PAPER Special Section on Acoustical Inverse Problems

Two-Channel Blind Deconvolution of Nonminimum Phase FIR Systems

Ken'ichi FURUYA† and Yutaka KANEDA†, *Members*

SUMMARY A new method is proposed for recovering an unknown source signal, which is observed through two unknown channels characterized by non-minimum phase FIR filters. Conventional methods cannot estimate the non-minimum phase parts and recover the source signal. Our method is based on computing the eigenvector corresponding to the smallest eigenvalue of the input correlation matrix and using the criterion with the multi-channel inverse filtering theory. The impulse responses are estimated by computing the eigenvector for all modeling orders. The optimum order is searched for using the criterion and the most appropriate impulse responses are estimated. Multi-channel inverse filtering with the estimated impulse responses is used to recover the unknown source signal. Computer simulation shows that our method can estimate non-minimum phase impulse responses from two reverberant signals and recover the source signal.

key words: *blind deconvolution, nonminimum phase, inverse filtering, impulse response, FIR filter*

1. Introduction

When a speaker is some distance away from the microphone in a teleconferencing situation, the speech signal is distorted by room reverberation, so it is less intelligible to the listeners. One way to achieve nearly perfect dereverberation of speech is to perform inverse filtering using two microphones [1]. This method requires the room impulse responses of sound transmission channels to be known in advance, but there has been no practical way to know the impulse responses between the human mouth and microphones.

A blind deconvolution method [2] based on multichannel inverse filtering has been proposed for estimating the impulse responses from the reverberant signals and recovering the source signal. The most significant problem with this method is that it is difficult to determine the order of the impulse response filter model. Wang [3] has proposed a criterion to determine the modeling order for minimum phase impulse responses, but estimating the impulse responses accurately is still difficult because room impulse responses are usually nonminimum phase.

This paper therefore proposes an approach to determining the order of the impulse response filter model and estimating impulse responses that may be nonminimum phase. This approach is based on a cost function that is minimized when the common zeros of the z-transforms of the two observed signals are extracted. If there are no common zeros between the system transfer functions of the two unknown channels, the common zeros of the observed signals represent the source signal and the noncommon zeros represent the characteristics of the two channels. The source signal can therefore be recovered by separating the common zeros from the other zeros; that is, by minimizing the cost function.

2. Principle

The proposed method consists of two stages, as shown in Figs. 1(a) and 1(b). After the impulse responses are estimated for various modeling orders, the optimum order is determined. Source recovery is done using multichannel inverse filters for the estimated impulse responses of the optimum order.

2.1 Estimation of Impulse Responses

Consider sound picked up by two microphones in a room, as shown in Fig. 1(a). Let $x(n)$ represent the sound-source signal, let $m_1(n)$ and $m_2(n)$ represent the signals received at the two microphones, and let $c_1(n)$ and $c_2(n)$ represent the impulse responses of the two acoustic paths. Signals $m_1(n)$ and $m_2(n)$ pass through FIR filters $h_2(n,i)$ and $h_1(n,i)$, respectively, where i represents the filter order. Note that the subscripts of $h_1(n,i)$ and $h_2(n,i)$ are reversed in Fig. 1(a). One of the filtered signals is subtracted from the other to generate error signal $e_a(n,i)$ for order i.

Let's assume that impulse responses $c_1(n)$ and $c_2(n)$ can be modeled using FIR filters with order j and there are no common zeroes in the z-transforms of $c_1(n)$ and $c_2(n)$. Then

$$\begin{aligned} e_a(n,i) &= m_1(n) * h_2(n,i) - m_2(n) * h_1(n,i) \\ &= x(n) * c_1(n) * h_2(n,i) - x(n) * c_2(n) * h_1(n,i) \\ &= x(n) * \{c_1(n) * h_2(n,i) - c_2(n) * h_1(n,i)\}, \end{aligned} \tag{1}$$

where the symbol * represents convolution.

