BLIND ALGORITHM FOR CALCULATING COMMON POLES BASED ON LINEAR PREDICTION

Takafumi Hikichi   *Masato Miyoshi*

NTT Communication Science Laboratories, NTT Corporation
2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237 Japan
{hikichi,miyo}@cslab.kecl.ntt.co.jp

ABSTRACT

This paper proposes a blind calculation method for the poles common to multiple signal transmission paths. In the field of room acoustics, the poles correspond to the mode frequencies that are determined by room size and shape, and they do not change when source and receiver locations change. Information on these acoustic poles is useful for many applications, including echo cancellation and sound field equalization in a room. Conventional pole estimation methods require *a priori* measurement of the room transfer functions. This paper proposes a new method for the blind calculation of the poles, where the poles are calculated solely from the observed signals. Simulation results show that the proposed algorithm provides precise estimates of the common poles.

1. INTRODUCTION

In the field of room acoustics, poles correspond to mode frequencies determined according to the room dimensions, and do not change when source and receiver locations change. Information on these acoustic poles is useful for various applications. For example, when designing an echo canceller [1], this information can be effectively utilized to reduce the number of parameters that represent echo paths, and hence to increase the convergence speed. As another example, acoustic common poles are also useful for multiple point equalization, which achieves flat sound pressure distribution [2].

This paper proposes a new method for blindly calculating the poles common to multiple transmission paths, without *a priori* measurements. This transfer function measurement is necessary in conventinal methods. Moreover, the pole estimates obtained with the proposed method are free from any influence of the zeros in the room transfer functions. These two features have not been achieved with conventional methods.

2. CALCULATION OF COMMON POLES OF MULTIPLE TRANSFER FUNCTIONS

We deal with a single-source and two-microphone system as shown in Figure 1. When the AR part that corresponds to acoustical poles is denoted as $1 - a(z)$, two transfer functions $g_1(z), g_2(z)$ are described as [1],

$$g_i(z) = \frac{h_i(z)}{1 - a(z)} \quad (i = 1, 2). \tag{1}$$

where $h_i(z)$ denotes J-th polynomials,
$a(z) = a_1 z^{-1} + a_2 z^{-2} + \cdots + a_K z^{-K}$.
We assume $h_1(z)$ and $h_2(z)$ have no common zeros [3].

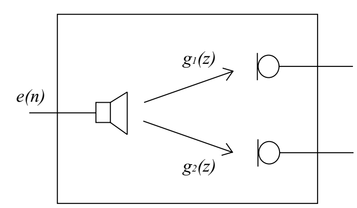


Fig. 1. A single-source two-microphone acoustic system.

2.1. Conventional method for calculating common poles

The conventional method for calculating common poles, reported in [2], is described in the following. This approach was originally derived for multiple point sound pressure equalization.

The common AR coefficients that correspond to the common poles can be estimated as those that minimize the mean squares cost function ϵ as,

$$\begin{aligned} \epsilon &= \sum_{i=1}^{M} \sum_{k=0}^{\infty} e_i^2(k), \\ &= \sum_{i=1}^{M} \sum_{k=0}^{\infty} \left[g_i(k) - \sum_{n=1}^{K} a_n g_i(k-n) \right]^2 . \end{aligned} \tag{2}$$

where
$g_i(n)$ is a measured impulse response obtained by i-th source-microphone pairs ($1 \le i \le M$),
a_n are common AR parameters,
M is the number of source-microphone pairs.

We deal with the case $M = 2$ hereafter. To minimize this cost function, all the partial derivatives of the cost function with respect to a_n must be equal to zero. Then, the common AR coefficients a_n can be expressed as

$$\mathbf{a} = (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\mathbf{v}, \tag{3}$$

