TM04-5

SICE02-0389

Minimal Distortion Principle for Blind Source Separation

K. Matsuoka

Kyushu Institute of Technology, Hibikino, Wakamatsu, Kitakyushu, Japan
matsuoka@brain.kyutech.ac.jp

Abstract: In blind source separation the number of the sensors is usually assumed to be equal to that of the sources. In this case an indeterminacy appears with which any linear transform of an estimated source signal can also be considered another estimation of the source signal. Moreover in the case that the number of the sensors is greater than that of the sources, another indeterminacy arises due to the redundancy of the sensors. Although these indeterminacies are often considered unsubstantial and have been eliminated without definite bases, an appropriate normalization of the separator is important to enhance the accuracy of the separation result, particularly in the case of convolutive mixture. This paper shows two principles for eliminating these indeterminacies: (i) minimal distortion principle and (ii) inverse minimal distortion principle.

Keywords: blind source separation, convolutive mixture, minimal distortion principle

1. Introduction

Blind source separation (BSS), or independent component analysis (ICA), is a method for recovering a set of statistically independent signals from the observation of their mixtures without any prior knowledge about the mixing process. It has been receiving a great deal of attention from various fields as a new signal processing technique. In view of the level of complexity, the mixing process can be classified into two types: instantaneous mixture and convolutive mixture. While early works for BSS considered the former type, recent works are mainly concerned with the latter type, which is much more difficult from theoretical and computational points of view. In this paper we deal with BSS of convolutive mixture in general, which includes instantaneous mixture as a special case.

Let $\mathbf{x}(t)=\mathbf{A}(z)\mathbf{s}(t)$ be a convolutive mixing process, where $\mathbf{s}(t)$ and $\mathbf{x}(t)$ are a source signal (vector) and a sensor signal (vector), respectively, and $\mathbf{A}(z)$ is an $M\times N$ transfer function matrix. Although typically the number M of the sensors is assumed to be equal to the number N of the sources, this paper deals with a more general case that M is equal to or greater than N. The demixing process (the separator) then takes the form $\mathbf{y}(t)=\mathbf{W}(z)\mathbf{x}(t)$, where $\mathbf{W}(z)$ is an $N\times M$ matrix.

As known well, two kinds of unavoidable indeterminacies exist in this case. One is the indeterminacy in scaling or normalization of the separator's output $\mathbf{y}(t)$; any linear transform of a source signal can also be considered a source signal in the 'blind' situation. So, if $\mathbf{W}(z)$ is a valid separator that recovers the source signal, then matrix $\mathbf{D}(z)\mathbf{W}(z)$ may well be valid, where $\mathbf{D}(z)$ is an arbitrary diagonal matrix. This type of indeterminacy always appears irrespectively of whether $M=N$ or $M>N$. In the case of instantaneous mixture, the indeterminacy is usually considered unsubstantial because it is only related to the scaling of the source signals. In the case of convolutive mixture, however, an appropriate choice of $\mathbf{D}(z)$ is very important for some reasons as shown later. [BSS also has indeterminacy in the numbering of the sources, but it will be discarded throughout the paper because it is not so essential.]

In the case of $M>N$, another indeterminacy arises due to the redundancy of the sensors. Let $\mathbf{Z}(z)$ be an arbitrary $N\times M$ matrix that satisfies $\mathbf{Z}(z)\mathbf{A}(z)=\mathbf{O}$. Then, if $\mathbf{W}(z)$ is a valid separator, $\mathbf{W}(z)+\mathbf{Z}(z)$ is also valid because $\left(\mathbf{W}(z)+\mathbf{Z}(z)\right)\mathbf{A}(z)\mathbf{s}(t)=\mathbf{W}(z)\mathbf{A}(z)\mathbf{s}(t)$. However, this argument holds only in the case that the mixing process is noiseless. In noisy situations, the separator $\mathbf{W}(z)$ must be chosen so as to reduce the noise to the utmost. In the case of instantaneous mixture, this issue is usually solved by mapping the M-dimensional vector $\mathbf{x}(t)$ to an N-dimensional vector space such that N dominant components in the data may be preserved, using principal component analysis. In the case of convolutive mixture, however, this mapping needs to be done in an infinite-dimensional space of time series, and hence the treatment becomes more complicated.

The purpose of this paper is to propose two natural and useful principles to eliminate the above-mentioned indeterminacies inherent in BSS. They are (i) the minimal distortion principle and (ii) the inverse minimal distortion principle.