If $i = j$ and $e_a(n,i) = 0$ for all n in (1), $h_1(n,i)$ and $h_2(n,i)$ satisfy

Manuscript received October 11, 1996.
Manuscript revised January 7, 1997.
† The authors are with the NTT Human Interface Laboratories, Musashino-shi, 180 Japan.

$$h_1(n,i) = \alpha\, c_1(n)$$
$$h_2(n,i) = \alpha\, c_2(n), \tag{2}$$

where α is an arbitrary constant. Thus $h_1(n,i)$ and $h_2(n,i)$ can be estimates of $c_1(n)$ and $c_2(n)$. However, since $e_a(n,i)$ does not reach zero exactly because of computation and measurement errors, we compute $h_1(n,i)$ and $h_2(n,i)$ that minimize the mean squared value of $e_a(n,i)$. The mean squared error $\mathrm{E}\{e_a^2(n,i)\}$ is written as

$$\begin{aligned}\mathrm{E}\{e_a^2(n,i)\} &= \mathrm{E}\{\boldsymbol{h}^{\mathrm{T}}(i)\, \boldsymbol{m}(i)\, \boldsymbol{m}^{\mathrm{T}}(i)\, \boldsymbol{h}(i)\} \\ &= \boldsymbol{h}^{\mathrm{T}}(i)\, \boldsymbol{R}(i)\, \boldsymbol{h}(i),\end{aligned} \tag{3}$$

where $\boldsymbol{h}(i)$ is the filter coefficient vector:

$$\boldsymbol{h}(i) = \begin{pmatrix} h_2(0,i) \\ h_2(1,i) \\ \vdots \\ h_2(i,i) \\ -h_1(0,i) \\ -h_1(1,i) \\ \vdots \\ -h_1(i,i) \end{pmatrix}, \tag{4}$$

$\boldsymbol{m}(i)$ is the input signal vector:

$$\boldsymbol{m}(i) = \begin{pmatrix} m_1(n) \\ m_1(n-1) \\ \vdots \\ m_1(n-i) \\ m_2(n) \\ m_2(n-1) \\ \vdots \\ m_2(n-i) \end{pmatrix}, \tag{5}$$

$\boldsymbol{R}(i) = \mathrm{E}\{\boldsymbol{m}(i)\, \boldsymbol{m}^{\mathrm{T}}(i)\}$ is the input correlation matrix, and E{ } represents expectation. The vector $\boldsymbol{h}(i)$ minimizing $\mathrm{E}\{e_a^2(n,i)\}$ keeping the norm $\|\boldsymbol{h}(i)\|$ constant can be derived as the eigenvector corresponding to the smallest eigenvalue of $\boldsymbol{R}(i)$.

2.2 Determination of Modeling Order and Source Recovery

The order of $c_1(n)$ and $c_2(n)$ is unknown. Thus, if order i is not an appropriate value to model the impulse responses $c_1(n)$ and $c_2(n)$, the $h_1(n,i)$ and $h_2(n,i)$ calculated by (3) do not satisfy (2).

First, a cost function is introduced to determine the optimum value of order i using Fig. 1(b) . This cost function is based on multichannel inverse filtering theory [3]. First, the multichannel inverse filters $g_1(n,i)$ and $g_2(n,i)$ for $h_1(n,i)$ and $h_2(n,i)$ are derived by solving the following diophantine equation:

$$G_1(z,i)H_1(z,i) + G_2(z,i)H_2(z,i) = 1, \tag{6}$$

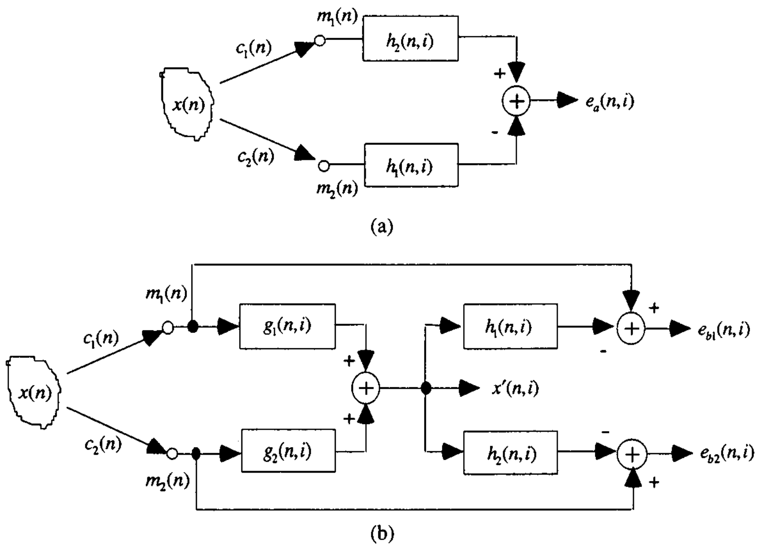


Fig. 1 Two-channel blind deconvolution framework for non-minimum phase impulse responses:
(a) estimating impulse responses for a given modeling order i;
(b) searching for optimum order and recovering source signal.

