Estimation of Acoustical Room Transfer Functions

T. Gustafsson[1] {tonyg@s2.chalmers.se} H. R. Pota[2] {h-pota@adfa.edu.au}
J. Vance[1] {jvance@ucsd.edu} B. D. Rao[1] {brao@ece.ucsd.edu}
M. M. Trivedi[1] {trivedi@ece.ucsd.ed}

Abstract

This paper presents a method to obtain room transfer functions for applications in the design of intelligent systems for video conferencing. In these applications the acoustical dynamics, for example, affects the quality of the recorded speech signal and the ability to achieve accurate source localization. The first part of the paper gives results on the identification of parameterized room transfer functions. The second part presents a method to extrapolate room transfer functions from the identified models. The extrapolation is achieved by getting a functional form of the transfer function parameters which depends on the speaker and the microphone locations.

1 Introduction

Room acoustics is a well researched phenomenon [6], and there are several well-known methods to obtain room acoustics, see for example [1, 5, 9]. These (numerical) methods have been used mainly to study the integral effects of sound, e.g., reverberation times and acoustic energy distribution. Our interest is in the transient response of the room acoustics. The methods used to study the integral acoustical effects do not work well for obtaining the transient response. Active noise control [7] and echo canceling [2] applications also require a suitable model to determine the transient response. In these applications the room transfer function is either obtained from the analytical solution of the wave equation or a black-box dynamic model is identified from the experimental data [7]. Our interest in obtaining room transfer functions stems from the desire to design intelligent systems for video conferencing, where the acoustical room transfer function, for example, affects the quality of the recorded speech signal, and the ability to achieve accurate source localization. For this type of application a few point-to-point room transfer functions, used in control applications, are not sufficient. For the ultimate aim of high fidelity sound reproduction it is envisaged that hundreds of room transfer functions between several points in the room will be needed. The use of the wave equation to derive these transfer functions becomes impractical when one realizes that even a modest room dynamics, in the 20–500 Hz range, can only be modeled with system orders around one thousand [8]. In practice one also knows that there is an overcrowding of modes and due to the damping provided by the absorbent room surfaces, only a fraction of the modes are prominent. This typically means that only $\frac{1}{4}^{th}$ or $\frac{1}{5}^{th}$ of the theoretically predicted modes need to be retained to capture the room acoustical dynamics with high fidelity. With this observation in mind, one can see that a practical way of obtaining room transfer functions should be based on an exchange of information between experimental data and the theoretical model. This paper presents an approach to estimating room transfer functions where the experimental data is gathered first and then the estimated modes are used to extrapolate room transfer functions at several locations in the room. The extrapolating functions are suggested by the analytical model.

A similar approach to estimate room transfer functions has been proposed in [3]. The method in [3] is based on the key observation that the zeros of the transfer function of the acoustical system depend on the speaker and the microphone location, while the poles depend only on the room geometry. See [4] for a detailed discussion on these aspects. In [3] the common-pole acoustical models are first derived from the experimental data. These models are then used to interpolate and extrapolate the room transfer functions between the speaker and an arbitrary microphone location. In this paper we primarily take the findings in [3] one step further. We investigate various identification procedures, and then we extrapolate the estimated transfer functions, not only in one dimension as in [3], but in two dimensions. We also highlight a couple of the difficulties in estimating these transfer functions.

The research presented in this paper is a part of a larger research thrust that is being pursued in the Computer Vision and Robotics Research Laboratory, University of California, San Diego. The overall objective of the research is to develop Intelligent Environments where the system is able to extract and dynamically maintain robust 3-D geometrical and compositional models of a

[1]Computer Vision and Robotics Research Laboratory, Electrical and Computer Engineering Department, University of California-San Diego, La Jolla CA 92093-0407

[2]School of Electrical Engineering, University College, University of New South Wales, ADFA Canberra Australia 2600

room and react to the activities in the room in an "intelligent" and efficient manner [10]. AVIARY (Audio Video Interactive Appliances, Rooms, and sYstems) is the main testbed built for this research project and it utilizes a network of cameras and microphone arrays for capturing the sensory information.