The first principle claims that among a set of valid separators the optimal separator should have the property that the observed signals are the least subjected to distortion by the separator. Separators obtained based on this principle have some favorable features, particularly for convolutive mixture of sources. The basic idea of this principle was first introduced in the author's recent paper [5], but it dealt with only the case of $M = N$. In the case of $M > N$, a generalization is required.

The second principle demands that applying the pseudo-inverse of $\mathbf{W}(z)$ to the separator's output must yield the least deformation relative to the observed signal. This principle has been utilized implicitly or explicitly in the case of instantaneous mixuture, but in the present paper we shall apply this principle to the case of convolutive mixture.

Mathematical Nomenclature

- $\|\mathbf{x}\|$ represents the Euclidean norm of vector $\mathbf{x}$.
- The conjugate of the transpose of transfer function matrix $\mathbf{X}(z)$ (with real matrix coefficients) is defined as $\mathbf{X}^H(z) \triangleq \mathbf{X}^T(z^{-1})$.
- $\mathrm{diag}\{d_i\}$ represents the diagonal matrix that has diagonal entries $d_1,\ldots,d_N$. $\mathrm{diag}\,\mathbf{X}$ (off-diag $\mathbf{X}$) sets every off-diagonal (diagonal) entry of matrix $\mathbf{X}$ to be zero.
- $\delta(\tau) = 1\ (\tau = 0), = 0\ (\tau \neq 0)$.

2. Mixing and Demixing Processes

2.1 Mixing process

Let us consider a situation where statistically independent random signals $s_i(t)$ $(i = 1,\cdots, N)$ are generated by N sources and their mixtures $x_i(t)$ $(i = 1,\cdots, M)$ are observed by M sensors. Throughout the paper we deal with the case that the number M of the sensors is equal to or larger than the number N of the sources ($M \geq N$). It is assumed that every source signal $s_i(t)$ is a (possibly nonstationary) random process with mean zero and the sensors' outputs $x_i(t)$ $(i = 1, \cdots, M)$ are given by a linear mixing process

$$\mathbf{x}(t) = \sum_{\tau=0}^{\infty} \mathbf{A}_\tau \mathbf{s}(t-\tau) = \mathbf{A}(z)\mathbf{s}(t), \qquad (1)$$

where $\mathbf{s}(t) \triangleq [s_1(t), \cdots, s_N(t)]^T$, $\mathbf{x}(t) \triangleq [x_1(t), \cdots, x_M(t)]^T$, and $\mathbf{A}(z) \triangleq \sum_{\tau=0}^{\infty} \mathbf{A}_\tau z^{-\tau}$. It is well known that, for BSS, at most one source signal is allowed to be Gaussian; we assume that this kind of conditions necessary for BSS are satisfied, of course. For the mixing process we assume that $\mathbf{A}(z)$ is column full rank, that is, $\mathbf{A}^H(z)\mathbf{A}(z)$ is nonsingular for every z on $|z| = 1$.

In BSS the definition of the source signals and correspondingly that of the mixing process have an indeterminacy. Namely, if $\{s_i(t)\}$ are source signals, their arbitrarily linear-filtered signals $\{e_i(z)s_i(t)\}$ can also be considered source signals. The mixing process is then $\mathbf{A}(z)\mathrm{diag}\{e_i^{-1}(z)\}$. There is no way to distinguish between $\{s_i(t)\}$ and $\{e_i(z)s_i(t)\}$ (or equivalently between $\mathbf{A}(z)$ and $\mathbf{A}(z)\mathrm{diag}\{e_i^{-1}(z)\}$).

2.2 Demixing process

To recover the source signals from the sensor signals, we consider a demixing process of the following form:

$$\mathbf{y}(t) = \sum_{\tau=-\infty}^{\infty} \mathbf{W}_\tau \mathbf{x}(t-\tau) = \mathbf{W}(z)\mathbf{x}(t), \qquad (2)$$

where $\mathbf{y}(t) \triangleq [y_1(t), \cdots, y_N(t)]^T$ and $\mathbf{W}(z) \triangleq \sum_{\tau=-\infty}^{\infty} \mathbf{W}_\tau z^{-\tau}$. Note that the impulse response $\{\mathbf{W}_\tau\}$ might in general need to take a noncausal form, i.e., $\mathbf{W}_\tau \neq \mathbf{O}$ $(\tau < 0)$. The problem of noncausality can approximately be solved by designing $\mathbf{W}(z)$ such that the source signals are reproduced with a time lag.