where

$$\begin{aligned}
\mathbf{a} &= [a_1, a_2, \cdots, a_K], \\
\mathbf{W} &= [\mathbf{G}_1, \mathbf{G}_2]^T, \\
\mathbf{v} &= [\mathbf{g}_1, \mathbf{g}_2]^T, \\
\mathbf{g}_i &= [g_i(1), g_i(2), \cdots, g_i(N-1), 0, 0, \cdots, 0]^T, \\
\mathbf{G}_i &= \begin{bmatrix}
g_i(0) & 0 & \cdots 0 \\
g_i(1) & g_i(0) & \cdots 0 \\
\vdots & \vdots & \vdots \\
g_i(K-1) & g_i(K-2) & \cdots g_i(0) \\
\vdots & \vdots & \vdots \\
g_i(N-1) & g_i(N-2) & \cdots g_i(N-K) \\
0 & g_i(N-1) & \cdots g_i(N-K-1) \\
\vdots & \vdots & \vdots \\
0 & 0 & \cdots g_i(N-1)
\end{bmatrix}.
\end{aligned}$$

Since this method uses actual impulse responses to estimate the poles, it is necessary to measure them in advance.

2.2. Proposed method

The proposed method blindly calculates common poles without measuring transfer functions $g_1(z)$ and $g_2(z)$, assuming that a stationary white noise is input into the system. Furthermore, common poles are expected to be estimated precisely even when using fewer transfer functions than required with the conventional method.

The signals observed by the two microphones are denoted as $x_1(n)$ and $x_2(n)$. Here, we introduce the two channel linear prediction matrix $\mathbf{Q}$ that is defined by the folowing equation,

$$\mathbf{x}_n^T = \mathbf{x}_{n-1}^T\mathbf{Q}, \tag{4}$$

where

$$\mathbf{x}_n = [x_1(n), \cdots, x_1(n-m), x_2(n), \cdots, x_2(n-m)]^T.$$

By multiplying $\mathbf{x}_{n-1}$ with this equation from the left-hand side and employing expectation, the above equation can be written as,

$$E < \mathbf{x}_{n-1}\mathbf{x}_n^T >= E < \mathbf{x}_{n-1}\mathbf{x}_{n-1}^T > \mathbf{Q}, \tag{5}$$

where $E < \cdot >$ is an expectation operator.

Here, it should be noted that the acoustic system shown in Fig. 1 is equivalent to that shown in Fig. 2. In Fig. 2, signal $u(n)$ produced by AR process $1/(1 - a(z))$ is input into the two transfer functions $h_1(z)$ and $h_2(z)$. Then, signal $u(n)$ is expressed by the following equation.

$$\mathbf{u}_n = \mathbf{C}^T\mathbf{u}_{n-1} + \mathbf{e}_n, \tag{6}$$

where

$$\mathbf{u}_n = [u(n), u(n-1), ..., u(n-m-J+1)]^T,$$

$$\mathbf{C} = \begin{bmatrix}
a_1 & a_2 & \cdots & a_K & 0 & 0 \\
1 & 0 & \cdots & 0 & & \\
0 & 1 & & 0 & \vdots & \vdots \\
 & & \ddots & & \vdots & \vdots \\
\vdots & & & 1 & \vdots & \vdots \\
\vdots & & & & \ddots & 0 \\
0 & 0 & \cdots & 0 & 1 & 0
\end{bmatrix}^T, \mathbf{e}_n = \begin{bmatrix} e(n) \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$e(n)$ is stationary white noise with unit variance, and
$m + J \ge K$.

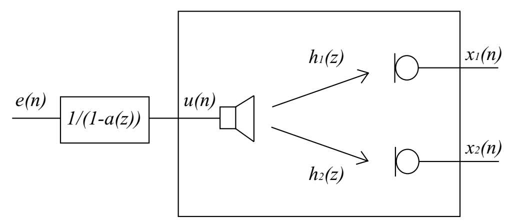


Fig. 2. An acoustic system equivalent to the system shown in Fig 1.

