Robustness of Spatial Averaging Equalization Methods: A Statistical Approach

Sunil Bharitkar* *Student Member, IEEE*, Philip Hilmes, and Chris Kyriakakis *Member, IEEE*
E-mail:bharitka@sipi.usc.edu

Section: 4. Speech and Audio Processing

Abstract: Traditionally, room response equalization is performed to improve sound quality at a given listener. However, room responses vary with source and listener positions. Hence, in a multiple listener environment, equalization may be performed through spatial averaging of room responses. However, the performance of averaging based equalization, at the listeners, may be affected when listener positions change. In this paper, we present a statistical approach to map variations in listener positions to performance of spatial averaging based equalization. The results indicate that, for the analyzed listener configurations, the zone of equalization depends on distance of microphones from a source and the frequencies in the sound.

I. Introduction

A typical room is an acoustic enclosure that can be modeled 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 $l_i \stackrel{\Delta}{=} (x_i, y_i, z_i)$. 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. 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 measurements, and inverting the stable component of 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 for evaluating the robustness of spatial averaging based equalization, 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. The proposed approach uses a statistical description for the reverberant components in the responses (viz., normalized correlation functions). A similar approach is followed in [3] for determining variations in performance for single listener equalization. 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 source. Section 5 concludes the paper.

II. Room Acoustics for Simple Sources

A. Sound Pressure at a Position in a Room

The sound pressure $p_{f,i}$ at location i and frequency f can be expressed as a sum of direct field component, $p_{f,d,i}$, and a reverberant field component, $p_{f,r,i}$, as given by

$$p_{f,i} = p_{f,d,i} + p_{f,r,i} \tag{1}$$

B. Free-field Pressure due to a Periodic Source

The direct field component for sound pressure, $p_{f,d,i}$, at location i, due to a sound source having frequency f located at i_0 can be expressed as [4]

$$\begin{aligned} p_{f,d,i} &= p_{f,d}(i|i_0)e^{\imath\omega t} \\ &= -jk\rho c S_f g_f(i|i_0)e^{-j\omega t} \\ g_f(i|i_0) &= \frac{1}{4\pi R}e^{jkR} \\ R^2 &= |i-i_0|^2 \end{aligned} \tag{2}$$

where $p_{f,d}(i|i_0)$ is the direct component sound pressure amplitude, S_f is the source strength, $k = 2\pi/\lambda$ is the wavenumber, $c = \lambda f$ is the speed of sound (343 m/s) and ρ is the density of the medium (1.4 for air).

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

IEEE 36th Asilomar Conference on Signals, Systems & Computers. Pacific Grove. CA. Nov. 3-6, 2002. http://web.nps.navy.mil/~asilomar

C. Reverberation Statistics due to a Periodic Source

The normalized correlation function [5] which expresses a statistical relation between sound pressures, of reverberant components, at separate locations i and j, is given by

$$\frac{E\{p_{f,r,i}p^*_{f,r,i}\}}{\sqrt{E\{p_{f,r,i}p^*_{f,r,i}\}}\sqrt{E\{p_{f,r,j}p^*_{f,r,j}\}}} = \frac{\sin kR_{ij}}{kR_{ij}} \quad (3)$$

where R_{ij} is the separation between the two locations i and j relative to an origin, and $E\{.\}$ is the expectation operator.

The reverberant-field mean square pressure is defined as,

$$E\{p_{f,r,i}p^*_{f,r,i}\} = \frac{4c\rho\Pi_a(1-\bar{\alpha})}{S\bar{\alpha}} \quad (4)$$

where, Π_a is the power of the acoustic source, $\bar{\alpha}$ is the average absorption co-efficient of the walls, and S is the wall area of the room.

The assumption of a statistical description (as given in (3), (4)) for reverberant fields in rooms is justified if certain conditions are fulfilled [6]. These conditions are typically fulfilled in rectangular rooms at frequencies above the Schroeder frequency, $f_s = 2000\sqrt{T_{60}/V}$ Hz (T_{60} is the reverberation time in seconds, and V is the volume in m^3). Furthermore, under the conditions in [6], the direct and reverberant sound pressure are uncorrelated.