This paper is organized as follows. The next section presents the derivation of the assumed modes model from the three-dimensional wave equation. In Section 3 the identification method for obtaining room transfer functions from the experimental data is presented. Section 4 gives the method used in extrapolating new room transfer functions from the identified transfer functions. Section 5 lists the conclusions of this work.

2 AVIARY Model

The wave equation for a three-dimensional acoustic system is [6]:

$$\frac{\partial^2 p(\boldsymbol{r},t)}{\partial x^2} + \frac{\partial^2 p(\boldsymbol{r},t)}{\partial y^2} + \frac{\partial^2 p(\boldsymbol{r},t)}{\partial z^2} - \frac{1}{c_0^2}\frac{\partial^2 p(\boldsymbol{r},t)}{\partial t^2} = \rho_0 \dot{u}_{\boldsymbol{r}_s}(t)\delta(\boldsymbol{r}-\boldsymbol{r}_s) \quad (1)$$

where $p(\boldsymbol{r},t)$ is the acoustic pressure at location $\boldsymbol{r} = \{x,y,z\}$, c_0 is the speed of sound, ρ_0 is the air density, $\boldsymbol{r}_s = \{x_s,y_s,z_s\}$ is the speaker location, $\dot{u}_{\boldsymbol{r}_s}(t)$ is the acceleration imparted to the medium by the speaker, and $\delta(\cdot)$ is the Dirac delta function. AVIARY has rigid boundary conditions:

$$\left.\frac{\partial p(\boldsymbol{r},t)}{\partial x}\right|_{\substack{x=0\\x=L_x}} = \left.\frac{\partial p(\boldsymbol{r},t)}{\partial y}\right|_{\substack{y=0\\y=L_y}} = \left.\frac{\partial p(\boldsymbol{r},t)}{\partial z}\right|_{\substack{z=0\\z=L_z}} = 0 \quad (2)$$

where AVIARY has dimensions $L_x \times L_y \times L_z$.

Let us assume that a solution to the wave equation (1) can be written in the following form:

$$p(\boldsymbol{r},t) = \sum_{i,j,k=0}^{\infty} \phi_i(x)\psi_j(y)\varphi_k(z)\, q_n(t). \quad (3)$$

This is known as the assumed modes solution; $\phi_i(x)$, $\psi_j(y)$, and $\varphi_k(z)$ are known as the mode shapes, $q_n(t)$ are the generalized co-ordinates, and n is a unique number for each triplet $\{i,j,k\}$. For the boundary conditions (2) using standard methods one can obtain [6]:

$$p(\boldsymbol{r},t) = \sum_{i,j,k=0}^{\infty} \sqrt{\frac{8}{L_xL_yL_z}} \cos\frac{i\pi x}{L_x}\cos\frac{j\pi y}{L_y}\cos\frac{k\pi z}{L_z} q_n(t). \quad (4)$$

Substituting the above expression (3) in the wave equation (1), multiplying by $\phi_i(x)\psi_j(y)\varphi_k(z)$, and then integrating both sides as

$$\int_0^{L_z}\int_0^{L_y}\int_0^{L_x} (.) dxdydz,$$

the following simplified relationship is obtained:

$$-\left(\frac{i^2\pi^2}{L_x^2} + \frac{j^2\pi^2}{L_y^2} + \frac{k^2\pi^2}{L_z^2}\right) q_n(t) - \frac{1}{c_0^2}\ddot{q}_n(t) = \rho_0 \dot{u}_{\boldsymbol{r}_s}(t)\phi_i(x_s)\psi_j(y_s)\varphi_k(z_s). \quad (5)$$

