

Common-Acoustical-Pole and Residue Model and Its Application to Spatial Interpolation and Extrapolation of a Room Transfer Function

Yoichi Haneda, *Member, IEEE*, Yutaka Kaneda, *Member, IEEE*, and Nobuhiko Kitawaki, *Member, IEEE*

Abstract—A method is proposed for modeling a room transfer function (RTF) by using common acoustical poles and their residues. The common acoustical poles correspond to the resonance frequencies (eigenfrequencies) of the room, so they are independent of the source and receiver positions. The residues correspond to the eigenfunctions of the room. Therefore, the residue, which is a function of the source and receiver positions, can be expressed using simple analytical functions for rooms with a simple geometry such as rectangular. That is, the proposed model can describe RTF variations using simple residue functions. Based on the proposed common-acoustical-pole and residue model, methods are also proposed for spatially interpolating and extrapolating RTF's. Because the common acoustical poles are invariant in a given room, the interpolation or extrapolation of RTF's is reformulated as a problem of interpolating or extrapolating residue values. The experimental results for a rectangular room, in which the residue values are interpolated or extrapolated by using a cosine function or a linear prediction method, demonstrate that unknown RTF's can be well estimated at low frequencies from known (measured) RTF's by using the proposed methods.

Index Terms—Extrapolation, interpolation, modeling, poles, residues, room transfer function.

I. INTRODUCTION

THE ROOM transfer function (RTF), which describes the sound transmission characteristics between a source and a receiver in a room, plays a very important role in acoustic signal processing and sound field control [1], [2]. For example, an acoustic echo canceller uses the estimated RTF to remove echo signals [3], [4], and an active noise controller uses inverse filters based on RTF's to reduce noise [5], [6]. Recently, a multiple-input, multiple-output sound control system has been investigated for these applications. In such a system, multiple RTF's between the sources and receivers are used. Because the RTF's strongly depend on the source and receiver positions [7], the RTF for every source-receiver configuration must

be measured and modeled when the conventional all-zero or pole/zero model is used.

We therefore previously proposed an efficient modeling method called the common-acoustical-pole and zero (CAPZ) model for multiple RTF's [8]. The common acoustical poles correspond to the resonance frequencies (eigenfrequencies) of the room. The zeros correspond to the time delay and anti-resonances. This model requires fewer variable parameters (zeros) than the conventional all-zero and pole/zero models to express the RTF's, because the common acoustical poles are common to all RTF's in the room. However, even when the CAPZ model is used, because of the complex variations in the zeros depending on the source and receiver positions, the RTF has to be measured for every source-receiver configuration. This is cumbersome. An interpolation or extrapolation technique to estimate an unknown RTF at an arbitrary position from known RTF's would thus be very attractive.

A promising approach to interpolating or extrapolating an RTF would be to use a model that can express the RTF variations as simple functions. However, the conventional model cannot do this. In this paper, we therefore propose a new RTF model that uses the common acoustical poles and their residues. In this model, the common acoustical poles correspond to the eigenfrequencies of the room, so their residues correspond to the eigenfunctions of the room. Therefore, the proposed model can express the RTF variations with simple analytical functions corresponding to the eigenfunctions for rooms with a simple geometry, such as rectangular. Furthermore, because this model corresponds to the partial fraction expansion of the CAPZ model [9], the residue values can be obtained from the CAPZ-modeled RTF.

Based on the proposed common-acoustical-pole and residue (CAPR) model, we also propose methods for interpolating and extrapolating the RTF at an arbitrary position from the known (measured) RTF's. In these methods, functions that describe residue variations (residue functions) are estimated using several measured RTF's. Then, we calculate the residue values for the target source-receiver positions from the estimated residue functions and obtain the target RTF by using the calculated residue values and the common acoustical poles.

This paper is organized as follows. Section II reviews the conventional models of the RTF. Section III explains the common-acoustical-pole and residue model. The relationship between the residue variations of the proposed model and the eigenfunctions of the room is discussed in Section IV. In

Manuscript received February 18, 1998; revised March 8, 1999. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Dennis R. Morgan.

Y. Haneda is with the Customer Premises Equipment Business Division, Nippon Telegraph and Telephone East Corporation, Tokyo 163-1441, Japan (e-mail: y.haneda@east.ntt.co.jp).

Y. Kaneda is with Cyber Space Laboratories, Nippon Telegraph and Telephone Corporation, Tokyo 180-8585, Japan.

N. Kitawaki is with the Institute of Information Science and Electronics, University of Tsukuba, Ibaraki 305-0006, Japan.

Publisher Item Identifier S 1063-6676(99)07980-8.

Section V, the proposed model is applied to the interpolation and extrapolation of an RTF, and the results for a rectangular room, discussed in Section VI, show the advantage of the proposed methods.

II. CONVENTIONAL MODELS OF THE ROOM TRANSFER FUNCTION

In this section, we consider whether the conventional models can effectively model RTF variations caused by changes in the source and receiver positions. The typical RTF model is a conventional all-zero model (moving-average (MA) model) [1]. This model has coefficients corresponding to the truncated impulse response of the RTF; it represents the RTF with either MA coefficients $h_{AZ}(i, \mathbf{r}_s, \mathbf{r}_o)$ or zeros $q_{AZi}(\mathbf{r}_s, \mathbf{r}_o)$:

$$\begin{aligned} H_{AZ}(z, \mathbf{r}_s, \mathbf{r}_o) &= \sum_{i=0}^{Q-1} h_{AZ}(i, \mathbf{r}_s, \mathbf{r}_o) z^{-i} \\ &= C_{AZ} z^{-Q_1} \prod_{i=1}^{Q_2} [1 - q_{AZi}(\mathbf{r}_s, \mathbf{r}_o) z^{-1}] \end{aligned} \quad (1)$$

where $H_{AZ}(z, \mathbf{r}_s, \mathbf{r}_o)$ represents the all-zero modeled RTF, \mathbf{r}_s and \mathbf{r}_o represent the position vectors of the source and receiver, C_{AZ} is a gain constant, and $Q = Q_1 + Q_2 + 1$ is the number of coefficients. Coefficient $h_{AZ}(i, \mathbf{r}_s, \mathbf{r}_o)$ represents the amplitude of the direct or reflected sound at discrete time i measured for the source-receiver positions $(\mathbf{r}_s, \mathbf{r}_o)$. This model can be interpreted as a geometrical expression of the RTF. However, formulating the variations in the MA coefficients and the zeros is not easy because the number of reflected sounds is large; i.e., the number of coefficients Q is large.

The RTF can be theoretically expressed by using the resonance frequencies (eigenfrequencies) ω_i and their eigenfunctions $P_i(\mathbf{r})$ of the room based on the wave equation [7]

$$H(\omega, \mathbf{r}_s, \mathbf{r}_o) = C_1 \sum_{i=1}^{\infty} \frac{P_i(\mathbf{r}_s) P_i(\mathbf{r}_o) j\omega}{\omega^2 - \omega_i^2 - 2j\delta_i \omega_i + \delta_i^2} \quad (2)$$

where ω is the angular frequency, δ_i is the damping constant (corresponding to the Q -factor), and C_1 is the gain constant. The parameters ω_i and δ_i are independent of the source and receiver positions; their values are determined by the room size, wall reflection coefficients, and room shape.