III. Robustness Analysis of Spatial Averaging Equalization

A performance function, J_f, that is used for analyzing the robustness, of spatial average equalization, to room response variations is given as

$$\begin{aligned}\bar{W}_f &= \frac{1}{N}\sum_{i=1}^{N}\epsilon_{f,i}(r)\\ \epsilon_{f,i}(r) &= E\{|\tilde{p}_f\bar{p}_f - p_{f,i}\bar{p}_f|^2\}\end{aligned} \quad (5)$$

In (4), $\epsilon_{f,i}(r)$ represents the equalization error in the r-neighborhood of the equalized location i (having response $p_{f,i}$) (r-neighourhood is defined as all points at a distance of r from location i). The neighboring response at distance r from location i is denoted by $\tilde{p}_f$, the spatial average equalization response is denoted by $\bar{p}_f$. The expectation is performed over all neighboring locations at a distance r from the spatial average equalized location i. The performance function $\bar{W}_f$ is the average of all the equalization errors in the vicinity of the N equalized locations.

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, (4) can be expanded to give,

$$\begin{aligned}\epsilon_{f,i}(r) &= E\{|\frac{\tilde{p}_f N}{\sum_{j=1}^{N}p_{f,j}} - \frac{p_{f,i}N}{\sum_{j=1}^{N}p_{f,j}}|^2\}\\ &= N^2E\{\frac{1}{\sum_j\sum_k p_{f,j}p^*_{f,k}}[\tilde{p}_f\tilde{p}^*_f - \tilde{p}_f p^*_{f,i}\\ &\quad -\tilde{p}^*_f p_{f,i} + p_{f,i}p^*_{f,i}]\}\end{aligned} \quad (6)$$

An approximate simplification for (5) can be done by using the Taylor series expansion [7]. Accordingly, if g is a function of random variables,x_i, with average values $E\{x_i\} = \bar{x}_i$, then $g(x_1, x_2, ..., x_n) = g(x)$ can be expressed as $g(x) = g(\bar{x}) + \sum_{i=1}^{n} g'_i(\bar{x})(x_i - \bar{x}_i) + g(\hat{x})$, where $g(\hat{x})$ is a function of order 2 (i.e., all its partial derivatives up to the first order vanish at $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$. Thus, to a zeroth order of approximation $E\{g(x)\} \approx g(\bar{x})$.

Hence, an approximation for (5) is given as,

$$\epsilon_{f,i}(r) = N^2\frac{E\{\tilde{p}_f\tilde{p}^*_f - \tilde{p}_f p^*_{f,i} - \tilde{p}^*_f p_{f,i} + p_{f,i}p^*_{f,i}\}}{\sum_j\sum_k E\{p_{f,j}p^*_{f,k}\}} \quad (7)$$

We use the following identities for determining the denominator of (7) ,

$$\begin{aligned}E\{p_{f,j}p^*_{f,k}\} &= E\{p_{f,d,j}p^*_{f,d,k} + p_{f,r,j}p^*_{f,r,k}\} &(8)\\ |kc\rho S_f|^2 &= 4\pi\Pi_a c\rho &(9)\\ E\{p_{f,d,j}p^*_{f,d,k}\} &= \frac{\Pi_a c\rho}{4\pi R_j R_k}e^{jk(R_j - R_k)} &(10)\\ E\{p_{f,r,j}p^*_{f,r,k}\} &= \frac{4c\rho\Pi_a(1-\bar{\alpha})}{S\bar{\alpha}}\frac{\sin kR_{jk}}{kR_{jk}} &(11)\\ R_{jk} &= \sqrt{R_j^2 + R_k^2 - 2R_jR_k\cos\theta_{jk}} &(12)\end{aligned}$$

In summary (8) is obtained by using (1) and knowing that the reverberant and direct field components of sound pressure are uncorrelated, (9) is derived in [4] (pp. 311), (10) is determined by using (2) and (9), and (11) is determined from (3) and (4). In (12), which is the cosine law, θ_{jk} is the angle, subtended at the source at i_0, between locations j and k.

Thus, the denominator term in (7) is

$$\begin{aligned}\sum_j\sum_k E\{p_{f,j}p^*_{f,k}\} &= \sum_j\sum_k(\frac{\Pi_a c\rho}{4\pi R_j R_k}e^{jk(R_j-R_k)}\\ &\quad + \frac{4c\rho\Pi_a(1-\bar{\alpha})}{S\bar{\alpha}}\frac{\sin kR_{jk}}{kR_{jk}})\end{aligned} \quad (13)$$