Then, $\mathbf{x}_{n-1}$ can be written as,

$$\mathbf{x}_{n-1}^T = \mathbf{u}_{n-1}^T\mathbf{H}, \tag{7}$$

where
$\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2]$,

$$\mathbf{H}_i = \begin{bmatrix}
h_i(0) & 0 & \cdots 0 \\
h_i(1) & h_i(0) & \cdots 0 \\
\vdots & \vdots & \vdots \\
h_i(m) & h_i(m-1) & \cdots h_i(0) \\
\vdots & \vdots & \vdots \\
h_i(J-1) & h_i(J-2) & \cdots h_i(J-m-1) \\
0 & h_i(J-1) & \cdots h_i(J-m) \\
\vdots & \vdots & \vdots \\
0 & 0 & \cdots h_i(J-1)
\end{bmatrix}.$$

It should be noted that the number of columns of matrix $\mathbf{H}$ is greater than the number of rows, i.e., $2(m+1) > m+J$. Matrix $\mathbf{Q}$ in Eq. (5) is calculated, in a minimum-mean-square-error (MMSE) sense, as follows.

$$\begin{aligned}\mathbf{Q} &= (E < \mathbf{x}_{n-1}\mathbf{x}_{n-1}^T > +\delta^2\mathbf{I})^{-1} E < \mathbf{x}_{n-1}\mathbf{x}_n^T >, \qquad (8)\\ &= (E < \mathbf{H}^T\mathbf{u}_{n-1}\mathbf{u}_{n-1}^T\mathbf{H} > +\delta^2\mathbf{I})^{-1}\\ &\quad E < \mathbf{H}^T\mathbf{u}_{n-1}\mathbf{u}_n^T\mathbf{H} >,\\ &= (\mathbf{H}^T E < \mathbf{u}_{n-1}\mathbf{u}_{n-1}^T > \mathbf{H} + \delta^2\mathbf{I})^{-1}\\ &\quad \mathbf{H}^T E < \mathbf{u}_{n-1}\mathbf{u}_{n-1}^T > \mathbf{C}\mathbf{H},\end{aligned}$$

where $\mathbf{I}$ is the unit matrix, and δ is a small positive number.

Since $E < \mathbf{u}_{n-1}\mathbf{u}_{n-1}^T >$ can be assumed to be positive-definite, it can be replaced by $\mathbf{U}^T\mathbf{U}$, where $\mathbf{U}$ is a matrix. Then, matrix $\mathbf{Q}$ can be written as,

$$\mathbf{Q} = (\mathbf{H}^T\mathbf{U}^T\mathbf{U}\mathbf{H} + \delta^2\mathbf{I})^{-1}\mathbf{H}^T\mathbf{U}^T\mathbf{U}\mathbf{C}\mathbf{H}. \qquad (9)$$

If we asumme that matrix $\mathbf{H}$ has a full row-rank, using the definition of the Moore-Penrose inverse matrix [4], we can rewrite the above equation as,

$$\begin{aligned}\mathbf{Q} &= \lim_{\delta^2\to 0}(\mathbf{H}^T\mathbf{U}^T\mathbf{U}\mathbf{H} + \delta^2\mathbf{I})^{-1}\mathbf{H}^T\mathbf{U}^T\mathbf{U}\mathbf{C}\mathbf{H},\\ &= (\mathbf{U}\mathbf{H})^+\mathbf{U}\mathbf{C}\mathbf{H},\\ &= \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{U}\mathbf{C}\mathbf{H},\\ &= \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{C}\mathbf{H}, \qquad (10)\end{aligned}$$

where $\mathbf{A}^+$ denotes the Moore-Penrose inverse matrix of $\mathbf{A}$.

The following relation holds for the non-zero eigenvalues of matrix $\mathbf{Q}$ [4],

$$\begin{aligned}\lambda(\mathbf{Q}) &= \lambda(\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{C}\mathbf{H}),\\ &= \lambda(\mathbf{H}\mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{C}),\\ &= \lambda(\mathbf{C}), \qquad (11)\end{aligned}$$

where $\lambda(\mathbf{A})$ are the non-zero eigenvalues of matrix $\mathbf{A}$. As a result, the following relationship is derived.

$$f_c(\mathbf{Q}) \equiv f_c(\mathbf{C}) \equiv 1 - a(z), \qquad (12)$$

where $f_c(\mathbf{A})$ is the characteristic polynomial of matrix $\mathbf{A}$. That is, the characteristic polynomial of matrix $\mathbf{Q}$ is equivalent to the AR polynomial $1 - a(z)$ [5]. Hence, the AR polynomial is blindly calculated from observed signals.