For a fixed mixing matrix $\mathbf{A}(z)$, any $N\times M$ matrix $\mathbf{W}(z)$ can uniquely be represented with respect to its pseudo-inverse $\mathbf{A}^{\dagger}(z) \triangleq (\mathbf{A}^{H}(z)\mathbf{A}(z))^{-1}\mathbf{A}^{H}(z)$ as

$$\mathbf{W}(z) = \mathbf{F}(z)\mathbf{A}^{\dagger}(z) + \mathbf{Z}(z), \tag{3}$$

where $\mathbf{Z}(z)$ is an $N\times M$ matrix satisfying $\mathbf{Z}(z)\,\mathbf{A}(z) = \mathbf{O}$.

In the particular case of $M = N$, the second term $\mathbf{Z}(z)$ in eqn (3) vanishes.

To achieve separation, the overall transfer function from the sources to the separator's output, $\mathbf{W}(z)\mathbf{A}(z) = \mathbf{F}(z)$, must be a nonsingular diagonal matrix $\mathbf{D}(z) = \mathrm{diag}\{d_i(z)\} = \sum_{\tau=-\infty}^{\infty} \mathbf{D}_\tau z^{-\tau}$. So, any desired separator needs to take the following form:

$$\mathbf{W}(z) = \mathbf{D}(z)\mathbf{A}^{\dagger}(z) + \mathbf{Z}(z) \quad (\mathbf{Z}(z)\,\mathbf{A}(z) = \mathbf{O}). \tag{4}$$

We call any separator of this form a valid separator. If the separator is valid, each source signal appears at an output terminal of the separator though it is subjected to a linear filtering $d_i(z)$. The numbers of degrees of freedom in matrices $\mathbf{D}(z)$ and $\mathbf{Z}(z)$ are N and $MN - N^2$, respectively, and hence a valid separator $\mathbf{W}(z)$ has $MN - N^2 + N$ degrees of freedom in total. Here, 'one' degree of freedom means an 'infinite' number of freedoms held by a scalar analytic function of z.

3. Principles for Eliminating Two Kinds of Indeterminacy

In this section we describe the principles for eliminating the indeterminacies in $\mathbf{D}(z)$ and $\mathbf{Z}(z)$. Since the two indeterminacies are independent of each other, they will be dealt with on different principles.

3.1 Minimal distortion principle

In the case of instantaneous mixture, indeterminacy in $\mathbf{D}(z)$ ($= \mathbf{D}$) is usually considered unsubstantial because that is only related to the scaling of the signals. In the case of convolutive mixture, however, an appropriate choice of $\mathbf{D}(z)$ is very important for some reasons described later.

Now we show a principle to eliminate this indeterminacy.

(P1) Among the set of valid separators, choose $\mathbf{W}(z)$ so that $E\left[\|\mathbf{y}(t) - \mathbf{Q}\mathbf{x}(t)\|^2\right]$ be minimized.

Matrix $\mathbf{Q}$ is an $N\times M$ matrix, being given somewhat arbitrarily by the designer. Note that $\mathbf{Q}$ may well not be full rank. One can show that the above principle leads to

$$\mathbf{D}(z) = \mathrm{diag}\left(\mathbf{Q}\mathbf{A}(z)\right). \tag{5}$$

An alternative representation of the proposed principle is:

(P1') Among the set of valid separators, choose $\mathbf{W}(z)$ so that $\mathrm{diag}\,E\left[\left(\mathbf{y}(t) - \mathbf{Q}\mathbf{x}(t)\right)\mathbf{y}^{H}(t,z)\right] = \mathbf{0}$.

Here, $\mathbf{y}^{H}(t,z)$ is formally defined as $\mathbf{y}(t,z) \triangleq \sum_{\tau=-\infty}^{\infty} \mathbf{y}(t+\tau)z^{-\tau}$ and $\mathbf{y}^{H}(t,z) \triangleq \mathbf{y}^{T}(t,z^{-1})$.

The meaning of the proposed principle can well be understood in the special case of $M = N$ and $\mathbf{Q} = \mathbf{I}$; then $\mathbf{W}(z) = \mathrm{diag}\,\mathbf{A}(z)\cdot\mathbf{A}^{-1}(z)$. In this case, the present principle claims that the sensor signal $\mathbf{x}(t)$ should be the least subjected to distortion by the separator. For this reason the author called this principle the minimal distortion principle in [5]. Normalization of $\mathbf{W}(z)$ based on the principle has some useful properties in actual applications.