Defining

$$\omega_n^2 = c_0^2\left[\frac{i^2\pi^2}{L_x^2} + \frac{j^2\pi^2}{L_y^2} + \frac{k^2\pi^2}{L_z^2}\right],$$

$$b_n(\boldsymbol{r}_s) = -c_0^2\rho_0\phi_i(x_s)\psi_j(y_s)\varphi_k(z_s), \text{ and}$$

$$c_n(\boldsymbol{r}_m) = \phi_i(x_m)\psi_j(y_m)\varphi_k(z_m),$$

where $\boldsymbol{r}_m = \{x_m, y_m, z_m\}$ is the microphone location, equation (5) can be written as:

$$\ddot{q}_n + \omega_n^2 q_n = b_n(\boldsymbol{r}_s)\dot{u}_{\boldsymbol{r}_s}(t), \quad n = 1\ldots\infty. \quad (6)$$

Let us define $p_n(\boldsymbol{r}_m,t) \triangleq c_n(\boldsymbol{r}_m)q_n(t)$. Physically speaking this means that $p_n(\boldsymbol{r},t)$ is the acoustic pressure due to n^{th} modal function. The acoustic pressure

$$p(\boldsymbol{r}_m,t) = \sum_{n=1}^{\infty} p_n(\boldsymbol{r}_m,t) = \sum_{n=1}^{\infty} c_n(\boldsymbol{r}_m)q_n(t)$$

is the sum of all the modal contributions.

In any practical model all the infinite modes cannot be included and only a finite number, say L, of modes are included. The above information can also be expressed as the transfer function between a speaker at location $\boldsymbol{r}_s$ and a microphone at location $\boldsymbol{r}_m$ as:

$$\frac{P(\boldsymbol{r}_m,s)}{sU_{\boldsymbol{r}_s}(s)} = \sum_{n=1}^{L}\frac{b_n(\boldsymbol{r}_s)c_n(\boldsymbol{r}_m)}{s^2+2\zeta_n\omega_n s+\omega_n^2}, \quad (7)$$

where ζ_n is the damping of the n^{th} mode. The values of these constants are arrived at from experimental data.

In a practical scenario, we of course have to work with sampled data. Using an assumption on "impulse-invariance", the discrete-time transfer function corresponding to (7) reads as

$$H(e^{j\omega};\boldsymbol{r}_s,\boldsymbol{r}_m) = \sum_{n=1}^{L}\frac{b_n(\boldsymbol{r}_s)c_n(\boldsymbol{r}_m)\xi_n e^{-j\omega}}{1-2e^{-\varsigma_n T_s}\cos(\kappa_n T_s)e^{-j\omega}+e^{-2\varsigma_n T_s}e^{-2j\omega}} \quad (8)$$

where T_s denotes the sampling interval and

$$\varsigma_n = \zeta_n \omega_n \quad (9)$$

$$\kappa_n^2 = \omega_n^2(1-\zeta_n^2) \quad (10)$$

$$\xi_n = \frac{T_s \sin(\kappa_n T_s) e^{-\varsigma_n T_s}}{\kappa_n}. \quad (11)$$

The terms $b_n(\boldsymbol{r}_s)c_n(\boldsymbol{r}_m)$, $n = 1, \ldots, L$, are of special interest. These terms depend only on the speaker and the microphone location. It is this parameterization which helps in extrapolating the available room transfer functions to obtain room transfer functions at new locations as shown in Section 4.

3 Identification

In this section we outline a procedure for estimating the room transfer functions from the experimentally obtained frequency response data. A sketch of the room with the microphone arrangement is shown in Figure 1. In our identification experiments we used a swept sinusoid input ($0 - 500$ Hz), with a sampling frequency of $F_s = 16$ kHz. The data are subsequently down-sampled with a factor 12. Suppose that we have collected N samples of the outputs of M microphones, denoted as

$$\boldsymbol{y}(t) = [y_1(t), \ \cdots, y_M(t)]^T, \ t = 0, \cdots, N-1. \quad (12)$$

The microphones are located at positions $\boldsymbol{r}_m^1, \cdots, \boldsymbol{r}_m^M$.