Because the RTF can be represented by a rational expression, as shown in (2), it can be modeled by the conventional pole/zero model [10], [11] and represented with poles $p_{PZi}(\mathbf{r}_s, \mathbf{r}_o)$ and zeros $q_{PZi}(\mathbf{r}_s, \mathbf{r}_o)$:

$$H_{PZ}(z, \mathbf{r}_s, \mathbf{r}_o) = \frac{C_{PZ} z^{-Q_1} \prod_{i=1}^{Q_2} [1 - q_{PZi}(\mathbf{r}_s, \mathbf{r}_o) z^{-1}]}{\prod_{i=1}^P [1 - p_{PZi}(\mathbf{r}_s, \mathbf{r}_o) z^{-1}]} \quad (3)$$

where $H_{PZ}(z, \mathbf{r}_s, \mathbf{r}_o)$ represents the pole/zero modeled RTF, P is the order of the poles, and C_{PZ} is a gain constant. This model needs fewer parameters than the all-zero model. In the pole/zero model, both the poles $p_{PZi}(\mathbf{r}_s, \mathbf{r}_o)$ and zeros $q_{PZi}(\mathbf{r}_s, \mathbf{r}_o)$ are estimated so as to minimize the squared error between the measured RTF and the modeled RTF at every source and receiver position. They are thus estimated as different values for every source and receiver position although

the physical poles are invariant. Tracking the pole and zero variations is difficult because they are not independent of each other; for example, they cancel each other out in some source-receiver configurations.

We previously proposed the common-acoustical-pole and zero (CAPZ) model for RTF's. The background of this model is that the resonance frequencies ω_i and their damping factors δ_i are independent of the source and receiver positions as shown in (2). This model uses the common acoustical poles p_{Ci} which correspond to the resonance frequencies and damping factors. The CAPZ-modeled RTF $H_{CAPZ}(z, \mathbf{r}_s, \mathbf{r}_o)$ is represented as

$$H_{CAPZ}(z, \mathbf{r}_s, \mathbf{r}_o) = \frac{C_C z^{-Q_1} \prod_{i=1}^{Q_2} [1 - q_i(\mathbf{r}_s, \mathbf{r}_o) z^{-1}]}{\prod_{i=1}^P (1 - p_{Ci} z^{-1})} \quad (4)$$

where C_C is a gain constant. Comparison of (4) and (3) shows that the position-dependent poles $p_{PZi}(\mathbf{r}_s, \mathbf{r}_o)$ are replaced by the position-independent poles p_{Ci} . The zeros $q_i(\mathbf{r}_s, \mathbf{r}_o)$ in (4) are generally different from the $q_{PZi}(\mathbf{r}_s, \mathbf{r}_o)$ in (3). The common acoustical poles p_{Ci} are estimated as common values for the RTF's measured at different source-receiver positions. Because only the zeros $q_i(\mathbf{r}_s, \mathbf{r}_o)$ depend on the source-receiver positions, this model needs fewer parameters to express the RTF variations than the conventional all-zero model or the pole/zero model (where the poles are estimated as different values for each RTF). However, it is still difficult to express the zero variations as explicit functions.

III. COMMON-ACOUSTICAL-POLE AND RESIDUE MODEL

In this section, we propose a new RTF model that uses the common acoustical poles and their residues to express the RTF variations with simple functions. The basis of this model is that a room transfer function can be expressed by using the eigenfrequencies (resonance frequencies) and the eigenfunctions as shown in (2).

We consider a new RTF model in a discrete time system by referring to (2). Resonance frequencies ω_i and their damping factors δ_i are represented by the common acoustical poles p_{Ci} , as in the CAPZ model. Because (2) is a partial fraction expansion for the resonance frequencies, our proposed model can be represented by a z -transform with common acoustical poles p_{Ci} :

$$H(z, \mathbf{r}_s, \mathbf{r}_o) = \sum_{i=1}^{P/2} \left[\frac{A_i(\mathbf{r}_s, \mathbf{r}_o)}{1 - p_{Ci} z^{-1}} + \frac{A_i^*(\mathbf{r}_s, \mathbf{r}_o)}{1 - p_{Ci}^* z^{-1}} \right] \quad (5)$$

where P is the number of poles in the objective frequency band, and function $A_i(\mathbf{r}_s, \mathbf{r}_o)$ is a residue function. The superscript * denotes the complex conjugate. In this model, the common acoustical poles p_{Ci} and their residues $A_i(\mathbf{r}_s, \mathbf{r}_o)$ are generally complex numbers. We call the expression in (5) the CAPR model. Although this model does not strictly correspond to (2), we show the approximated deviation of it in Appendix. From Appendix, the residue function $A_i(\mathbf{r}_s, \mathbf{r}_o)$ can be expressed using eigenfunctions $P_i(\mathbf{r}_s)$ and $P_i(\mathbf{r}_o)$ as

$$A_i(\mathbf{r}_s, \mathbf{r}_o) = \frac{1}{2} C_{2i} P_i(\mathbf{r}_s) P_i(\mathbf{r}_o) \quad (6)$$

where C_{2i} is a constant. We verified the validity of this model by the experiments discussed in the next section.

Because this proposed model corresponds to the partial fraction expansion of the CAPZ model in (4), the specific residue value $A_i(\mathbf{r}_s, \mathbf{r}_o)$ for the i -th common acoustical pole p_{Ci} at the source and receiver positions $(\mathbf{r}_s, \mathbf{r}_o)$ can be calculated using

$$A_i(\mathbf{r}_s, \mathbf{r}_o) = \frac{C_C z^{-Q_1} \prod_{n=1}^{Q_2} [1 - q_n(\mathbf{r}_s, \mathbf{r}_o) z^{-1}]}{\prod_{\substack{n=1 \\ n \neq i}}^P (1 - p_{Cn} z^{-1})} \Big|_{z=p_{Ci}}. \quad (7)$$

Since the residue variations due to changes in the source and receiver positions are characterized by the eigenfunction of the room, as shown in (6), this model can express the RTF variations by using the expressions of the eigenfunctions. Although the eigenfunctions depend on the physical characteristics of the room, formulating the residue variations is easy when the eigenfunctions are well understood as is the case for a rectangular room. Thus, in the following section, we discuss the residue variations of our CAPR model in a rectangular room.