First, the separator's output then becomes

$$\mathbf{y}(t) = \mathrm{diag}\mathbf{A}(z)\cdot\mathbf{A}^{-1}(z)\mathbf{A}(z)\mathbf{s}(t) = \mathrm{diag}\mathbf{A}(z)\cdot\mathbf{s}(t). \tag{6}$$

This implies that output $y_i(t)$ of the separator becomes $a_{ii}(z)s_i(t)$, which is the i-th source that would be observed at the i-th sensor in the absence of other source signals than $s_i(t)$. Namely, the i-th source signal observed at the i-th sensor is subjected to no distortion by the separator. This property may be convenient for evaluating the separation result and for later processing.

Second, the obtained separator does not depend on the properties of the sources. Namely, it is independent of the indeterminacy in the definition of the source signals because the following holds for any diagonal matrix $\mathbf{E}(z)$:

$$\operatorname{diag}\mathbf{A}(z)\mathbf{E}(z)\cdot(\mathbf{A}(z)\mathbf{E}(z))^{-1} = \operatorname{diag}\mathbf{A}(z)\cdot\mathbf{A}^{-1}(z).$$

So, even for such nonstationary signals as speech, the separator obtained based on this principle is invariant with time as long as the mixing process is fixed, which is a favorable property in actual implementation. Usual normalizations as $E[y_i(t)y_i(t-\tau)] = \delta(\tau)$ and $E[\varphi_i(y_i(t))y_i(t-\tau)] = \delta(\tau)$ do not have this property.

Third, in actual implementation it is required to construct the separator with a (multi-dimensional) FIR filter. Then, it is desirable that the filter's degree is as low as possible. The present principle determines a valid separator so that its output may become as close to the input as possible. So, it can be expected that the separator will be realized with a relatively short filter length.

Another natural selection for $\mathbf{Q}$, which can also be used in the case of $M > N$, is $\mathbf{Q} = \mathbf{J}/M$, where $\mathbf{J}$ is the matrix whose entries are all unity. This intends that output $y_i(t)$ of the separator be $\frac{1}{M}\sum_{j=1}^{M} a_{ji}(z)s_i(t)$, which is the average of all the observations of the i-th source detected at the sensors.

3.2 Inverse minimal distortion principle

As for the indeterminacy in $\mathbf{Z}(z)$, if there is no noise in the mixing process any selection of $\mathbf{Z}(z)$ gives the same result (i.e., the same overall transfer function matrix). There is no necessity even to deal with all the M sensor signals; we have only to utilize N signals among them. Since however the observed data inevitably contain some noise in actual situations, it is desirable to utilize all the data available.

Consider an additive noise $\mathbf{n}(t)$ to $\mathbf{x}(t)$, then a valid separator's output becomes

$$\begin{aligned}\mathbf{y}(t) &= \mathbf{W}(z)(\mathbf{x}(t)+\mathbf{n}(t)) \\ &= \mathbf{D}(z)\mathbf{s}(t)+\left(\mathbf{D}(z)\mathbf{A}^{\dagger}(z)+\mathbf{Z}(z)\right)\mathbf{n}(t)\end{aligned} \tag{7}$$

One can show that, if $\mathbf{n}(t)$ is spherically symmetric and temporally white, then $\mathbf{Z}(z) = \mathbf{0}$ is the best choice under the condition that $\mathbf{s}(t)$ is recovered at the separator, though with indeterminacy in $\mathbf{D}(z)$ and inevitable noise $\mathbf{D}(z)\mathbf{A}^{\dagger}(z)\mathbf{n}(t)$.

In order to realize $\mathbf{Z}(z) = \mathbf{0}$ without directly manipulating $\mathbf{Z}(z)$, we introduce the following principle:

(P2) Among the set of valid separators, choose $\mathbf{W}(z)$ so that the $E\left[\left\|\mathbf{W}^{\dagger}(z)\mathbf{y}(t)-\mathbf{x}(t)\right\|^2\right]$ be minimized.