where $G_1(z,i)$, $G_2(z,i)$, $H_1(z,i)$, and $H_2(z,i)$ are the z-transforms of $g_1(n,i)$, $g_2(n,i)$, $h_1(n,i)$ and $h_2(n,i)$. Then the recovered source signal for order i is calculated using

$$x'(n,i) = m_1(n) * g_1(n,i) + m_2(n) * g_2(n,i) \qquad (7)$$

Now, the cost function $PE(i)$ is defined as

$$PE(i) = \frac{E\{e_{b1}^2(n,i)\}}{E\{m_1^2(n)\}} + \frac{E\{e_{b2}^2(n,i)\}}{E\{m_2^2(n)\}}, \qquad (8)$$

where

$$\begin{aligned} e_{b1}(n,i) &= m_1(n) - x'(n,i) * h_1(n,i) \\ e_{b2}(n,i) &= m_2(n) - x'(n,i) * h_2(n,i). \end{aligned} \qquad (9)$$

This cost function evaluates how well the estimated impulse responses $h_1(n,i)$ and $h_2(n,i)$ and the recovered source signal $x'(n,i)$ approximate the actual reverberant signals $m_1(n) = x(n) * c_1(n)$ and $m_2(n) = x(n) * c_2(n)$.

If and only if $PE(i) = 0$; that is, $e_{b1}(n,i) = 0$ and $e_{b2}(n,i) = 0$,

$$\begin{aligned} h_1(n,i) &= \alpha\, c_1(n) \\ h_2(n,i) &= \alpha\, c_2(n) \\ x'(n,i) &= \frac{1}{\alpha}\, x(n), \end{aligned} \qquad (10)$$

where α is an arbitrary constant. The proof of (10) is given in the Appendix.

Since we assume that there are no common zeros between $c_1(n)$ and $c_2(n)$, the common zeros of the observed signals $m_1(n)$ and $m_2(n)$ represent the source signal $x(n)$ and the noncommon zeros represent the characteristics of the channels $c_1(n)$ and $c_2(n)$. Thus the cost function $PE(i)$ may be considered to be minimized by extracting the common zeros from the observed signals. If $c_1(n)$ and $c_2(n)$ have common zeros, the zeros are extracted and $c_1(n)$ and $c_2(n)$ are estimated as the remainders.

The optimum order I of the estimated impulse responses is determined by the following procedure:

(i) The filter coefficient vector $\boldsymbol{h}(i)$ is computed as the eigenvector corresponding to the smallest eigenvalue of $\boldsymbol{R}(i)$, then we have the estimates $h_1(n,i)$ and $h_2(n,i)$ using (4).

(ii) The inverse filters $g_1(n,i)$ and $g_2(n,i)$ are computed using (6).

(iii) The recovered signal $x'(n,i)$ is computed using (7).

(iv) The cost function $PE(i)$ is computed using (8).

(v) The above computations (i) - (iv) are done for various values of order i. Then, the order that minimizes the cost function $PE(i)$ is selected as the optimum order I.

The estimated impulse responses $h_1(n,I)$ and $h_2(n,I)$ and the recovered source signal $x'(n,I)$ with the optimum order I are used as the final estimates.

3. Computer Simulation

To confirm the validity of the proposed method, we simulated two-channel blind deconvolution for nonminimum phase impulse responses. The reverberant signals were obtained by convolving the source signal with the two nonminimum phase impulse responses $c_1(n)$ and $c_2(n)$ shown in Fig. 3. The order of the impulse responses was 30.

The optimum order was searched for from 3 to 50. Figure 2 shows that the optimum order I, the one that minimizes the cost function $PE(i)$, is 28. Although this optimum order differs from the original order 30, it is a valid value because there was a two-tap delay at the head of the impulse responses, as shown in Fig. 3, and the delay was extracted as common zeros through the computation.

Figure 3 compares the original impulse responses with the estimates containing the two-tap delay and shows that the estimated impulse responses $h_1(n,i)$ and $h_2(n,i)$ are good approximations of the original responses $c_1(n)$ and $c_2(n)$.

The result of deconvolution for an impulse source signal is shown in Fig. 4, the reverberant signal $m_1(n)$ (dashed line) is overlaid on the recovered source signal. The reverberant part of the recovered signal is suppressed well, which demonstrates that the proposed method can help overcome reverberation problems.