IV. RESIDUE VARIATIONS IN A RECTANGULAR ROOM

A. Theoretical Residue Variations

For a rectangular room, the eigenfunction $P_i(\mathbf{r}_\xi)$ ($\xi = s, o$) can be decomposed into three eigenfunctions corresponding to the x -, y -, and z -axes [7] as follows:

$$P_i(\mathbf{r}_\xi) = P_{n_x}(x_\xi) P_{n_y}(y_\xi) P_{n_z}(z_\xi) \quad (\xi = s, o) \quad (8)$$

where n_x , n_y , and n_z are integers representing the index of each eigenfunction, and i corresponds to a set of (n_x, n_y, n_z) . The eigenfunction along the u -axis ($u = x, y, z$) is expressed as

$$P_{n_u}(u_\xi) = C_{n_u} \exp(-jk_{n_u} u_\xi) + D_{n_u} \exp(jk_{n_u} u_\xi) \quad (9)$$

where C_{n_u} and D_{n_u} are constants, and k_{n_u} is a wave number expressed as

$$k_{n_u} = \frac{n_u \pi}{L_u} + j\gamma_u \quad (n_u = 0, 1, 2, \dots) \quad (10)$$

where L_u is the dimension along the u -axis of the room, and γ_u corresponds to the acoustic absorption coefficient of the walls. Based on (6) and (8), the residue $A_i(\mathbf{r}_s, \mathbf{r}_o)$ is expressed as

$$A_i(\mathbf{r}_s, \mathbf{r}_o) = \frac{1}{2} C_{2i} P_{n_x}(x_s) P_{n_y}(y_s) P_{n_z}(z_s) P_{n_x}(x_o) P_{n_y}(y_o) P_{n_z}(z_o). \quad (11)$$

For simplicity, we consider the residue variation when the source location \mathbf{r}_s is fixed and the receiver position \mathbf{r}_o is moved parallel to the x -axis. In this case, \mathbf{r}_o is a function of only x ; i.e., residue $A_i(\mathbf{r}_s, \mathbf{r}_o)$ can be represented by $A_i(x)$. Wave number k_{n_x} is replaced by k_i to allow use of the same index (i) of $A_i(x)$. Because all of the other eigenfunctions,

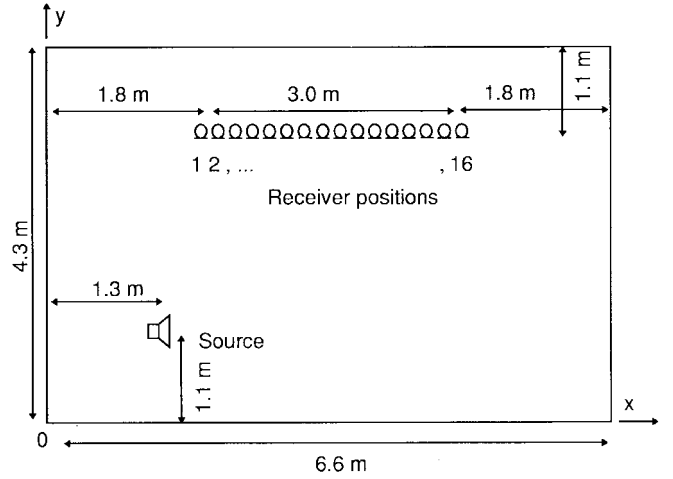


Fig. 1. Arrangement of source and receivers. The height of the room was 3.1 m, and the reverberation time was about 0.5 s.

except $P_{n_x}(x)$, can be treated as constants, the residue $A_i(x)$ can be expressed as

$$A_i(x) = C_{3i} P_{n_x}(x) = C_{4i} \exp(-jk_i x) + C_{5i} \exp(jk_i x). \quad (12)$$

The wave number k_i ($=k_{n_x}$) of residue $A_i(x)$ is a wave number of the x -axis. The wave number k_{Pi} (corresponding to the resonance frequency) of the common acoustical poles p_{Ci} is a three-dimensional (3-D) space wave number. Therefore, k_i and k_{Pi} are usually different. The relationship between k_i and k_{Pi} is

$$k_{Pi} = [k_i^2 + k_{n_y}^2 + k_{n_z}^2]^{1/2}. \quad (13)$$

Moreover, several sets of (n_x, n_y, n_z) can occasionally satisfy one k_{Pi} . That is, when the resonance frequencies degenerate, several wave numbers k_i can correspond to one 3-D space wave number k_{Pi} .

Based on [7], the residue $A_i(x)$ of (12) is a cosine function whose amplitude initially decreases away from the wall boundary. By assuming the acoustic absorption coefficient γ_u is small, we can treat the wave number k_i as a real number. In this case, the residue function $A_i(x)$ can be expressed as a simple cosine function

$$A_i(x) = B_1 \cos(k_i x) \quad (14)$$

where the constant B_1 is a complex number and $x = 0$ corresponds to the wall boundary of the room. The residue variation can thus be expressed as an explicit function, while the zero variations in the conventional common-acoustical-pole and zero model cannot.

B. Experimental Results

In practice, the residue values are calculated from measured RTF's. We calculated the residue values in a rectangular room based on our proposed CAPR modeling method to investigate the relationship between their variations and the theoretical residue variations. The room was 6.6 m (w) \times 4.3 m (d) \times 3.1 m (h) with a reverberation time of 0.5 s. The source location was fixed, and 16 receivers were set parallel to the x -axis at intervals of 20 cm (Fig. 1). In Fig. 1, the origin is

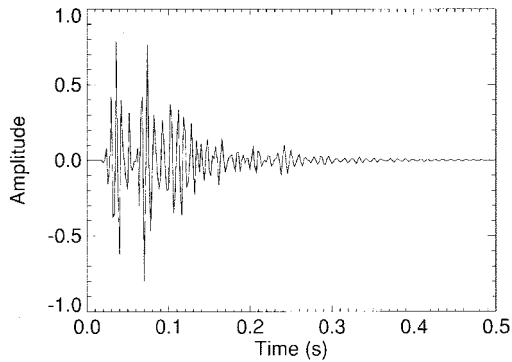


Fig. 2. Example of a measured impulse response.

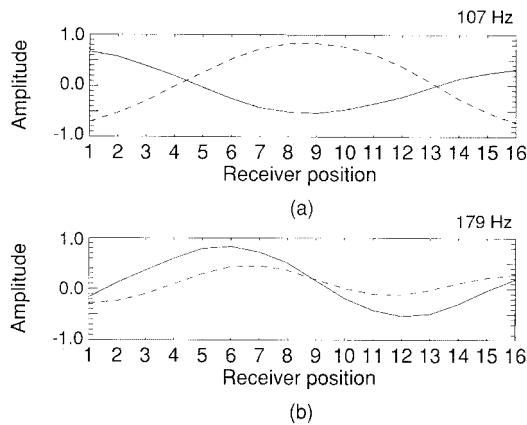


Fig. 3. Residue variations due to changes in receiver position for resonance frequencies of (a) 107 and (b) 179 Hz. The solid and dashed lines indicate the real and imaginary parts of the residues, respectively.

at the lower left corner. We numbered the receiver positions from 1 to 16 starting at the end nearest to the source. The middle point between receiver positions 8 and 9 corresponds to the center of the room along the x -axis. We measured the 16 impulse responses by using a maximum-length sequence with a period of 16 383. A loudspeaker with a diameter of 16 cm, an omnidirection microphone, and 16-b A/D and D/A converters were used for the measurements. The frequency range was limited to a low range (80 to 200 Hz), where there are not so many resonance frequencies, to avoid a high computational load. The sampling frequency was set to 500 Hz. The average signal-to-noise ratio (SNR) of the measured impulse responses was over 40 dB [12]. Fig. 2 shows an example of a measured impulse response.