Now, the first numerator term in (7) is,

$$\begin{aligned}E\{\tilde{p}_f\tilde{p}^*_f\} &= E\{\tilde{p}_{f,d}\tilde{p}^*_{f,d} + \tilde{p}_{f,r}\tilde{p}^*_{f,r}\}\\ E\{\tilde{p}_{f,d}\tilde{p}^*_{f,d}\} &= |k\rho cS_f|^2E\{g_f(\tilde{i}|i_0)g^*_f(\tilde{i}|i_0)\}\\ &= |k\rho cS_f|^2E\{\frac{1}{(4\pi)^2|\tilde{R}|^2}\}\end{aligned} \quad (14)$$

where, $\tilde{R}$ is the distance from a source at i_0 relative to spatial average equalized location i, and is determined by using cosine law (viz., $\tilde{R} = \sqrt{R_i^2 + r^2 - 2R_i r\cos\theta_i}$, where θ_i is the angle subtended at the source between location i and the location in the r-neighborhood of location i). The result from applying the expectation can be found by averaging over all locations in a circle in the r-neighborhood of location i (since for simplicity we have assumed mismatch in the horizontal or x-y plane). Thus,

$$E\{\frac{1}{|4\pi\tilde{R}|^2}\} = \frac{1}{2}\frac{1}{(4\pi)^2}\int_{-1}^{1}\frac{d(\cos\theta_i)}{R_i^2 + r^2 - 2R_i r\cos\theta_i} \quad (15)$$

Simplifying (15) and substituting the result in (14) gives

$$\begin{aligned} E\{\tilde{p}_{f,d}\tilde{p}^*_{f,d}\} &= \frac{|k\rho c S_f|^2}{2(4\pi)^2 R_i r}\log|\frac{R_i+r}{R_i-r}| \\ &= \frac{\Pi_a \rho c}{8R_i r\pi}\log|\frac{R_i+r}{R_i-r}| \quad (16) \\ E\{\tilde{p}_{f,r}\tilde{p}^*_{f,r}\} &= \frac{4c\rho\Pi_a(1-\bar{\alpha})}{S\bar{\alpha}} \quad (17) \end{aligned}$$

The result in (16) is obtained by using (9), whereas (17) is a re-statement of (4). Thus,

$$\begin{aligned} E\{\tilde{p}_f\tilde{p}^*_f\} &= \frac{\Pi_a\rho c}{8R_i r\pi}\log|\frac{R_i+r}{R_i-r}| \\ &\quad + \frac{4c\rho\Pi_a(1-\bar{\alpha})}{S\bar{\alpha}} \quad (18) \end{aligned}$$

The correlation, $E\{\tilde{p}_{f,d}p^*_{f,d,i}\}$, in the direct-field component for the second term in the numerator of (7) is

$$\begin{aligned} &|k\rho c S_f|^2\frac{1}{2(4\pi)^2}\int_{-1}^{1}\frac{e^{jk(\sqrt{R_i^2+r^2-2R_i r\cos\theta_i}-R_i)}d\cos\theta_i}{R_i\sqrt{R_i^2+r^2-2R_i r\cos\theta_i}} \\ &= \frac{\Pi_a\rho c}{4\pi R_i^2}\frac{1}{(4\pi R_i)^2}\frac{\sin kr}{kr} \quad (19) \end{aligned}$$

The reverberant field correlation for the second term in the numerator of (7) can be found using (3), and is

$$\begin{aligned} E\{\tilde{p}_{f,r}p^*_{f,r,i}\} &= \sqrt{E\{\tilde{p}_{f,r}\tilde{p}^*_{f,r}\}}\sqrt{E\{p_{f,r,i}p^*_{f,r,i}\}} \\ &= \frac{4c\rho\Pi_a(1-\bar{\alpha})}{S\bar{\alpha}}\frac{\sin kr}{kr} \quad (20) \end{aligned}$$

The third numerator term in (7) can be found in a similar manner as compared to the derivation for (19), and (20).