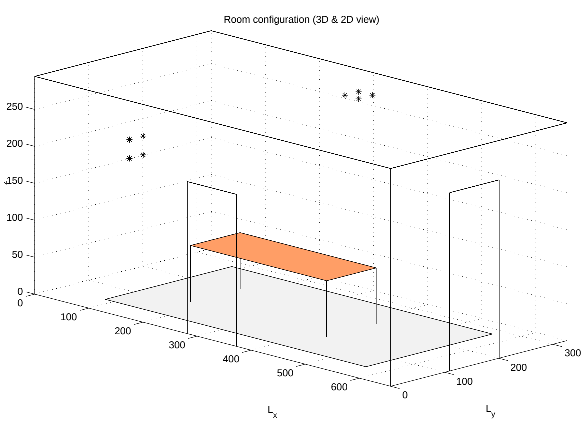

Figure 1: AVIARY (dimensions in cm). + - microphones)

In principle, the identification can be performed by minimizing the following least squares criterion,

$$V(\boldsymbol{\theta}) = \frac{1}{N}\sum_{t=0}^{N-1} \|\boldsymbol{y}(t) - \begin{bmatrix} H(q;\boldsymbol{r}_s,\boldsymbol{r}_m^1) \\ \vdots \\ H(q;\boldsymbol{r}_s,\boldsymbol{r}_m^M)\end{bmatrix} u(t)\|^2 \quad (13)$$

with respect to

$$\boldsymbol{\theta} = [\omega_1 \ \cdots \ \omega_L \ \zeta_1 \cdots \zeta_L \ \mu_1^1 \cdots \ \mu_L^M]^T, \quad (14)$$

where $H(q;\boldsymbol{r}_s,\boldsymbol{r}_m^k)$ are the transfer functions defined in (8), $\mu_n^k = \xi_n b_n(\boldsymbol{r}_s)c_n(\boldsymbol{r}_m^k)$, and $u(t)$ denote the known input signal (i.e. the swept sinusoid). Since the source location and the microphone location are known, we could also parameterize the transfer functions using the assumed modes. However, at this stage we believe that such precise modeling complicates the matter. Therefore, the "unstructured" parameters μ_n^k are introduced. Note that the least squares criterion (13) includes M experimentally observed responses. The transfer functions chosen to match these M responses are so parameterized that they all have common poles. It should be emphasized that this method is different from fitting a transfer function to each response separately. There are several pole-zero cancellations for a given location of the speaker and the microphone. This means that only one experimentally observed response is incapable of capturing the room acoustical dynamics. The method suggested in this paper overcomes this difficulty by firstly choosing a suitable parameterization and secondly using data from several responses.

3.1 Initialization

Since the criterion (13) is non-linear in the parameters of interest, we have to deal with non-linear optimization, and it becomes important to obtain an accurate "guesstimate" to initialize the non-linear optimization. For that purpose, we first assume the following "black-box" model structure:

$$\boldsymbol{y}(t) + a_1\boldsymbol{y}(t-1) + \cdots + a_{2L}\boldsymbol{y}(t-2L) = \boldsymbol{B}_1 u(t-1) + \cdots + \boldsymbol{B}_{2L}u(t-2L) + \boldsymbol{E}(t), \quad (15)$$

where $\boldsymbol{B}_k \in \mathbb{R}^{M\times 1}$, and $a_k \in \mathbb{R}^1$. The parameters $\boldsymbol{\theta}_{bb} = [a_1 \ \cdots a_{2L} \ \boldsymbol{B}_1^T \ \cdots \boldsymbol{B}_{2L}]^T$ can then be estimated as the minimizing argument of

$$V_{bb}(\boldsymbol{\theta}_{bb}) = \frac{1}{T}\sum_{t=0}^{T-1} \|\boldsymbol{E}(t)\|^2. \quad (16)$$