Here, $\mathbf{W}^{\dagger}(z)$ is the pseudo-inverse of $\mathbf{W}(z)$; $\mathbf{W}^{\dagger}(z) \triangleq \mathbf{W}^{H}(z)\left(\mathbf{W}(z)\mathbf{W}^{H}(z)\right)^{-1}$. From this we find that $\mathbf{Z}(z) = \mathbf{O}$ gives the minimum of $E\left[\left\|\mathbf{W}^{\dagger}(z)\mathbf{y}(t)-\mathbf{x}(t)\right\|^2\right]$, actually zero, in the noiseless case. The principle gives the same number of constraints as the number of degrees of freedom in $\mathbf{Z}(z)$, being $MN - N^2$.

3.3 The optimal separator

Thus, the separator satisfying both the two principles is

$$\mathbf{W}(z) = \mathbf{W}^{*}(z) \triangleq \operatorname{diag}\mathbf{Q}\mathbf{A}(z)\cdot\mathbf{A}^{\dagger}(z). \tag{8}$$

We call this the optimal separator. The optimal separator is indifferent to the indeterminacy in the definition of the source signals because the following holds for any diagonal matrix $\mathbf{E}(z)$:

$$\operatorname{diag}\mathbf{Q}\mathbf{A}(z)\mathbf{E}(z)\cdot(\mathbf{A}(z)\mathbf{E}(z))^{\dagger} = \operatorname{diag}\mathbf{Q}\mathbf{A}(z)\cdot\mathbf{A}\ (z).$$

4. Implementation of the Principles

4.1 Algorithm

For actual implementation of the separator, we consider a multi-dimensional FIR filter:

$$\mathbf{y}(t-L) = \sum_{\tau=-L}^{L}\mathbf{W}_{\tau}\mathbf{x}(t-L-\tau), \tag{9}$$

that is, $\mathbf{W}(z) = \sum_{\tau=-L}^{L}\mathbf{W}_{\tau}z^{-\tau}$. This aims to recover the source signals with a time lag of around L. The principles described in the last section leads to the following algorithm:

$$
\begin{aligned}
\Delta \mathbf{W}_\tau = & -\alpha \sum_{r=-L}^{L} \text{off-diag}\, \varphi(\mathbf{y}(t-3L))\mathbf{y}^T(t-3L-\tau+r)\cdot \mathbf{W}_r \\
& -\beta \sum_{r=-L}^{L} \text{diag}\left(\mathbf{y}(t-3L)-\sum_{r'=-L}^{L} \mathbf{Q}_{r'}\mathbf{x}(t-3L-r')\right)\mathbf{y}^T(t-3L-\tau+r)\cdot \mathbf{W}_r \\
& +\gamma\left\{\mathbf{y}(t-3L)\mathbf{x}^T(t-3L-\tau)-\sum_{r=-L}^{L}\sum_{r'=-L}^{L} \mathbf{y}(t-3L)\mathbf{y}^T(t-3L-\tau+r)\bar{\mathbf{V}}_{r-r'}\mathbf{W}_{r'}\right\}.
\end{aligned}
\tag{10}
$$

Here, $\varphi_i(y)$ is defined as $\varphi_i(y) \triangleq -d\log q_i(y)/dy$, where $q_i(y)$ is the pdf of the i-th source signal. Matrix $\bar{\mathbf{V}}(z)=\sum_{\tau=-\infty}^{\infty}\bar{\mathbf{V}}_\tau z^{-\tau}$ is defined as $\bar{\mathbf{V}}(z)=\left(\mathbf{W}(z)\mathbf{W}^H(z)\right)^{-1}$; an efficient method for this inverse calculation exists.

The first term in the right hand side of this equation is to achieve separation, and the second and third terms to realize the minimal distortion and the inverse minimal distortion principle. Detailed explanation on the derivation of the algorithm has been omitted for the shortage of space.

4.2 An example

To demonstrate the effectiveness of the proposed method, we here show a result of computer simulation. The mixing process assumed was

$$
\mathbf{A}(z)=\begin{bmatrix} z^{-2} & 0.3z^{-2} \\ 0.7z^{-1} & 0.7z^{-1} \\ 0.3 & 1 \end{bmatrix}.
$$

The sources $s_i(t)$ (i = 1, 2) were both iid signals that obeyed the uniform distribution in [-1, 1]. The degree of the separator was set as L = 5, and matrix $\mathbf{Q}$ and the initial value of $\mathbf{W}(z)$ were chosen respectively as

$$
\mathbf{Q}=\frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{W}(z)=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
$$