The last term in the numerator of (7) is computed to yield,

$$E\{p_{f,i}p^*_{f,i}\} = \frac{\Pi_a\rho c}{4\pi R_i^2} + \frac{4\rho c\Pi_a(1-\bar{\alpha})}{S\bar{\alpha}} \quad (21)$$

Substituting the computed results into (7), and simplifying by cancelling certain common terms in the numerator and the denominator, the resulting equalization error due to displacements (viz., mismatch in responses) is,

$$\begin{aligned} \epsilon_{f,i}(r) &\approx \frac{N^2}{\psi_1}[\frac{1}{8R_i r\pi}\log|\frac{R_i+r}{R_i-r}| + 2\psi_2 \\ &\quad + \frac{1}{2\psi_3} - (\frac{1}{\psi_3}+2\psi_2)\frac{\sin kr}{kr}] \quad (22) \\ \psi_1 &= \sum_j\sum_k(\frac{1}{4\pi R_j R_k}e^{jk(R_j-R_k)} \\ &\quad + \psi_2\frac{\sin kR_{jk}}{kR_{jk}}) \\ \psi_2 &= \frac{4(1-\bar{\alpha})}{S\bar{\alpha}} \\ \psi_3 &= 2\pi R_i^2 \\ R_{jk} &= \sqrt{R_j^2+R_k^2-2R_jR_k\cos\theta_{jk}} \end{aligned}$$

Finally, substituting (22) into (5) yields the necessary equation for $\bar{W}_f$.

IV. RESULTS

We simulated (5) for frequencies above the Schroeder frequency $f_s = 102$ Hz (i.e., $T_{60} = 1.342s, V = 512m^3$). The average absorption coefficient, $\bar{\alpha}$, was set at 0.16 Sabines. Specifically, the frequencies of test were 500 Hz, 2 kHz, 4 kHz, and 10 kHz.

In the first setup, we simulated a circular arrangement relative to a source. Specifically, there were six microphones arranged symmetrically around a central source equidistant from a central source. The three distances to the six microphones, from a central source, used for evaluating (5) were (i) 1 m, (ii) 2 m, and (iii) 3 m. The results are plotted in Figs. 1-3. No significant dependencies on frequencies are seen for any given plot. However, the zone of equalization, where more than 10 dB of reverberation reduction is achieved, shrinks as the distance from the source increases. Specifically, the zone of equalization is a sphere of a radius 0.075λ for $R_i = R = 1m, (i = 1, 2.., 6)$, whereas it is 0.045λ for $R_i = R = 3m, (i = 1, 2.., 6)$. This is to be expected, since as the distance from a source increases to beyond a threshold distance, the reverberation field, which is characterized by a random model, has a larger impact than the direct field. Thus, this result indicates the relative lower efficacy of spatial average equalization for reverberation reduction at larger distances. It would be interesting to see how different multiple listener equalization algorithms perform by comparing their spheres of reverberation reduction. This is one of the goals for the future.

In the second setup, we simulated a rectangular arrangement of six microphones, with a source in the front of the arrangement. Specifically, microphones 1 and 3 were at a distance of 3 m from the source, microphone 2 was at 2.121 m, microphones 4 and 6 were at 4.743 m, and microphone 5 was at 4.242 m. The result is depicted in Fig. 4. A clear frequency dependence can be seen in this configuration. Higher frequencies have a smaller zone of equalization (due to their shorter wavelengths) as compared to lower frequencies.

V. CONCLUSIONS

In this paper, we proposed a performance function and a statistical formulation for determining robustness of spatial average equalization for simultaneous multiple listener room response compensation. We found the influence of frequency and distance on the size of equalization zones for different microphone configurations. Future goals will be directed to using the proposed method for comparing different multiple listener equalization techniques in terms of their robustness to response mismatch.

VI. ACKNOWLEDGEMENTS

This research has been funded in part by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152.

This research was also funded in part by the United States Department of the Army under contract number DAAD 19-99-D-0046. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the Department of the Army.

REFERENCES

[1] H. Kuttruff, *Room Acoustics,* Elseiver Applied Science, 3rd ed., New York. 1991.

[2] J. Mourjopoulos, "On the variation and invertibility of room impulse response functions," *Journal of Sound and Vibration,* vol. 102(2), pp. 217–228, 1985.

[3] B. D. Radlovic, R. C. Williamson, and R. A. Kennedy, "Equalization in acoustic reverberant environment:robustness results," *IEEE Trans. on Speech and Audio Proc.,* 8(3), pp. 311–319, May 2000.

[4] P. M. Morse and K. Uno Ingard, *Theoretical Acoustics,* Princeton Univ. Press, Princeton. 1986.

[5] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson, "Measurement of correlation coefficients in reverberant sound fields," submitted *J. Acoust. Soc. Amer.,* vol. 27(6), pp. 1072–1077, 1955.

[6] M. R. Schroeder, "Frequency correlation of frequency responses in rooms," submitted *J. Acoust. Soc. Amer.,* vol. 34(12), pp. 1819–1823, 1962.