The advantage with the model structure (15) is that this criterion is linear in $\boldsymbol{\theta}_{bb}$. The initial estimate of ω_n and ζ_n can then be computed from

$$\hat{\zeta}_n = \frac{1}{\sqrt{1+\frac{\hat{\phi}_n^2}{\log^2(\hat{r}_n)}}} \quad (17)$$

$$\hat{\omega}_n = -\frac{\log(\hat{r}_n)}{\hat{\zeta}_n F_s}, \quad (18)$$

where the n^{th} root of $(1 + \hat{a}_1 z^{-1} + \cdots + \hat{a}_{2L}z^{-2L}) = 0$ is denoted as $\hat{p}_n = \hat{r}_n e^{j\hat{\phi}_n}$. The so obtained values of $\hat{\zeta}_n$ and $\hat{\omega}_n$ are then used as an initial guess when optimizing the criterion (13). To further simplify the problem, note that (13) is linear in the parameters μ_n^k. The cost-function $V(\boldsymbol{\theta})$ can hence be concentrated with respect to μ_n^k, $n = 1, \cdots, L$, $k = 1, \cdots, M$, leaving us with the problem of minimizing $V(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}_c = [\omega_1 \ \cdots \ \omega_L \ \zeta_1 \cdots \zeta_L]$.

3.2 Model Order

Let us now investigate how to choose the model order L. For that purpose, we used the first 2000 samples of the data to estimate $\boldsymbol{\theta}_{bb}$. We then use the next 2000 samples for validation. To determine the model order, we compute the simulated output as

$$\begin{aligned} \boldsymbol{y}^s(t) + \hat{a}_1 \boldsymbol{y}^s(t-1) + \cdots + \hat{a}_{2N} \boldsymbol{y}^s(t-2N) \\ = \hat{\boldsymbol{B}}_1 u(t-1) + \cdots + \hat{\boldsymbol{B}}_{2N} u(t-2N), \end{aligned} \quad (19)$$

for various model orders. To decide on the model order we compute the quantity $\sum_{m=1}^{M} \epsilon_m$, where

$$\epsilon_m = \frac{\sum_{t=2000}^{3999} (y_m(t) - y_m^s(t))^2}{\sum_{t=2000}^{3999} (y_m(t))^2}. \quad (20)$$

The outcome of this experiment is illustrated in Figure 2. Based on Figure 2 we conclude that L should be chosen from the interval $\{70, 90\}$.

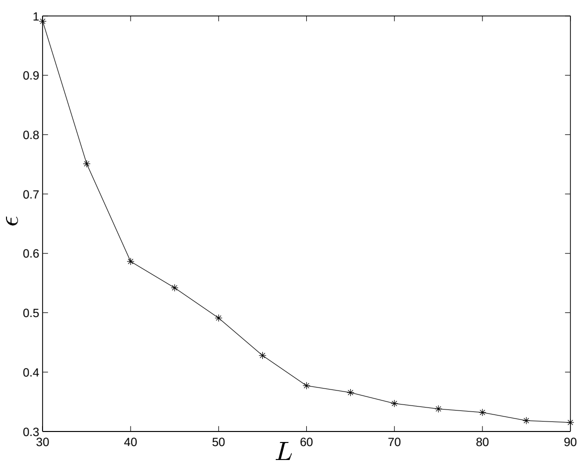

Figure 2: Evaluation of (20) for different values of L.