First, we estimated the common acoustical poles from the measured impulse responses at seven positions: 1, 2, 3, 11, 12, 13, and 14. The number of estimated poles was 60 based on the number of resonance frequencies and their degeneration [13]. Next, we estimated the zeros in each of the 16 RTF's by using the estimated common acoustical poles. That is, we obtained sixteen CAPZ-modeled RTF's. The number of zeros was the same as the number of poles. Then, the residue values in each RTF were calculated using the partial fraction expansion of those CAPZ-modeled RTF's.

Examples of the residue variations, for resonance frequencies of 107 and 179 Hz, are shown in Fig. 3(a) and (b), respectively, with the residue values plotted continuously.

Both observed residue values varied as a cosine-like function corresponding to (14). They were symmetrical in absolute amplitude with respect to the center of the room (the point between receiver positions 8 and 9). The residue variation for the resonance frequency of 107 Hz corresponds to the $(4, 0, 0)$ mode. Although several modes degenerate at a resonance frequency of 179 Hz, the observed residue variation corresponds to the $(5, n_y, n_z)$ mode. That is, even if the poles degenerate, the residue variations can be obtained at one particular mode (a particular function). When the resonance frequencies degenerate, the correspondence between a particular observed function and the theoretical eigenfunction requires further analysis. Nevertheless, these experimental results show that our proposed CAPR modeling method can express RTF variations as simple residue variations corresponding to the eigenfunctions.

V. PRINCIPLE OF INTERPOLATION AND EXTRAPOLATION OF RTF'S

Because the common acoustical poles do not depend on the source and receiver positions, the RTF variations can be expressed by the residue variations in our proposed model. Therefore, interpolating or extrapolating the RTF can be reformulated as a problem of interpolating or extrapolating residue functions. That is, it becomes a problem of estimating the residue functions.

We will discuss interpolation and extrapolation assuming a rectangular room, because the eigenfunctions of a rectangular room are well understood. For such a room, we need to estimate only the parameters of the eigenfunction. Although the room shape is simple, it provides a good approximation for many rooms, especially at low frequencies. Also, we assume that the source location is fixed and the receiver moves parallel to the x -axis which simplifies the residue variation as discussed in Section IV.

The proposed interpolation method is outlined in Fig. 4. In this figure, the impulse response $\mathbf{h}(x_{\text{IN}})$ at receiver position x_{IN} is interpolated by using, as an example, the four impulse responses $\mathbf{h}(x_1)$ to $\mathbf{h}(x_4)$ measured at x_1 to x_4 . The number of impulse responses is required to exceed the number of parameters in the residue functions. First, the common acoustical poles p_{Ci} are estimated from the measured impulse responses. Next, each RTF is CAPZ modeled by using the estimated poles p_{Ci} . The residue values $A_i(x_m)$ ($i = 1, 2, \dots, P; m = 1, 2, 3, 4$) are calculated using the partial fraction expansion of the CAPZ-modeled RTF's as shown in (7). Then the parameters of the residue functions are determined based on the calculated residue values $A_i(x_m)$ at the four positions. The residue function $\hat{A}_i(x)$ is thus expressed by this parametric model. Residue value $\hat{A}_i(x_{\text{IN}})$ at receiver position x_{IN} is calculated by evaluating the estimated residue function $\hat{A}_i(x)$ for $x = x_{\text{IN}}$. These steps are repeated for all i ($i = 1, 2, \dots, P$). Finally, using all of the estimated residue values $\hat{A}_i(x_{\text{IN}})$ ($i = 1, 2, \dots, P$) and the common acoustical poles p_{Ci} ($i = 1, 2, \dots, P$), we obtain the interpolated impulse response $\hat{\mathbf{h}}(x_{\text{IN}})$ at x_{IN} . Extrapolation of the RTF can be done in a similar manner.

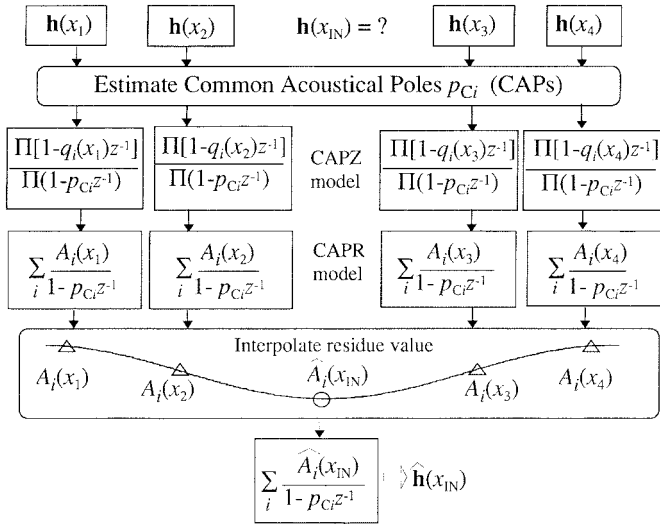


Fig. 4. In this example, RTF $\hat{h}(x_{IN})$ is estimated from RTF's $h(x_1)$ to $h(x_4)$ by using the proposed interpolation method. First, the common acoustical poles p_{Ci} are estimated from RTF's $h(x_1)$ to $h(x_4)$. The actual residue values of $A_i(x_m)$ are obtained by using the partial fraction expansion of the CAPZ-modeled RTF's (CAPR modeling). The residue values $\hat{A}_i(x_{IN})$ ($i = 1, 2, \dots, P$) at receiver position x_{IN} are estimated using the actual residue values of $A_i(x_m)$ ($i = 1, 2, \dots, P; m = 1, 2, 3, 4$) and an interpolation method. The impulse response $\hat{h}(x_{IN})$ at receiver position x_{IN} is obtained based on the common acoustical poles p_{Ci} and estimated residue values $\hat{A}_i(x_{IN})$.

As shown in the previous section, the residue variations are cosine-like functions in a rectangular room. Therefore, we propose using a simple cosine function or a linear prediction method to estimate the residue values at the target position. The details of these parameter estimation methods are explained below.

A. Residue Function Estimation as a Cosine Function

When the acoustic absorption coefficients of the walls are small, wave number k_i can be treated as a real number. In this case, the real and imaginary parts of the residue functions can be approximated by cosine functions

$$\hat{A}_i(x) = \hat{A}_{Rei}(x) + j\hat{A}_{Imi}(x) \quad (15)$$

$$\hat{A}_{\gamma i}(x) = B_{\gamma 1} \cos(k_i x) + B_{\gamma 2} \quad (\gamma = \text{Re}, \text{Im}). \quad (16)$$

The real and imaginary parts of the residue function are assumed to be independent, and $B_{\gamma 2}$ is used to remove any bias components.