In consideration of sub-Gaussianness of the sources, $\varphi(u)$ in eqn (10) was given as $\varphi(u)=u^3$. Learning coefficients α and β were set as 0.0 1 while γ was set as 0.0001; γ must be small enough to avoid the instability possibly induced by the inverse calculation for $\mathbf{V}(z)$. Fig.1 shows the impulse response of the overall transfer function matrix $\mathbf{W}(z)\mathbf{A}(z)$ at the start and after 5000 iterations of time, respectively. Theoretically, $\mathbf{W}(z)$ must converge to

$$
\mathbf{W}(z)\mathbf{A}(z)=\frac{1}{3}\begin{bmatrix} 0.3+0.7z^{-1}+z^{-2} & 0 \\ 0 & 1+0.7z^{-1}+0.3z^{-2} \end{bmatrix}.
$$

Fig.1 (b) shows that the desired separator was obtained successfully.

5. Conclusion

We have shown two principles to eliminate the indeterminacies inherent in BSS with $M \geq N$ and how to implement them. In the history of BSS, the ideas presented in this paper might have been used implicitly or explicitly, but they have usually been limited in the case of instantaneous mixture.

As for the minimal distortion principle, if the feedback-type separator is adopted, the principle can be attained automatically. In the case of convolutive mixture, however, the feedback-type separator is hard to design so as to guarantee its stability. In this paper we have shown a method for realizing the principle in the feedforward-type separator. As for the inverse minimum distortion principle, in the case of instantaneous mixture it can be performed by usual principal component analysis of the observed data. Our method can apply to the case of convolutive mixture. Finally we want to stress that the both principles are combined into an adaptive separation algorithm, which can cope with a (slow) change in the mixing process.

References

[1] S. Amari, S. C. Douglas, A. Cichocki and H. H.Yang, "Multichannel blind deconvolution and equalization using the natural gradient," Proc. IEEE International Workshop on Wireless Communication, pp. 101-104, 1997.

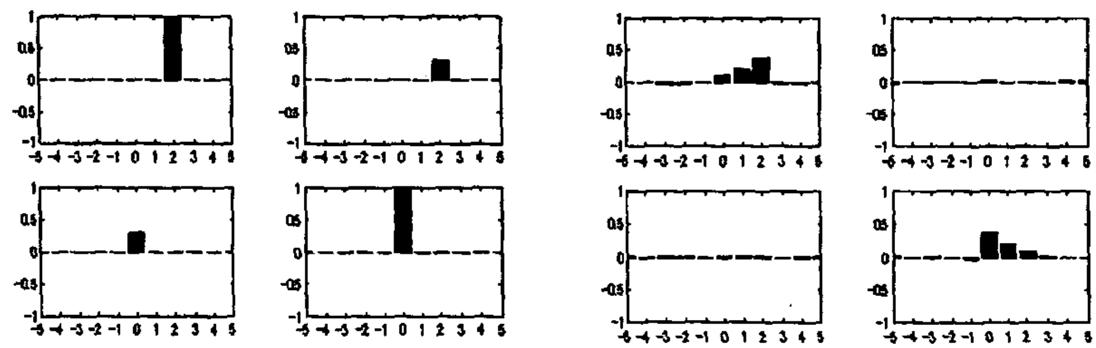
[2] N. Charkani, and Y. Deville, "Self-adaptive separation of convolutively mixed signals with a recursive

structure. Part I: stability analysis and optimization of asymptotic behavior," Signal Processing, vol. 73, pp. 225-254, 1999.

[3] S. Choi, S. Amari, A. Cichocki, and R. Liu, "Natural gradient learning with a nonholonomic constraint for blind deconvolution of multiple channels," Proc. International Workshop on Independent Component Analysis and Blind Signal Separation (ICA'99), pp. 371-376, 1999.

[4] Y. Guo, F. Sattar, and C. Koh, "Blind separation of temporomandibular joint sound signals," Proc. ICASSP, 1999.

[5] K. Matsuoka and S. Nakashima, "Minimal distortion principle for blind source separation," Proceedings of ICA2001, pp.722-727, 2001.



(a) (b)

Fig. 1 The impulse responses of $\mathbf{W}(z)\mathbf{A}(z)$ at the starting (a) and after 5,000 iterations (b) of the computation. The (i, j)-th block represents the impulse response of the (i, j)-th element of $\mathbf{W}(z)\mathbf{A}(z)$ ((i, j)= 1, 2).