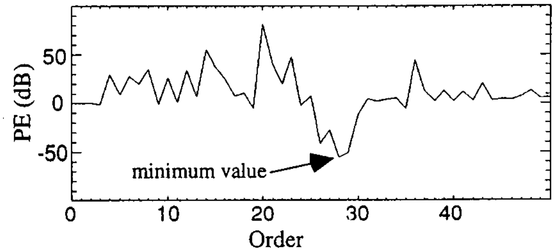


Fig. 2 Optimum order of the estimated impulse responses. The arrow points to the minimum value at the optimum order.

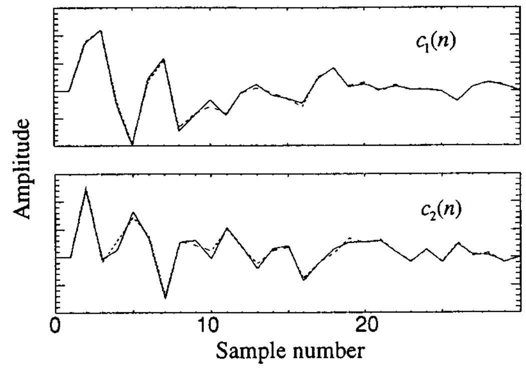


Fig. 3 Estimates (solid lines) and originals (dashed lines) of two unknown impulse responses that are nonminimum phase.

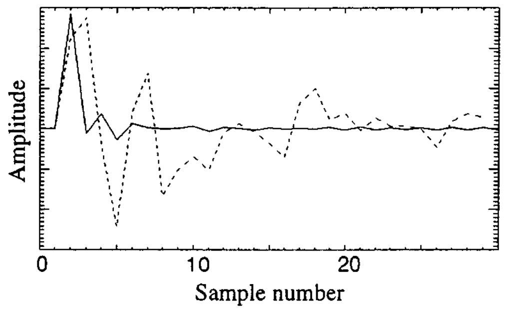


Fig. 4 Result of deconvolution using the proposed method. The solid line is the recovered signal of an impulse source and the dashed line is the original impulse signal distorted by reverberation.

4. Conclusions

The method for blind deconvolution proposed here can determine the order of the impulse response filter model and estimate impulse responses which may be nonminimum phase. The proposed method is based on a cost function that is minimized when the common zeros of the z-transforms of the two observed signals are extracted. Multichannel inverse filtering with the estimated impulse responses of the optimum order is used to recover the unknown source signal. Computer simulation showed that this method can estimate nonminimum phase impulse responses from two reverberant signals and recover the source signal.

Acknowledgments

We thank Dr. Nobuhiko Kitawaki, Director of the Speech and Acoustics Laboratory, and Mr. Junji Kojima, Leader of the Audio Information Processing Group, for their helpful guidance in this research.

References

[1] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," IEEE Trans. Acoust., Speech & Signal Process., vol.36, no. 2, pp. 145-152, 1988.

[2] M. I. Gurelli and C. L. Nikias, " EVAM: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals," IEEE Trans. Signal Process., vol.43, no.1, pp. 134-149, 1995.

[3] H. Wang, "Multi-channel deconvolution using Pade approximation," Proc. ICASSP 95, pp. 3007-3010, 1995.

Appendix

The z-transforms of $e_{b1}(n,i)$ and $e_{b2}(n,i)$ are written as

$$\begin{aligned}E_{b1}(z,i) &= M_1(z) - X'(z,i)H_1(z,i)\\ &= C_1(z)X(z) - H_1(z,i)\{G_1(z)C_1(z) + G_2(z)C_2(z)\}X(z)\\ &= \{C_1(z) - H_1(z,i)\{G_1(z)C_1(z) + G_2(z)C_2(z)\}\}X(z)\\ E_{b2}(z,i) &= M_2(z) - X'(z,i)H_2(z,i)\\ &= C_2(z)X(z) - H_2(z,i)\{G_1(z)C_1(z) + G_2(z)C_2(z)\}X(z)\\ &= \{C_2(z) - H_2(z,i)\{G_1(z)C_1(z) + G_2(z)C_2(z)\}\}X(z),\end{aligned} \tag{A$\cdot$1}$$