We consider the determination of the parameters of cosine function $\hat{A}_{\gamma i}(x)$ in (16) based on the residue values $A_{\gamma i}(x_m)$ ($m = 1, 2, \dots, M$) observed at M receiver positions. We assume that the relative position Δx_m from the first receiver position x_1 is known for each receiver position, but x_1 is unknown. Moreover, we assume that the room size is unknown. Because $x_m = x_1 + \Delta x_m$, the approximated residue function can be represented as

$$\begin{aligned} \hat{A}_{\gamma i}(x_m) &= B_{\gamma 1} \cos[k_i(x_1 + \Delta x_m)] + B_{\gamma 2} \\ &= B_{\gamma 1} \cos(k_i x_1) \cos(k_i \Delta x_m) \\ &\quad - B_{\gamma 1} \sin(k_i x_1) \sin(k_i \Delta x_m) + B_{\gamma 2}. \end{aligned} \quad (17)$$

The wave number of 3-D space k_{Pi} can be calculated based on the estimated common acoustical pole. However, the wave number for each axis cannot be estimated from the wave number of 3-D space k_{Ti} because the room size is unknown. That is, the wave number k_i for the x -axis in (17) is an unknown parameter. Thus, it is difficult to determine all the parameters in (17) at the same time. Therefore, by setting values from 0 Hz up to the maximum objective frequency at intervals of 1 Hz for the wave number k_i , we can determine the optimum set of wave numbers k_i and the other parameters as follows.

First, the wave number k_i is set to one value from among the objective frequencies. Then, because Δx_m is already known, the values of $\cos(k_i \Delta x_m)$ and $\sin(k_i \Delta x_m)$ can be obtained. That leaves only $B_{\gamma 1} \cos(k_i x_1)$, $-B_{\gamma 1} \sin(k_i x_1)$, and $B_{\gamma 2}$ as unknown parameters. By representing these three parameters as β_1 , β_2 , and β_3 , we can describe the relationship between the M actual (=observed) residues and these parameters as

$$\begin{pmatrix} \cos(k_i \Delta x_1) & \sin(k_i \Delta x_1) & 1 \\ \cos(k_i \Delta x_2) & \sin(k_i \Delta x_2) & 1 \\ \vdots & \vdots & \vdots \\ \cos(k_i \Delta x_M) & \sin(k_i \Delta x_M) & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} A_{\gamma i}(x_1) \\ A_{\gamma i}(x_2) \\ \vdots \\ A_{\gamma i}(x_M) \end{pmatrix}. \quad (18)$$

This is an overdetermined matrix equation. By writing it as $\mathbf{WB} = \mathbf{A}$, least-squares solutions for β_1 , β_2 , and β_3 can be calculated $\mathbf{B} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{A}$. The squared error between the actual residue values $A_{\gamma i}(x_m)$ and the residue values $\hat{A}_{\gamma i}(x_m)$ calculated by \mathbf{WB} using the least-squares solutions is expressed as

$$\varepsilon_{\gamma}(k_i) = \sum_{m=1}^M [A_{\gamma i}(x_m) - \hat{A}_{\gamma i}(x_m)]^2. \quad (19)$$

Changing the value of wave number k_i within the objective frequencies, $\mathbf{B} = [\beta_1, \beta_2, \beta_3]$, and their squared error $\varepsilon_{\gamma}(k_i)$ are calculated for each wave number k_i . The optimum set of parameters $\mathbf{B}_{\gamma}^{\text{opt}}$ and $k_{\gamma i}^{\text{opt}}$ is determined so as to minimize the squared error (19). Finally, the approximated residue function is given as

$$\begin{aligned} \hat{A}_{\gamma i}(x_1 + \Delta x) &= \beta_{\gamma 1}^{\text{opt}} \cos(k_{\gamma i}^{\text{opt}} \Delta x) + \beta_{\gamma 2}^{\text{opt}} \sin(k_{\gamma i}^{\text{opt}} \Delta x) + \beta_{\gamma 3}^{\text{opt}}. \end{aligned} \quad (20)$$

These steps are repeated for all i ($i = 1, 2, \dots, P$).

B. Residue Value Estimation Based on a Linear Prediction Method

The linear prediction method [14] corresponds to approximating the residue function as an exponentially increasing/decreasing cosine function. Therefore, this method can approximate the residue variations better than the simple cosine function approximation. However, the impulse responses should be measured at equal intervals.

When the receivers are set at intervals of Δ , we assume that the real and imaginary part of the residue at position

$x_m = x_0 + m\Delta$ can be expressed by using those of positions $x_{m-1} = x_0 + (m-1)\Delta$ and $x_{m-2} = x_0 + (m-2)\Delta$:

$$\hat{A}_{\gamma i}(x_m) = \alpha_{\gamma 1} \hat{A}_{\gamma i}(x_{m-1}) + \alpha_{\gamma 2} \hat{A}_{\gamma i}(x_{m-2}). \quad (21)$$

This equation is equivalent to the linear prediction equation in the time domain with prediction coefficients $\alpha_{\gamma 1}$ and $\alpha_{\gamma 2}$. The matrix formulation of (21) for $m = 3, 4, \dots, M$ is

$$\begin{pmatrix} A_{\gamma i}(x_2) & A_{\gamma i}(x_1) \\ A_{\gamma i}(x_3) & A_{\gamma i}(x_2) \\ \vdots & \vdots \\ A_{\gamma i}(x_{M-1}) & A_{\gamma i}(x_{M-2}) \end{pmatrix} \begin{pmatrix} \alpha_{\gamma 1} \\ \alpha_{\gamma 2} \end{pmatrix} = \begin{pmatrix} A_{\gamma i}(x_3) \\ A_{\gamma i}(x_4) \\ \vdots \\ A_{\gamma i}(x_M) \end{pmatrix}. \quad (22)$$

When this matrix equation is written as $\mathbf{WB} = \mathbf{A}$, the least-squares solutions can again be obtained by calculating $\mathbf{B} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{A}$, as in the previous method. By using the determined $\alpha_{\gamma 1}$ and $\alpha_{\gamma 2}$ and the known $A_{\gamma i}(x_M)$ and $A_{\gamma i}(x_{M-1})$, the residue values $\hat{A}_{\gamma i}(x)$ ($x = x_M + \Delta, x_M + 2\Delta, \dots$) can be estimated recursively.

Comparing these two methods, the cosine function approximation can be used for unequal intervals between receiver positions and for both interpolation and extrapolation, although it ignores the damping effect. In contrast, the linear prediction method requires that the receiver positions be set at equal intervals. Moreover, when there are few measured RTF's, it can only be used for extrapolation. Nevertheless, it enables increases or decreases in the amplitude of the residue function to be expressed, and it requires fewer operations to estimate the parameters than does the cosine function approximation.

