

# TWO-CHANNEL BLIND DECONVOLUTION FOR NON-MINIMUM PHASE IMPULSE RESPONSES

*Ken'ichi Furuya and Yutaka Kaneda*

NTT Human Interface Laboratories  
3-9-11, Midori-cho, Musashino-shi, Tokyo 180, Japan  
furuya@splab.hil.ntt.co.jp

## ABSTRACT

A new blind deconvolution method is proposed for recovering an unknown source signal, which is observed through two unknown channels characterized by non-minimum phase impulse response filters. Conventional methods cannot estimate the non-minimum phase parts. Our method is based on computing the eigenvector corresponding to the smallest eigenvalue of the input correlation matrix and using a cost function to determine the order of the impulse response filter model. Multi-channel inverse filtering with the estimated impulse responses is used to recover the unknown source signal. Sub-band processing is also used to reduce the complexity of dealing with long impulse responses such as room impulse responses. Computer simulation shows that the effectiveness of our method.

## 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] proposed a criterion for determining the modeling order for minimum phase impulse responses, but estimating the impulse responses accurately is still difficult because room impulse responses are usually non-minimum phase.

This paper proposes a method for determining the order of the impulse response filter model and estimating impulse responses that may be non-minimum phase. This method is based on a cost function that is minimized when there are no non-common zeros between the two observed signals. If there are no common zeros between the system transfer functions of the two unknown channels, the source signal can be recovered by minimizing the cost function. Since the impulse response of a room is usually

too long to deal with, sub-band processing [4] is also introduced to reduce the computational complexity.

## 2. PRINCIPLE

The proposed method consists of two stages, as shown in Figs. 1(a) and (b). First, the impulse responses are estimated for various modeling orders. Next, the optimum order is determined, and then the source signal is recovered by 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

$$e_a(n,i) = x(n) * \{c_1(n) * h_2(n,i) - c_2(n) * h_1(n,i)\}, \quad (1)$$

where the symbol  $*$  represents convolution.

Here, if  $i = j$  and  $e_a(n,i) = 0$  for all  $n$ , then  $h_1(n,i)$  and  $h_2(n,i)$  satisfy

$$\begin{aligned} h_1(n,i) &= \alpha c_1(n) \\ h_2(n,i) &= \alpha c_2(n), \end{aligned} \quad (2)$$

where  $\alpha$  is an arbitrary constant. Thus,  $h_1(n,i)$  and  $h_2(n,i)$  can be considered estimates of  $c_1(n)$  and  $c_2(n)$ .

### 2.2 Computation based on eigenvector

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

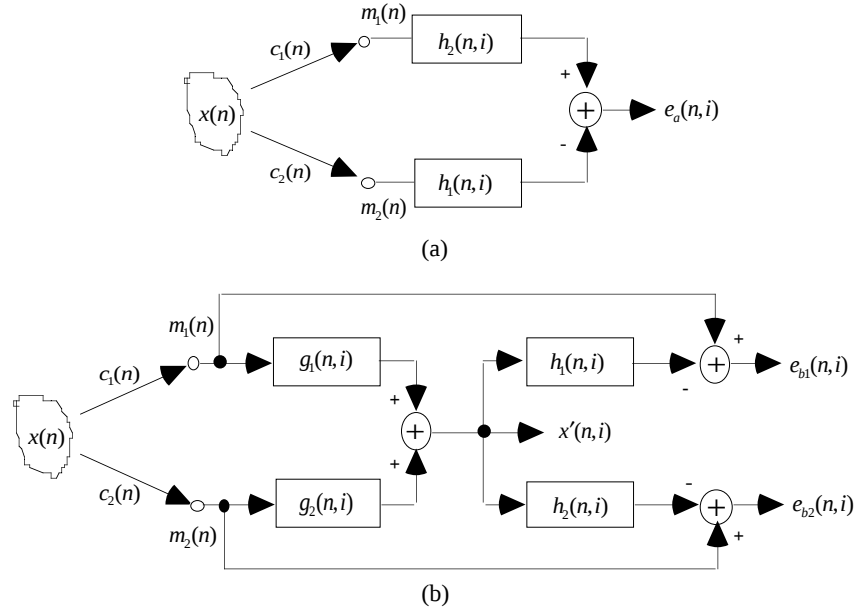


Figure 1. Two-channel blind deconvolution framework for non-minimum phase impulse responses:  
(a) estimating impulse responses for a given modeling order  $i$  ;  
(b) searching for optimum order and recovering the source signal.

