polynomial expansion of the cosine function, and the first term in the polynomial expansion of the sine function since \( (\epsilon/L_x, \epsilon/L_y, \epsilon/L_z) << 1 \).

Thus,
\[
E\left\{ \cos \left( \frac{n_x \pi x_{i}^{(i)}}{L_x} \right) \cos \left( \frac{n_y \pi y_{i}^{(i)}}{L_y} \right) \cos \left( \frac{m_x \pi x_{i}^{(i)}}{L_x} \right) \cos \left( \frac{m_y \pi y_{i}^{(i)}}{L_y} \right) \right\} = \frac{1}{2\pi} (A + B + C)
\]

(9)

where
\[
A = \pi \cos \left( \frac{n_x \pi x_{i}^{(i)}}{L_x} \right) \cos \left( \frac{n_y \pi y_{i}^{(i)}}{L_y} \right) \cos \left( \frac{m_x \pi x_{i}^{(i)}}{L_x} \right) \cos \left( \frac{m_y \pi y_{i}^{(i)}}{L_y} \right)
\]

1 - 722
where, 
\[
\epsilon_x = \frac{\pi \epsilon}{\sqrt{2L_x}} ; \quad \epsilon_y = \frac{\pi \epsilon}{\sqrt{2L_y}}
\]
\[
u_x = \nu_x \epsilon_x \quad ; \quad \nu_y = \nu_y \epsilon_y \quad (11)
\]

Thus (10) can be substituted in (9) and subsequently in (4) to determine Term I.

Now Terms (II) and (III) in (4) can be combined to give,
\[
-\frac{8Q\omega\rho}{V}|\sum_{m,n}p_m(q)E\left\{\cos\left(\frac{n_x \pi x}{L_x}\right) \cos\left(\frac{n_y \pi y}{L_y}\right) \right\} =
\]
\[
-2|2S\omega\rho| \sum_{m,n}p_m(q)E\left\{\cos\left(\frac{n_x \pi x}{L_x}\right) \cos\left(\frac{n_y \pi y}{L_y}\right) \right\} (12)
\]

Now again using \(\phi_i^{(1)} = x_i + \epsilon \cos \theta ; \phi_i^{(2)} = y_i + \epsilon \sin \theta\), we have
\[
E\left\{\cos\left(\frac{n_x \pi x}{L_x}\right) \cos\left(\frac{n_y \pi y}{L_y}\right) \right\} =
\]
\[
\frac{1}{2\pi} \int_0^{2\pi} \cos\left(\frac{n_x \pi x}{L_x} + \epsilon \cos \theta \right) \cos\left(\frac{n_y \pi y}{L_y} + \epsilon \sin \theta \right) d\theta (13)
\]

Thus, upon again using the fact that \((\epsilon/L_x, \epsilon/L_y, \epsilon/L_z) < < 1\), we can solve (13) as
\[
E\left\{\cos\left(\frac{n_x \pi x}{L_x}\right) \cos\left(\frac{n_y \pi y}{L_y}\right) \right\} =
\]
\[
\frac{1}{2\pi} \int_0^{2\pi} \cos\left(\frac{n_x \pi x}{L_x} + \epsilon \cos \theta \right) \cos\left(\frac{n_y \pi y}{L_y} + \epsilon \sin \theta \right) d\theta (13)
\]

Substituting (14) in (12) and subsequently into (4) gives

Substituting (14) in (12) and subsequently into (4) gives

V. 5. Conclusions

In this paper we presented a statistical approach using modal equations for evaluating the robustness of equalization based on magnitude response averaging, due to the introduction of variations in room responses, for different listener arrangements relative to a fixed sound source. The simulations were performed for a six listener setup with a simple source in a cubic room.

Future research will be directed towards, (i) obtaining the results for the performance measure, \(W_\omega^{(i)}(\epsilon)\), for rooms of arbitrary dimensions, (ii) extending the simulations over more frequencies and positions, (iii) forming a psycho-acoustically motivated composite measure using a weighted combination of \(W_\omega^{(i)}(\epsilon)\) over specific frequencies.

VI. Acknowledgements

This research was funded in part by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152. The project was also sponsored in part by the U.S. Army. The content of the information does not necessarily reflect the position or the policy of the Government of the United States of America, and no official endorsement should be inferred.

References

Fig. 1. Simulated setup for a six position displacement effects analysis. The six equalized positions are marked by an asterisk and the source is denoted by a circle (at a distance of 1m from the listeners).

Fig. 2. Results from the simulation for the setup in Fig. 1 for frequency $f=150$ Hz. The results indicate that the equalization performance is affected for each listener.

Fig. 3. Average equalization performance for the setup in Fig. 1 for frequency $f=150$ Hz. The results indicate that the zone of equalization degrades with increasing listener displacement.

Fig. 4. Average equalization performance for the setup in Fig. 1 (except that the listeners are now at a distance of 2m from the source). Comparing with Fig. 3, it can be seen that the zone of equalization degrades with increasing distance from the source.

Fig. 5. Simulated setup for a six position displacement effects analysis in a rectangular configuration in front of a source.

Fig. 6. Average equalization performance, for the setup in Fig. 5, for source signal having frequency of 150 Hz.

Fig. 7. Average equalization performance, for the setup in Fig. 5, for source signal with frequency 300 Hz. Comparing with Fig. 6, it can be seen that the zone of equalization degrades with decreasing wavelength of the source signal.