The proposed algorithm is summarized as follows.

1. The two-channel linear prediction matrix $\mathbf{Q}$ is calculated using output signal vectors $\mathbf{x}_n$ and $\mathbf{x}_{n-1}$ by Eq. (8).

2. The characteristic polynomial of matrix $\mathbf{Q}$ is calculated to obtain the AR polynomial $1 - a(z)$ that corresponds to the common poles.

3. SIMULATION

A simulation was performed to show the validity of the proposed algorithm. Figure 3 shows the simulation setup, where a simple sound field is surrounded by three reflectors. The transfer functions $g_i(z)$ in Eq. (1) were simulated as follows. As the MA part, $h_i(z)$, we simulated the initial reflections by the image method [6]. The AR part, $1 - a(z)$, was obtained based on the theoretical resonance frequencies calculated by $f_n = nc/2L$, where L is the distance between the two opposing walls, c is sound velocity, and n is an integer (e.g. [7]). These conditions are shown in Table 1.

Figure 4 (a) shows the impulse responses of $g_1(z)$ and $g_2(z)$, and (b) shows their frequency responses. The impulse responses are truncated at the point where the magnitude of the response seems sufficiently small. We use these two impulse responses to estimate the common AR parameters by the conventional method.

The Spectral Distortion (SD) [8], shown in the following, was used to evaluate the results.

$$SD = \sqrt{\frac{1}{F}\sum_{f=0}^{F-1}\{20\log|P(f)| - 20\log|\hat{P}(f)|\}^2}, \qquad (13)$$

where
$P(f) = 1/(1 - a(e^{jf\pi/F}))$ is the given all-pole spectrum
$\hat{P}(f) = 1/(1 - \hat{a}(e^{jf\pi/F}))$ is the estimated all-pole spectrum.

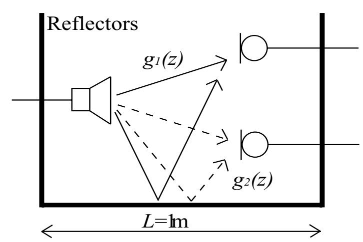


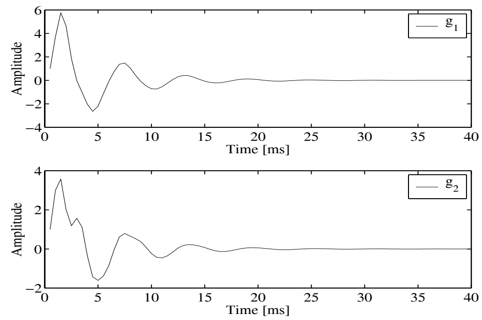
Fig. 3. Simulation setup

Table 1. Simulation conditions

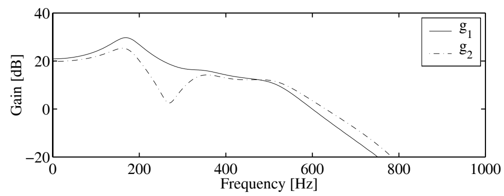
J	Order of zeros	10
K	Order of poles	6
f_1	Lowest resonance frequency	171.5 Hz
f_n	Resonance frequencies	$f_1, 2f_1, 3f_1$
N	Order of impulse responses	80
f_s	Sampling frequency	2 kHz

Figure 5 (a) shows the given and estimated all-pole spectra provided by the conventional method. Here, the AR order is set to match the actual AR order, i.e., $K = 6$. The figure shows the discrepancy between the estimated and reference responses. In this case, SD = 3.8 [dB] was obtained.

Next, the proposed calculation was tested under the same situation. We used a stationary white noise signal with the

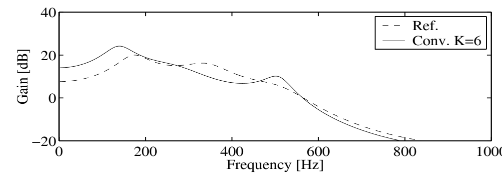


(a) Impulse responses.