3.3 Experimental Data

The experiments were performed in the AVIARY shown in Figure 1. The dimensions of the AVIARY, i.e., $L_x \times L_y \times L_z$, are $6.6m \times 3.3m \times 2.9m$. The loudspeaker is located at $\boldsymbol{r}_s = \{1.7m, 1.6m, 1.0m\}$, and is facing the microphones on the left wall in the $\{y, z\}$ plane. There are eight microphones in the room (marked by pluses) but only the four on the left hand side of the room are used for these experiments. The response was recorded between the speaker and these four microphones. In the work in this paper it's assumed that the transfer function of the microphone can be modeled by a constant gain in the frequency range of interest. In Figure 3 we illustrate the estimated transfer functions for $L = 70$, where the estimated transfer functions are obtained by minimizing (13), using $\hat{\boldsymbol{\theta}}_{bb}$ as the starting point. Clearly, we obtain quite a good fit between the experimental data and the identified model.

Due to the non-linear nature of the optimization problem the identification process is a bit tedious. With the knowledge of the physics of the problem and our initialization method we were able to get a very good match but this is only the first step. The second and the crucial step is the ability of our identified model in extrapolating room transfer functions for locations at which no experimental data is available. The next section discusses the extrapolation problem.

4 Extrapolation

Even if we obtained a good fit in Figure 3, it is important to remember that the involved model orders are quite high, and the minimization of the criterion (13) can easily end up in a local minimum. Primarily as a means for validating the estimated model, but also for validation of the theoretical model (8), we will next attempt to extrapolate the estimated transfer functions.

4.1 Fitting a function to μ_n^k

Based on the assumed modes solution we can assume that for every mode ω_n and a fixed x-coordinate, the terms μ_n^k can be written as:

$$\underbrace{\begin{bmatrix} \mu_n^1 \\ \vdots \\ \mu_n^{M-1} \end{bmatrix}}_{\boldsymbol{\mu}_n^{1:M-1}} = C \cos(\beta_n y_s) \cos(\gamma_n z_s) \times \begin{bmatrix} \cos(\beta_n y_m^1) \cos(\gamma_n z_m^1) \\ \vdots \\ \cos(\beta_n y_m^{M-1}) \cos(\gamma_n z_m^{M-1}) \end{bmatrix}. \quad (21)$$

Here C, β_n and γ_n are unknown constants. Note that we have only included the first $M-1$ microphones in the above relationship (21) since our aim is to extrapolate these terms for the microphone M. In a practical scenario, we have an estimate $\hat{\boldsymbol{\mu}}_n^{1:M-1}$, from the identification step in the previous section, at our disposal, but the values β_n and γ_n are generally unknown. We next propose to estimate these quantities as

$$\{\hat{\beta}_n, \hat{\gamma}_n\} = \arg \min_{\beta_n, \gamma_n} \|\hat{\boldsymbol{\mu}}_n^{1:M-1} - \frac{1}{\boldsymbol{\mu}_n^T \boldsymbol{\mu}_n} \boldsymbol{\mu}_n \boldsymbol{\mu}_n^T \hat{\boldsymbol{\mu}}_n^{1:M-1}\|^2, \quad (22)$$

where

$$\boldsymbol{\mu}_n = \boldsymbol{\mu}_n(\beta_n, \gamma_n) = \cos(\beta_n y_s) \cos(\gamma_n z_s) \times \begin{bmatrix} \cos(\beta_n y_m^1) \cos(\gamma_n z_m^1) \\ \vdots \\ \cos(\beta_n y_m^{M-1}) \cos(\gamma_n z_m^{M-1}) \end{bmatrix}. \quad (23)$$