mean squared error  $E\{e_a^2(n,i)\}$  is written as

$$E\{e_a^2(n,i)\} = E\{\mathbf{h}^T(i)\mathbf{m}(i)\mathbf{m}^T(i)\mathbf{h}(i)\} = \mathbf{h}^T(i)\mathbf{R}(i)\mathbf{h}(i), \quad (3)$$

where  $\mathbf{h}(i)$  is the filter coefficient vector  $[h_2(0,i) \ h_2(1,i) \ \dots \ h_2(i,i) \ -h_1(0,i) \ -h_1(1,i) \ \dots \ -h_1(i,i)]^T$ ,  $\mathbf{m}(i)$  is the input signal vector  $[m_1(n) \ m_1(n-1) \ \dots \ m_1(n-i) \ m_2(n) \ m_2(n-1) \ \dots \ m_2(n-i)]^T$ ,  $\mathbf{R}(i) = E\{\mathbf{m}(i)\mathbf{m}^T(i)\}$  is the input correlation matrix, and  $E\{\}$  represents expectation. The vector  $\mathbf{h}(i)$  that minimizes  $E\{e_a^2(n,i)\}$  keeping the norm  $\|\mathbf{h}(i)\|$  constant can be derived as the eigenvector corresponding to the smallest eigenvalue of  $\mathbf{R}(i)$ .

### 2.3 Determination of optimum order and source recovery

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

A cost function is introduced to determine the optimum value of filter order  $i$ . It is based on multichannel inverse filtering theory [1]. Figure 1(b) shows the scheme for deriving the cost function. 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:

$$G_1(z,i)H_1(z,i) + G_2(z,i)H_2(z,i) = 1, \quad (4)$$

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 inverse filters  $g_1(n,i)$ ,  $g_2(n,i)$  are applied to the received signals  $m_1(n)$  and  $m_2(n)$  in order to generate the recovered signal  $x'(n,i)$ . The recovered source signal  $x'(n,i)$  for order  $i$  is written as

$$x'(n,i) = m_1(n)*g_1(n,i) + m_2(n)*g_2(n,i). \quad (5)$$

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

$$PE(i) = E\{e_{b1}^2(n,i)\}/E\{m_1^2(n)\} + E\{e_{b2}^2(n,i)\}/E\{m_2^2(n)\}, \quad (6)$$

where

$$e_{b1}(n,i) = m_1(n) - x'(n,i)*h_1(n,i) \\ e_{b2}(n,i) = m_2(n) - x'(n,i)*h_2(n,i).$$

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  $e_{b2}(n,i)=0$ , then

$$h_1(n,i) = \alpha c_1(n) \\ h_2(n,i) = \alpha c_2(n) \\ x'(n,i) = 1/\alpha x(n), \quad (7)$$

where  $\alpha$  is an arbitrary constant. The optimum order of the estimated impulse responses is determined such that it minimizes the cost function  $PE(i)$ . 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 are used as the final estimates.

## 2.4 Simulation for short impulse responses

To confirm the validity of the proposed method, we simulated two-channel blind deconvolution for non-minimum phase impulse responses. The reverberant signals were obtained by convolving the source signal with the two non-minimum 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 between 3 and 50. Figure 2 shows that the optimum order, i.e., the one that minimizes the cost function  $PE(i)$ , is 28. Although this optimum order differs from the original order 30, it is reasonable 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 is overlaid on the recovered source signal. The reverberant part of the recovered signal is suppressed well.