VI. INTERPOLATION AND EXTRAPOLATION EXPERIMENTS

We interpolated and extrapolated unknown RTF's from measured RTF's by using our proposed methods. The impulse responses used for these experiments were the same as those described in Section IV. To evaluate the effectiveness of our method, we compared the estimation error with that obtained using conventional methods. A linear interpolation method was used as a conventional interpolation method. Since there is no specific conventional extrapolation method, we used the RTF of the nearest known position as a conventional method for comparison.

A. Interpolation

First, we estimated the residue function $\hat{A}_i(x)$ by using the seven actual residue values $A_i(x_m)$ ($m = 1, 2, 3, 11, 12, 13$, and 14) for each i ($i = 0, 1, \dots, 59$). The actual residue values $A_i(x_m)$ and the common acoustical poles p_{Ci} were calculated based on the proposed CAPR modeling method from the measured RTF's. Here, the number of common acoustical poles was set to 60. The residue function $\hat{A}_i(x)$ was approximated as a simple cosine function, and the parameters of the approximated residue function were estimated using the method described in the previous section.

Examples of the estimated residue function $\hat{A}_i(x)$ and actual residue variations $A_i(x_m)$ corresponding to pole frequencies of 107 and 179 Hz are shown in Fig. 5(a) and (b), respectively.

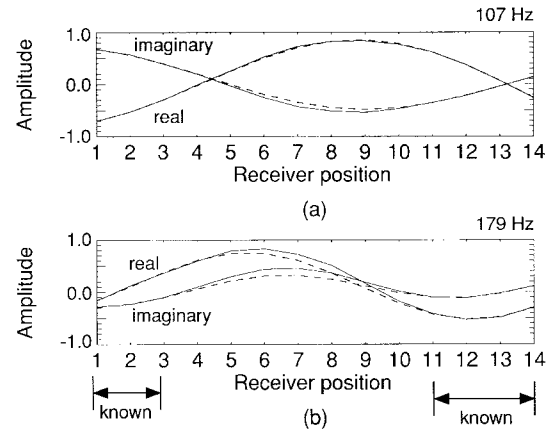


Fig. 5. Examples of actual residue variations and estimated residue functions for resonance frequencies of (a) 107 and (b) 179 Hz. The solid lines indicate the real and imaginary parts of the actual residue variations. The dashed lines indicate the real and imaginary parts of the estimated residue functions. The estimated residue functions were estimated from the actual residue values at receiver positions 1, 2, 3, 11, 12, 13, and 14.

In these figures, the actual residue variations were obtained by continuously plotting the actual residue values $A_i(x_m)$ calculated from the RTF's measured at all receiver positions 1, 2, \dots , and 14. The estimated residue functions (dashed lines) agree closely with the actual residue variations (solid lines).

Next, we interpolated the RTF at receiver position 7 by using the estimated residue functions $\hat{A}_i(x)$ and the common acoustical poles p_{Ci} . This corresponds to interpolating the RTF at receiver position 7 by using the RTF's at receiver positions 1, 2, 3, 11, 12, 13, and 14. Fig. 6(a) and (b) shows the frequency responses of the actual RTF, the RTF interpolated using the proposed method, and the RTF interpolated using the conventional linear interpolation method. Although receiver position 7 was 80 cm from both positions 3 and 11, the RTF interpolated using our proposed method agreed well with the actual RTF. In contrast, the RTF interpolated by the conventional linear interpolation method using complex values of the RTF's at receiver positions 3 and 11 had large errors.

We also compared the actual RTF and interpolated RTF in the time domain. The impulse response of our interpolated RTF was derived from the inverse z -transform of (5). As shown in Fig. 7(a) and (b), the impulse responses of the RTF interpolated using the proposed method also agreed closely with that of the actual RTF. To show the effectiveness of our method, the impulse responses of the actual RTF's at receiver positions 3 and 11, and the impulse responses of the RTF's at receiver position 7, which were interpolated by using the conventional interpolation method from the RTF's at the receiver position 3 and 11 are shown in Fig. 7(c), (d), and (e).

Next, we investigated the relationship between the interpolation distance and the time domain estimation error. Here, the interpolation distance is the distance between the position of the interpolated RTF and the nearest receiver among the known RTF's. The error power was defined as

$$\text{Error Power} = 10 \log_{10} \frac{\sum_{n=0}^{L-1} [h(n) - \hat{h}(n)]^2}{\sum_{n=0}^{L-1} h^2(n)} \quad (\text{dB}) \quad (23)$$

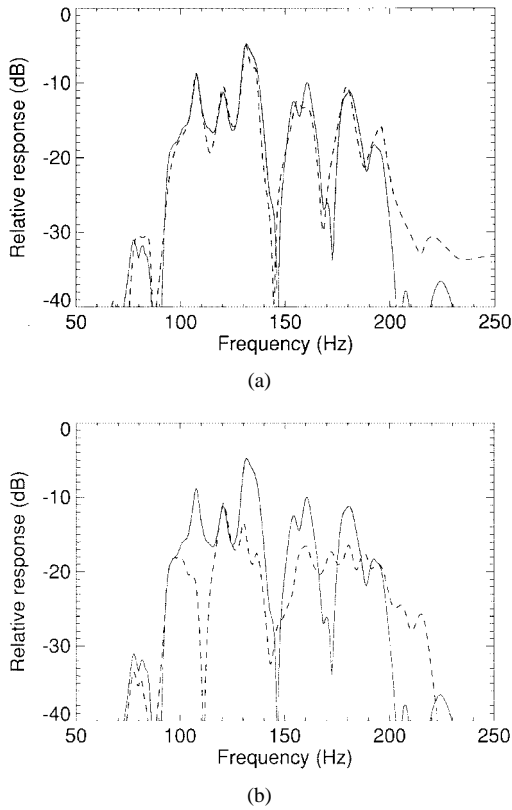


Fig. 6. Frequency responses of (a) an actual RTF (solid line) and an RTF interpolated using the proposed method (dashed line), and (b) an actual RTF (solid line) and an RTF interpolated using conventional linear interpolation (dashed line) at receiver position 7.