[7] M. Kendall and A. Stuart, *The Advancedd Theory of Statistics,* 4th Ed. Griffin Press, London, 1976, pp. 246

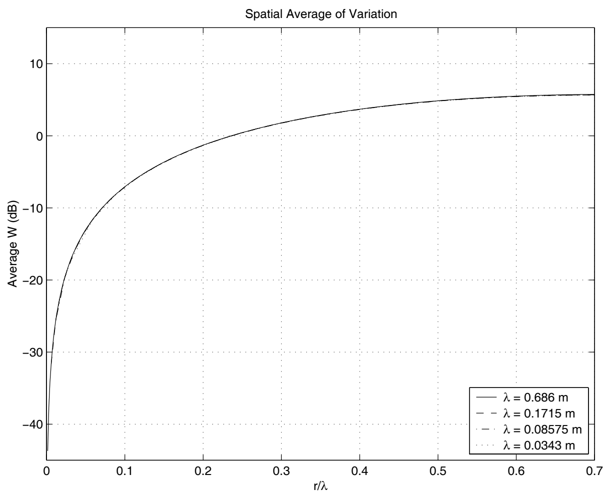

Fig. 1. $\bar{W}$ for six microphones equidistant at 1m from source.

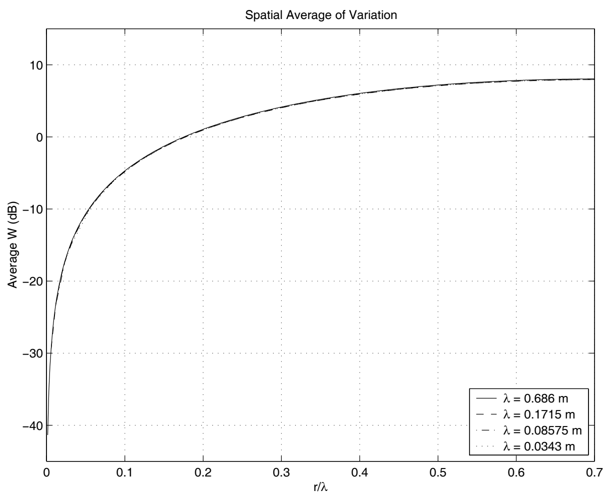

Fig. 2. $\bar{W}$ for six microphones equidistant at 2m from source.

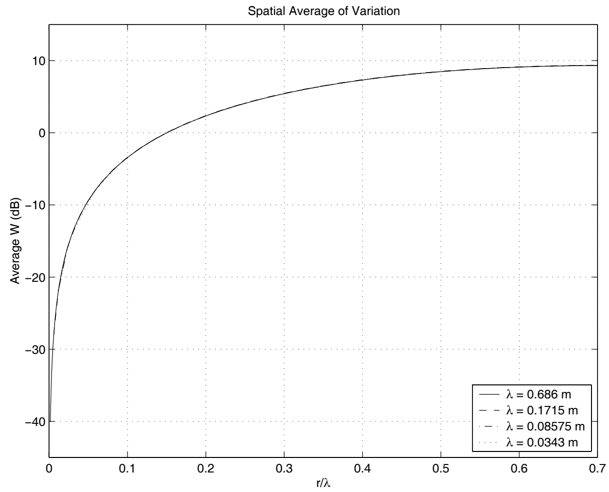

Fig. 3. $\bar{W}$ for six microphones equidistant at 3m from source.

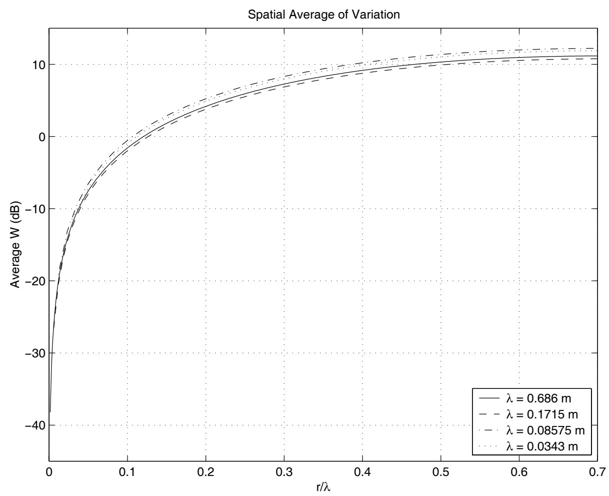

Fig. 4. $\bar{W}$ for six microphones in a rectangular arrangement.