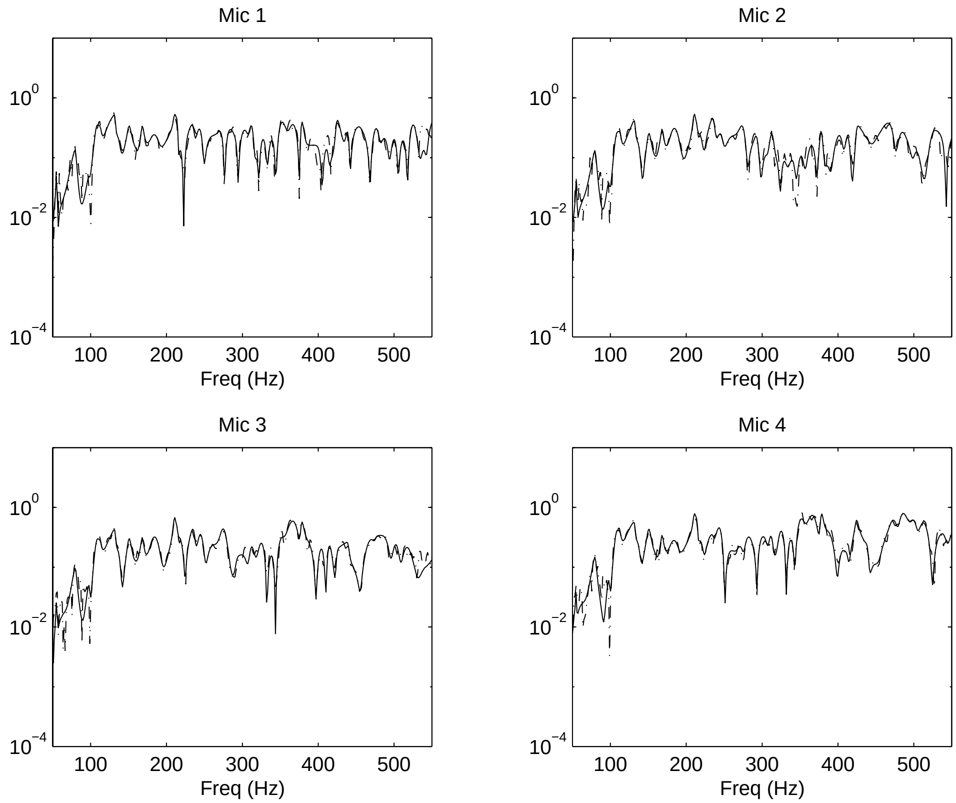

Figure 3: Magnitude of estimated transfer functions. Solid line: model based estimate. Dash-dot: Fourier-based estimate.

The above minimization is performed subject to $\beta_n = \frac{j\pi}{L_y}$, and $\gamma_n = \frac{i\pi}{L_y}$, where i and j are positive integers. Hence, in practice we have to perform a grid-search over a set of feasible integers i and j. The scalar C can subsequently be estimated as

$$\hat{C} = \frac{\boldsymbol{\mu}_n^T(\hat{\beta}_n, \hat{\gamma}_n)\hat{\boldsymbol{\mu}}_n^{1:M-1}}{\boldsymbol{\mu}_n^T(\hat{\beta}_n, \hat{\gamma}_n)\boldsymbol{\mu}_n(\hat{\beta}_n, \hat{\gamma}_n)}. \qquad (24)$$

Now, if the estimated transfer functions are accurate, we should be able to use the above findings for extrapolating the transfer function at a position $\boldsymbol{r}_m^M$. More precisely, let

$$\tilde{\mu}_n^M = \hat{C}\cos(\hat{\beta}_n y_s)\cos(\hat{\gamma}_n z_s)\cos(\hat{\beta}_n y_m^M)\cos(\hat{\gamma}_n z_m^M). \qquad (25)$$

4.2 Experimental Data

To illustrate the extrapolation procedure we consider Figure 4, which shows the microphone arrangement on the left wall ($\{y, z\}$ plane) of the AVIARY shown in Figure 1. Here, $y_0 = 1.65$, $z_0 = 1.55$, and $\Delta y = \Delta z = 0.25$ (all measures are in meters). Our task is now to apply the above procedure for extrapolating the frequency response of the upper right microphone, using the estimated responses at the remaining three positions.