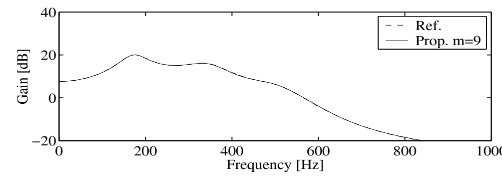


(b) Frequency responses.

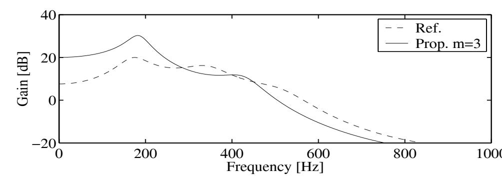
Fig. 4. Time and frequency responses of $g_1(z)$ and $g_2(z)$.



(a) Spectrum with the conventional method (SD=3.8 dB).



(b) Spectrum with the proposed method (SD=0.1 dB).



(c) Spectrum with the proposed method in the case of improper parameter setting (SD= 6.2 dB).

Fig. 5. Estimated all-pole spectra (The reference spectrum is shown by a dashed line in each figure).

normal distribution, zero mean, and unit variance. Using this random signal as the input, we estimated the AR coefficients from the two channel output signals. Since this input signal is not perfectly random, it is necessary to suppress the effect of imperfect randomness on the estimation precision. To achieve this, the estimated AR parameters were obtained by averaging over 20 repetitions.

Figure 5 (b) shows the given and estimated all-pole spectra obtained with the proposed method. The output signal length m was adjusted to 9 to meet the above mentioned requirement, i.e., $m \geq J - 1$. It is shown that the estimated spectrum agrees quite well with the actual response, and the obtained SD value is much smaller than the conventional result. In this case, the estimated AR order is $\hat{K}$=18. We have also confirmed that the similar accuracy was achieved in the case of $\hat{K} \geq 18$.

For comparison, a simulation was also performed for $\hat{K}$=6, which corresponds to the estimated order matched with the actual order. It should be noted that this does not meet the above requirement for the proposed method. The result is shown in Fig. 5 (c). In this case, the response exhibits degradation. This result suggests that the proposed method must meet the above mentioned requirement, $m \geq J - 1,\ m \geq K - J$, even when the estimated AR order is greater than the actual AR order.

4. CONCLUSIONS

This paper proposed a blind calculation method for estimating common poles based on linear prediction, where no *a priori* measurements of the transfer functions are needed. Furthermore, by utilizing the characteristic polynomial of the linear prediction matrix, the AR polynomial is directly estimated so that the pole estimation precision is unaffected by the zeros that may exist in the transfer functions. Simulation results showed that the proposed algorithm provides precise estimates of the common poles. Future work will include improvements designed to cope with severe calculation errors, for example, when the impulse responses become much longer.

5. REFERENCES

[1] Y. Haneda, S. Makino, and Y. Kaneda, "Common acoustical pole and zero modeling of room transfer functions," *IEEE Trans. SAP*, vol. 2, no. 2, pp. 320–328, 1994.

[2] Y. Haneda, S. Makino, and Y. Kaneda, "Multiple-point equalization of room transfer functions by using common acoustical poles," *IEEE Trans. SAP*, vol. 5, no. 4, pp. 325–333, 1997.

[3] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. ASSP*, vol. 36, no. 2, pp. 145–152, 1988.

[4] S. Kodama and N. Suda, *Matrix theory for systems and control*, Corona Publishing Co., 1978.

[5] M. Miyoshi, "Estimating ar parameter-sets for linear-recurrent signals in convolutive mixtures," *Proceedings of the ICA2003*, pp. 585–589, 2003.

[6] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, 1979.

[7] H. Kuttruff, *Room Acoustics*, Elsevier Applied Science, 1973.

[8] W. B. Klein and K. K. Paliwai Eds., *Speech Coding and Synthesis*, Elsevier, 1995.