where $M_1(z)$, $M_2(z)$, $C_1(z)$, $C_2(z)$, $X(z)$, and $X'(z,i)$ are the z-transforms of $m_1(n)$, $m_2(n)$, $c_1(n)$, $c_2(n)$, $x(n)$, and $x'(n,i)$. If $e_{b1}(n,i) = 0$ and $e_{b2}(n,i) = 0$, then

$$\begin{aligned}0 &= \{C_1(z) - H_1(z,i)\{G_1(z)C_1(z) + G_2(z)C_2(z)\}\}X(z)\\ 0 &= \{C_2(z) - H_2(z,i)\{G_1(z)C_1(z) + G_2(z)C_2(z)\}\}X(z).\end{aligned} \tag{A$\cdot$2}$$

When (A·2) is satisfied for any source signal $X(z)$, we have

$$\begin{aligned}C_1(z) &= H_1(z,i)\{G_1(z)C_1(z) + G_2(z)C_2(z)\}\\ C_2(z) &= H_2(z,i)\{G_1(z)C_1(z) + G_2(z)C_2(z)\}.\end{aligned} \tag{A$\cdot$3}$$

Equations (A·3) indicate that $G_1(z)C_1(z) + G_2(z)C_2(z)$ is a common term of$C_1(z)$ and $C_2(z)$. Since $G_1(z,i)$, $G_2(z,i)$, $C_1(z)$, and $C_2(z)$ are the z-transforms of the FIR filters $g_1(n,i)$, $g_2(n,i)$, $c_1(n)$, and $c_2(n)$, $G_1(z)C_1(z) + G_2(z)C_2(z)$ is a finite polynomial:

$$G_1(z)C_1(z) + G_2(z)C_2(z) = \beta_0+\beta_1 z+ \cdots +\beta_k z^k+ \cdots +\beta_M z^M \tag{A$\cdot$4}$$

where β_k is the coefficient of the kth order term of z. However, since we assume that there is no common zero in $C_1(z)$ and $C_2(z)$, $G_1(z)C_1(z) + G_2(z)C_2(z)$ does not have zeros. Thus β_k is zero for all k except β_0.

$$G_1(z)C_1(z) + G_2(z)C_2(z) = \beta_0 \tag{A$\cdot$5}$$

Substituting (A·5) into (A·3), we have

$$\begin{aligned}H_1(z,i) &= \alpha\, C_1(z)\\ H_2(z,i) &= \alpha\, C_2(z),\end{aligned} \tag{A$\cdot$6}$$

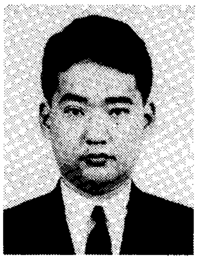
where α is $1/\beta_0$. The z-transform of the recovered source signal $x'(n,i)$ in (7) is written as

$$\begin{aligned} X'(z) &= M_1(z)G_1(z) + M_2(z)G_2(z) \\ &= G_1(z)C_1(z)X(z) + G_2(z)C_2(z)X(z) \\ &= \{G_1(z)C_1(z) + G_2(z)C_2(z)\}X(z). \end{aligned} \tag{A·7}$$

Substituting (A·5) into (A·7), we have

$$\begin{aligned} X'(z) &= \beta_0 X(z) \\ &= \frac{1}{\alpha} X(z). \end{aligned} \tag{A·8}$$

The inverse z-transforms of (A·6) and (A·8) are (10).



Ken'ichi Furuya was born in Fukuoka, Japan in 1962. He graduated in 1985 from the Department of Acoustic Design, Kyushu Institute of Design, where he obtained a Master's degree in 1987 and affiliated with Nippon Telegraph and Telephone Corporation (NTT). He is now a Senior Research Engineer at the Speech and Acoustics Laboratory of the NTT Human Interface Laboratories. His current resarch interests include signal processing in acoustic engineering. He was awarded the Sato Prize of the Acoustical Society of Japan in 1991. He is a member of the Acoustical Society of Japan, the Acoustical Society of America, and IEEE.

Yutaka Kaneda was born in Osaka, Japan, on February 20, 1951. He received the B.E., M.E. and Doctor of Engineering degrees from Nagoya University, Nagoya, Japan, in 1975, 1977 and 1990. In 1977, he joined 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 Speech and Acoustics Laboratory of the NTT Human Interface Laboratories. He 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. Dr. Kaneda is a member of the Acoustical Society of Japan, and the Acoustical Society of America.