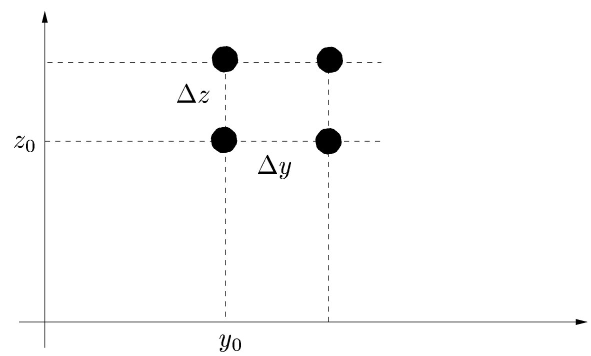

Figure 4: Microphone arrangement.

In the first set of experiments it was difficult to get good extrapolations with the data obtained from only three microphones.

Right now we can only speculate about the sources for errors. A likely problem is that the chosen model order ($L = 70$) is not sufficient to give a good fit for all frequencies. Even if the model order is not correct, we can still get a good match in the frequency domain, but the individual parameters cannot be given the proper physical interpretation. A second pitfall is that our optimization procedure got stuck in a local minimum. However, the most probable explanation is that the ideal conditions assumed to obtain the analytical solution to the wave equation are not completely true. For instance, when we performed our experiments, a table along with some standard office equipment was present in the room. The wave equation solution assumes a perfectly rectangular room, with no obstructing objects. Continuing from the very encouraging results of this paper, our research is presently concentrating on improving the proposed method in view of the afore mentioned limitations.

5 Conclusions

Applications in need of acoustical modeling is increasing. The theoretical treatment given in this paper serves as a framework for designing new powerful signal processing methods. Of special interest in this contribution was the problem of identifying the transfer function from a single source to multiple microphones.

The main contribution of this paper is the proposed procedure for simultaneous estimation of multiple acoustical transfer functions. Using real data recordings, we unfortunately had limited success with the extrapolation method. The reason for the failure is probably the large order of the involved transfer function, and also the fact that the geometry of the studied room does not match the ideal conditions assumed in the analysis.

Acknowledgement
This research was supported by the Digital Media Innovation Program of the University of California. Sponsors for this project include Sony Electronics, Compaq Computers, and the State of California.

References

[1] Jont B. Allen and A. Berkeley. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. America*, 65(4):943–950, April 1979.

[2] S. Furui and M. M. Sondhi. *Advances in Speech Signal Processing.* Marcel Dekker, New York, 1991.

[3] Yoichi Haneda, Yutaka Kaneda, and Nobuhiko Kitawaki. Common-acoustical-pole and residue model and its application to spatial interpolation and extrapolation of a transfer function. *IEEE Transactions on Speech and Audio Processing*, 7(6):709–717, November 1999.

[4] Heinrich Kuttruff. *Room Acoustics.* John Wiley, 1973.

[5] Heewon Lee and Byung-Ho Lee. An efficient algorithm for the image model technique. *Applied Acoustics*, 24:87–115, 1988.

[6] Philip M. Morse and K. Uno Ingard. *Theoretical Acoustics.* McGraw-Hill Book Company, Sydney, 1968.

[7] P. A. Nelson and S. J. Elliott. *Active Control of Sound.* Academic Press, London, 1992.

[8] H. R. Pota, J. Vance, T. Gustafsson, and M. M. Trivedi. Transfer functions for intelligent conference rooms. In *Submitted to the IEEE International Conference on Systems, Man, and Cybernetics*, Nashville, TN, 8–11 Oct 2000.

[9] S. Strom, A. Krokstad, and S. Sordal. Fifteen years experience with computerized ray tracing. *Applied Acoustics*, 16:291–312, 1983.

[10] M. M. Trivedi, B. Rao, and K. C. Ng. Camera networks and microphone arrays for video conferencing. In *Proceedings of Multimedia Systems and Applications Conference*, Boston, September 1999.