## 3. COMBINATION WITH SUB-BAND PROCESSING

### 3.1 Procedure of combination

Room impulse responses are often too long (more than 1000 samples) to deal with. It is difficult to compute the eigenvector and inverse filters of this order with sufficient precision and reasonable speed.

Sub-band processing [4] is a way of dealing with the problems mentioned above. The down-sampling of sub-band signals by the sub-band processing reduces the effective order of impulse responses. In the combination of the two-channel blind deconvolution and sub-band processing shown in Fig. 5, the process is: (i) The signals are divided into hundreds of sub-bands. (ii) The sub-band signals are down-sampled. (iii) The two-channel blind deconvolution is applied to each sub-band signal. (iv) The full-band signal is reconstructed by up-sampling and summing up the recovered sub-band signals. An arbitrary constant of each recovered signal is determined by equalizing the power of the recovered signal to the power of the source signal in each sub-band. Note that this step is not blind, i.e., it uses the information about the source signal power. However, if we can approximate the power of the source signal by the power of the received signals in each sub-band, this step becomes completely blind.

### 3.2 Simulation for room impulse responses

We performed computer simulation using room impulse responses was conducted to investigate the effectiveness of our method

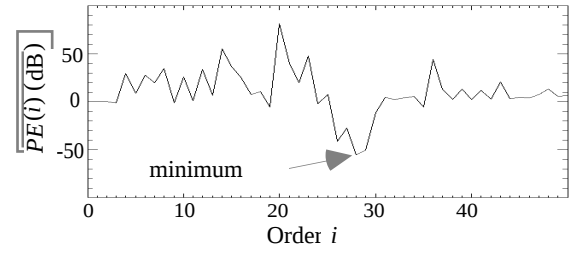


Figure 2. Optimum order of the estimated impulse responses. The arrow points to the minimum value at the optimum order.

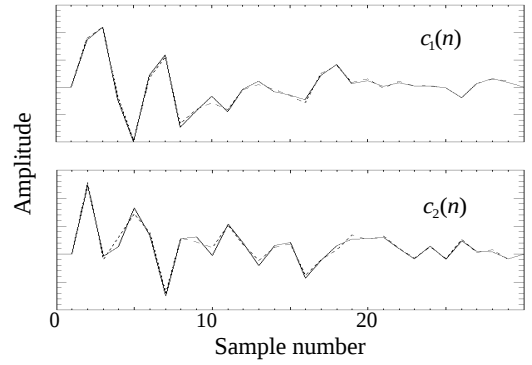


Figure 3. Estimates (solid lines) and originals (broken lines) of two unknown impulse responses that are non-minimum phase.

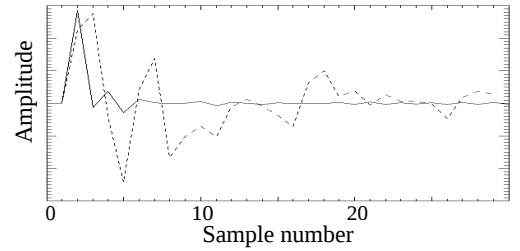


Figure 4. Result of deconvolution using the proposed method. The solid line is the recovered signal of an impulse source and the broken line is the original impulse signal distorted by reverberation.

with sub-band processing. The impulse responses were measured in a room whose reverberation time was 0.5 s. The distances between the sound source and microphones were 3 m and 4 m, and the sampling frequency was 12 kHz. The impulse responses were truncated to be 3000 samples, i.e. 250 ms as shown in Fig. 6. The number of sub-bands was 200, and the down-sampling and up-sampling rates were both 100.

Figure 7 shows the recovered impulse signal. The power of the non-pulsive part or reverberant part was decreased about 5 dB by our method. However, the recovered impulse signal was delayed by more than 500 ms and included a non-causal reverberant part which caused unnatural sound. These phenomena are due to the sub-band processing.