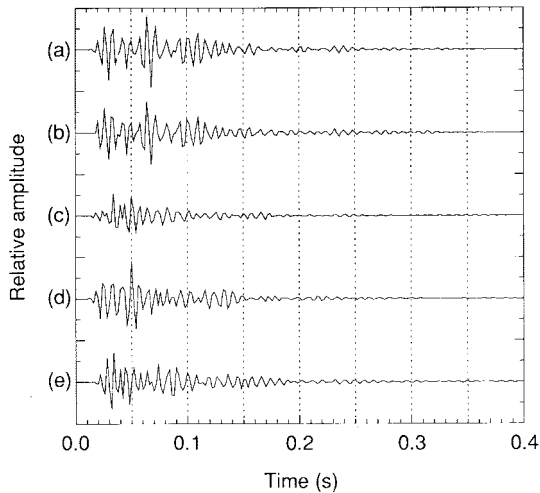


Fig. 7. Impulse responses of (a) an actual RTF at receiver position 7, (b) an RTF interpolated using the proposed method at receiver position 7, (c) an actual RTF at receiver position 3, (d) an actual RTF at receiver position 11, and (e) an RTF interpolated using the conventional linear interpolation method at receiver position 7 from the RTF's at receiver positions 3 and 11.

where, $h(n)$ is the actual impulse response and $\hat{h}(n)$ is the impulse response of the interpolated RTF. We plotted the error power against the interpolation distance (Fig. 8). For comparison, we also plotted the results for the linear interpolation method. The proposed method interpolated the

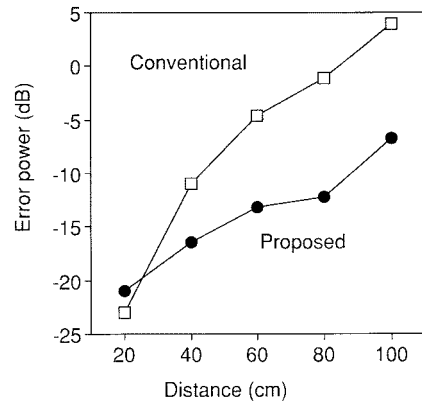


Fig. 8. Estimation errors between actual and interpolated impulse responses for various interpolation distances. The interpolations were done using the proposed and conventional linear interpolation methods.

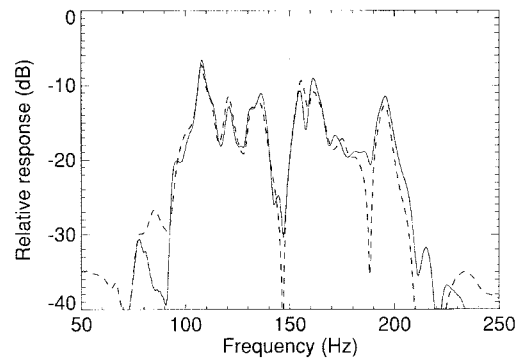


Fig. 9. Frequency responses of actual (solid line) and extrapolated (dashed line) RTF's at receiver position 9 using the linear prediction method.

RTF's with better accuracy than did the conventional linear interpolation method.

B. Extrapolation

We estimated the RTF's at receiver positions 8 to 12 by using the RTF's at positions 1 to 7. Two approximate residue functions [a cosine function (16) and a linear prediction method (21)] were used to extrapolate the RTF's.

First, we estimated the common acoustical poles by using the seven known RTF's, then we calculated the residue values at receiver positions 1 to 7. The estimation conditions were the same as for the interpolation. Next, the parameters of the approximated residue function $\hat{A}_i(x)$ were determined by using the actual residue values $A_i(x_m)$ ($i = 0, 1, \dots, 59; m = 1, 2, \dots, 7$) for each method: the cosine function approximation and the linear prediction method. Finally, we obtained the extrapolated RTF's $\hat{h}(x_m)$ ($m = 8, 9, \dots, 12$) by using the common acoustical poles p_{C_i} and estimated residue values $\hat{A}_i(x_m)$ ($m = 8, 9, \dots, 12$).

Fig. 9 shows the frequency responses of the actual (solid) and extrapolated (dashed line) RTF's at receiver position 9 (40 cm from receiver position 7) when the residues were extrapolated using the linear prediction method. The peaks of the extrapolated RTF agree well with those of the actual RTF. The large amplitude estimation error (the dip) at about 190 Hz was caused by a misestimation of the residue function.

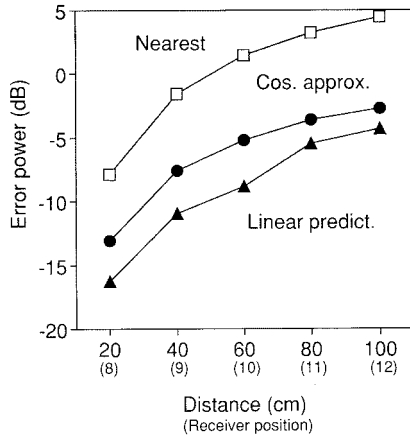


Fig. 10. Estimation error between actual and extrapolated impulse responses at each receiver position. The extrapolations were done using a cosine function, a linear prediction method, and the nearest RTF (i.e., the RTF at receiver position 7).

The error powers of the RTF's extrapolated using the cosine function and linear prediction method at each receiver position are shown in Fig. 10. The error power when the nearest RTF (that is, the RTF measured at receiver position 7) was used as the estimated RTF for all receiver positions is shown for reference. These results show that the errors for all extrapolation methods decreased as the distance increased, but both proposed methods had lower error power than when using the nearest RTF. To quantify the relationship between the estimation error and the interpolation or extrapolation distance, and the effective frequency range of the proposed method, will need further investigation under various conditions.

VII. CONCLUSION

We have proposed using the common acoustical poles and their residues to model the room transfer function. This model corresponds to the theoretical expression of the room transfer function, which is based on the wave equation. The common acoustical poles correspond to the resonance frequencies (eigenfrequencies) of the room; they are independent of the source and receiver positions. The residues correspond to the eigenfunctions in the room. Therefore, our proposed model can express the RTF variations by using analytical residue functions corresponding to the eigenfunctions for rooms with a simple geometry, such as rectangular.

We also proposed methods for interpolating and extrapolating RTF's by using known (measured) RTF's based on our proposed model. The residue variation can be approximated as a cosine function or a linear prediction method in a rectangular room when the source location is fixed and the receiver moves parallel to one axis. The parameters of the approximated residue functions were estimated from the actual residues, which were calculated from the measured impulse responses. The room transfer functions were then interpolated and extrapolated based on the estimated residue values and the common acoustical poles. Our experiments showed that at low frequencies the proposed interpolation and extrapolation methods have much smaller errors than conventional methods.

The proposed common-acoustical-pole and residue model thus provides a promising approach to interpolating and extrapolating RTF's. Furthermore, we expect that our proposed model can be applied to simulations of continuous RTF variations caused by movement of the source.

APPENDIX

APPROXIMATED DEVIATION METHOD OF THE CAPR MODEL

Equation (2) can be rewritten in the s -plane (the Laplace transform domain) with the limitation of the frequency range as

$$H_1(s, \mathbf{r}_s, \mathbf{r}_o) = \sum_{i=1}^{P/2} \left[\frac{A_i(\mathbf{r}_s, \mathbf{r}_o)}{s - s_i} + \frac{A_i(\mathbf{r}_s, \mathbf{r}_o)}{s + s_i} \right] \quad (24)$$

where $s = j\omega$, $s_i = j\omega_i - \delta_i$, and $P/2$ is the number of resonance frequencies in the objective frequency band. From (2) and (24), function $A_i(\mathbf{r}_s, \mathbf{r}_o)$ is expressed using eigenfunctions $P_i(\mathbf{r}_s)$ and $P_i(\mathbf{r}_o)$ as

$$A_i(\mathbf{r}_s, \mathbf{r}_o) = \frac{1}{2} C_{2i} P_i(\mathbf{r}_s) P_i(\mathbf{r}_o). \quad (25)$$

The poles in (24) correspond to the double-sided waveform of sound pressure in the space domain. So, the transfer function should be represented as a causal transfer function in the time domain.