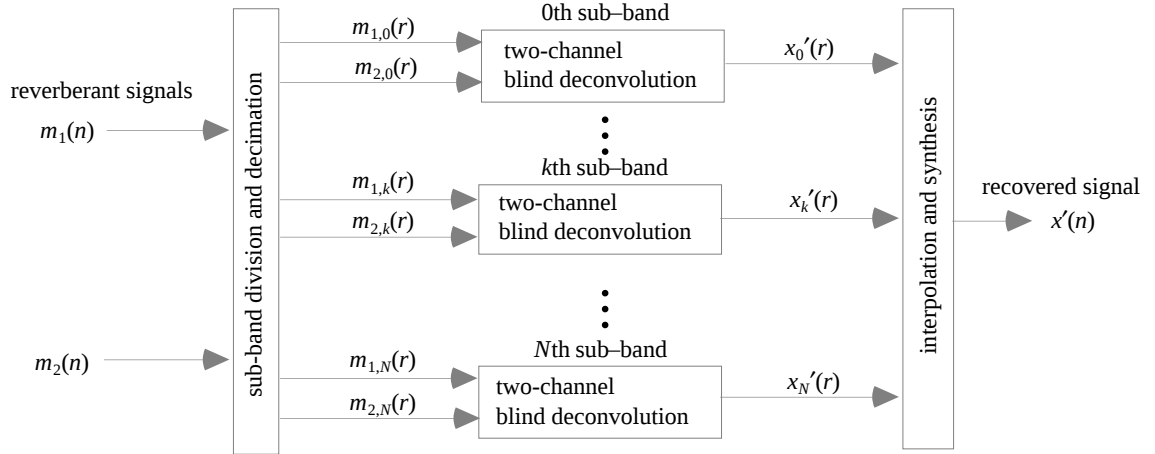


Figure 5. Block diagram of combination of two-channel blind deconvolution and sub-band processing.

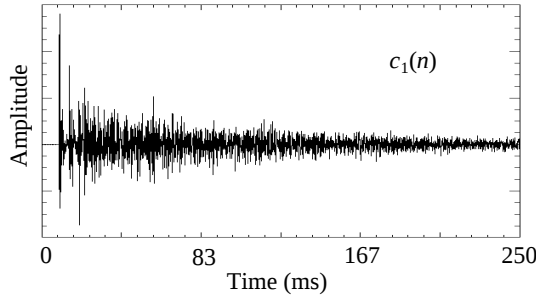


Figure 6. Room impulse response.

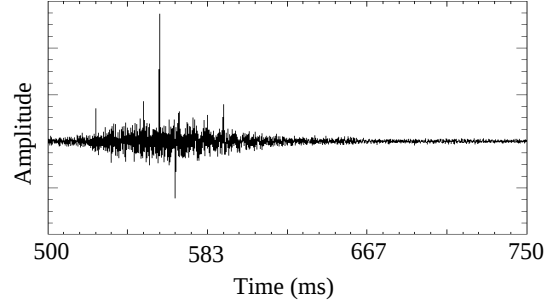


Figure 7. Recovered impulse signal.

#### 4. CONCLUSIONS

Our new method for achieving blind deconvolution can estimate non-minimum phase impulse responses from two-channel reverberant signals and recover the source signal. Computer simulation showed that it can achieve nearly perfect blind deconvolution for short impulse responses. For long (several thousand) impulse responses, such as room impulse responses, we introduced sub-band processing to reduce the computational complexity. The power of the reverberant part of a room impulse response was decreased about 5 dB by this method. For further improvement in performance, we plan to study optimizing the sub-band processing.

#### 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. ASSP*, 36 (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. SP*, 43 (1), pp. 134-149, 1995.
- [3] H. Wang, "Multi-channel deconvolution using Pade approximation", *Proceedings of the ICASSP 95*, pp. 3007-3010, 1995.
- [4] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Printice-Hall, 1983.