Now, (24) can be rewritten as follows:

$$H_1(s, \mathbf{r}_s, \mathbf{r}_o) = \sum_{i=1}^{P/2} \left\{ \frac{A_i(\mathbf{r}_s, \mathbf{r}_o)}{s - s_i} - \left[\frac{A_i^*(\mathbf{r}_s, \mathbf{r}_o)}{s - s_i^*} \right]^* \right\}. \quad (26)$$

By assuming $\delta_i \ll \omega_i$, the first term on the right side in (26) contributes significantly for $\omega \approx \omega_i$, and the second term does so significantly for $\omega \approx -\omega_i$. Therefore, the first and second terms on the right side do not interfere very much with each other, and can be treated separately. Since $-\left[\frac{A_i^*(\mathbf{r}_s, \mathbf{r}_o)}{s - s_i^*} \right]^*$ and $\left[\frac{A_i^*(\mathbf{r}_s, \mathbf{r}_o)}{s - s_i^*} \right]$ have the same amplitude frequency response, the second term can be substituted by $\left[\frac{A_i^*(\mathbf{r}_s, \mathbf{r}_o)}{s - s_i^*} \right]$:

$$H(s, \mathbf{r}_s, \mathbf{r}_o) = \sum_{i=1}^{P/2} \left[\frac{A_i(\mathbf{r}_s, \mathbf{r}_o)}{s - s_i} + \frac{A_i^*(\mathbf{r}_s, \mathbf{r}_o)}{s - s_i^*} \right]. \quad (27)$$

This will guarantee that the transfer function is a causal and real response in the time domain, and (24) and (27) have the same amplitude frequency response. The CAPR model can be derived by z -transforming (27) using an impulse invariant method [15].

ACKNOWLEDGMENT

The authors are grateful to Y. Nishino, J. Kojima, and S. Makino for their support and suggestions.

REFERENCES

- [1] M. Tohyama, H. Suzuki, and Y. Ando, *Acoustic Space*. New York: Academic, 1995.
- [2] P. A. Nelson and S. J. Elliot, *Active Control of Sound*. New York: Academic, 1993.

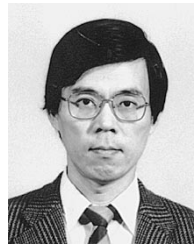
- [3] S. Furui and M. M. Sondhi, *Advances in Speech Signal Processing*. New York: Marcel Dekker, 1991.
- [4] S. Makino, Y. Kaneda, and N. Koizumi, "Exponentially weighted stepsize NLMS adaptive filter based on the statistics of a room impulse response," *IEEE Trans. Speech Audio Processing*, vol. 1, pp. 101–108, Jan. 1993.
- [5] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [6] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 145–152, Feb. 1988.
- [7] H. Kuttruf, *Room Acoustics*. New York: Elsevier, 1991.
- [8] Y. Haneda, S. Makino, and Y. Kaneda, "Common acoustical pole and zero modeling of room transfer functions," *IEEE Trans. Speech Audio Processing*, vol. 2, pp. 320–328, Apr. 1994.
- [9] Y. Haneda, "Interpolation and extrapolation of the room transfer functions based on common acoustical poles and their residues," in *Proc. IEICE General Conf.*, A-277, Mar. 1996.
- [10] J. Mourjopoulos and M. Paraskevas, "Pole and zero modeling of room transfer functions," *J. Sound Vibr.*, vol. 146, pp. 281–302, Apr. 1991.
- [11] G. Long, D. Shwed, and D. D. Falconer, "Study of a pole-zero adaptive echo canceller," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 765–769, July 1987.
- [12] Y. Kaneda, "A study of nonlinear effect on acoustic impulse response measurement," *J. Acoust. Soc. Jpn. E*, vol. 16, pp. 193–195, Mar. 1994.
- [13] Y. Haneda, S. Makino, Y. Kaneda, and N. Koizumi, "ARMA modeling of a room transfer function at low frequencies," *J. Acoust. Soc. Jpn. E*, vol. 15, pp. 353–355, Sept. 1994.
- [14] S. Haykin, *Adaptive Filter Theory*, 2nd Ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [15] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.



Yoichi Haneda (A'92–M'97) was born in Sendai, Japan, on June 17, 1964. He received the B.S., M.S., and Ph.D. degrees from Tohoku University, Sendai, in 1987, 1989, and 1999, respectively.

Since joining Nippon Telegraph and Telephone Corporation (NTT) in 1989, he has been investigating the modeling of acoustic transfer functions, acoustic signal processing and acoustic echo cancellers. He is now an Associate Manager with the Business Division of NTT-East.

Dr. Haneda is a member of the Acoustical Society of Japan and the Institute of Electronics, Information, and Communication Engineers of Japan.



Yutaka Kaneda (M'80) was born in Osaka, Japan, on February 20, 1951. He received the B.E., M.E., and Dr. Eng. degrees from Nagoya University, Nagoya, Japan, in 1975, 1977, and 1990, respectively.

In 1977, he joined the Electrical Communication Laboratory of Nippon Telegraph and Telephone Corporation (NTT), Musashino, Tokyo, Japan. He has since been engaged in research on acoustic signal processing. He is now a Senior Research Engineer at the Media processing project of the NTT Cyber Space Laboratories. His recent research interests include microphone array processing, adaptive filtering, and sound field control.

Dr. Kaneda received the IEEE ASSP Senior Award in 1990 for an article on inverse filtering of room acoustics, and paper awards from the Acoustical Society of Japan in 1990 and 1992 for articles on adaptive microphone arrays and active noise control. He is a member of the Acoustical Society of Japan, the Acoustical Society of America, and the Institute of Electronics, Information and Communication Engineers of Japan.



Nobuhiko Kitawaki (M'87) was born in Aichi, Japan, on September 27, 1946. He received the B.E.E., M.E.E., and Dr. Eng. degrees from Tohoku University, Sendai, Japan, in 1969, 1971, and 1981, respectively.

From 1971 to 1997 he was engaged in research on speech and acoustics information processing at the Laboratories in Nippon Telegraph and Telephone Corporation. From 1993 to 1997, he was Executive Manager of Speech and Acoustics Laboratory. He currently serves as Professor at the Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki, Japan. He joined to ITU-T Study Group 12 at 1981, and serves as Rapporteur from 1985.

Prof. Kitawaki is a member of the Institute of Electronics, Information, and Communication Engineers of Japan (IEICEJ), the Acoustical Society of Japan, and Information Processing Society of Japan. He received Paper Awards from IEICEJ in 1979 and 1984.