Blind Separation for Convolutive Mixtures of Non-stationary Signals

Mitsuru KAWAMOTO[1], Allan Kardec BARROS[1], Ali MANSOUR[1],
Kiyotoshi MATSUOKA[2], and Noboru OHNISHI[1,3]

1. Bio-Mimetic Control Research Center, RIKEN,
2271-130 Anagahora, Shimoshidami, Moriyama-ku, Nagoya, 463, Japan
E-mail: kawa@nagoya.riken.go.jp,
2. Department of Control Engineering, Kyushu Institute of Technology,
3. Graduate School of Engineering, Nagoya University

Abstract-This paper proposes a method of "blind separation" which extracts non-stationary signals (e.g., speech signals, music) from their convolutive mixtures. The function is acquired by modifying a network's parameters so that a cost function takes the minimum at any time. The cost function is the one introduced by Matsuoka *et al*. [15]. The learning rule is derived from the natural gradient [1] minimization of the cost function. The validity of the proposed method is confirmed by computer simulation.

1. Introduction

This paper deals with *blind separation* of non-stationary signals. Blind separation is a signal processing technique that extracts the original signals (source signals) from their mixtures observed by sensors. An attractive feature of blind separation is that the source signals can be recovered without strong assumption about the source signals or the mixing process, i.e., the transfer function between signal sources and sensors. The only a priori knowledge is, basically, that the source signals are mutually statistically independent. Therefore, this technique can be applied to various fields, e.g., speech recognition, image processing, medical measurement, etc.

Since Herault *et al.* [9,10] presented the problem of blind separation, various methods for blind separation have been proposed. In those methods, as a model of the transfer function between signal sources and sensors, there exit two types of mixing models: instantaneous mixture and convolutive mixture. Instantaneous mixture is a mixing model which does not take into account any delays, whereas convolutive mixture involves some delays in the mixing process. This paper deals with a convolutive mixture in the mixing process. That is, we consider the problem of how the source signals can be separated from their convolutive mixtures.

If the transfer function between signal sources and sensors is estimated, then the source signals can be extracted by applying the inverse dynamics of the transfer function to the observed signals. Therefore, the problem is to estimate the unknown transfer function (or its inverse) by only using the observed signals. The conventional methods [1,3,8,14,17] stipulate that the source signals are non-Gaussian, and utilize information about whether the source signals are super-Gaussian or sub-Gaussian. In these methods, in order to implement blind separation, the algorithms corresponding to the statistical properties of the source signals (super- or sub-Gaussian signals) must be selected by using the observed signals. However, it is difficult to accurately estimate from the observed signals whether the source signals are super- or sub-Gaussian signals because, in a real world, there are a lot of random signals whose statistical properties change with time [16].

In this paper, we assume that the source signals are non-stationary signals (e.g., speech signals, music), and propose a method using non-stationarity of the signals. The proposed method does not require any additional information about whether the source signals are super- or sub-Gaussian. We only make use of the second-order moments of the observed signals. Methods using second-order moments for estimating the unknown transfer function have been proposed by D. C. B. Chan *et al*. [2] and S. V. Gerven *et al*. [6,7]. In [2] source signals are assumed to be stationary signals, and their method must use the same number of cross-correlation data as the degree of freedom of the unknown transfer function. Our method, differently from that, uses only one set of cross-correlation data. The method proposed in [6,7] can not deal with non-minimum phase systems.

The method proposed in this paper estimates the unknown transfer function by modifying the parameters of an adaptive network. The learning rule of the network's parameters is derived from the natural gradient [1] minimization of a cost function. The validity of the proposed method is confirmed by computer simulation. Simulation result will show that our method can be applied to non-minimum phase systems.

2. Source Signals

Suppose that random signals $s_i(t)$ ($i = 1,\dots,N$; $t = \dots, -1,0,1,\dots$) are generated by N mutually independent sources. Henceforth, these signals will be considered as the source signals. The source signals are mixed with the following process, and their mixed signals $x_i(t)$ ($i = 1,\dots,N$) are observed by N sensors:

$$x_i(t) = \sum_{j=1}^{N} \sum_{k=-\infty}^{\infty} a_{ij}(k)s_j(t-k), \quad (i = 1,\dots,N) \tag{1}$$

$$= \sum_{j=1}^{N} \bar{a}_{ij}(z)s_j(t), \tag{2}$$

where $\bar{a}_{ij}(z) = \sum_{k=-\infty}^{\infty} a_{ij}(k)z^{-k}$ is the transfer function from the j-th input (source signal) to the i-th output (mixed signal), and z^{-k} is a delay operator, i.e., $s_i(t)z^{-k} = s_i(t-k)$. Equation (2) can be rewritten in vector notation as

$$\mathbf{x}(t) = \bar{\mathbf{A}}(z)\mathbf{s}(t), \qquad (3)$$

where $\mathbf{x}(t) = [x_1(t),\ldots,x_N(t)]^T$, $\bar{\mathbf{A}}(z) = [\bar{a}_{ij}(z)]$, and $\mathbf{s}(t) = [s_1(t),\ldots,s_N(t)]^T$. We refer to $\mathbf{x}(t)$ (or $x_i(t)$) as observed signal.

Our aim is to extract source signals from the observed signals $x_i(t)$ $(i = 1,\ldots,N)$. To this end, we make the following assumptions.

Assumption 1 $\bar{\mathbf{A}}(z)$ does not have poles or zeros on the unit circle $|z| = 1$.

Assumption 2 $s_i(t)$ $(i = 1,\ldots,N)$ are mutually independent signals with zero mean.

From Assumption 2, the auto-correlation matrix $\mathbf{R}(t,\tau)$ of $\mathbf{s}(t)$ becomes a diagonal matrix:

$$\begin{aligned}\mathbf{R}(t,\tau) &= E[\mathbf{s}(t)\,\mathbf{s}(t-\tau)^T] \\ &= \mathrm{diag}\{E[s_1(t)s_1(t-\tau)],\ldots,E[s_N(t)s_N(t-\tau)]\} \\ &\equiv \mathrm{diag}\{r_1(t,\tau),\ldots,r_N(t,\tau)\} \qquad (4)\end{aligned}$$

where diag{…} represents a diagonal matrix with the diagonal element {…}, and $E[x]$ is the ensemble average of x. The source signals $s_i(t)$ $(i = 1,\ldots,N)$ are assumed to be non-stationary signals, that is,

Assumption 3 For all τ, each $r_i(t,\tau)$ (see eqn (4)) changes independently with time t.

3. Separation Process

An adaptive feedforward network (see Figure 1) is used to separate the source signals from the observed signals. The input signals of the network are the observed signals $x_i(t)$ $(i = 1,\ldots,N)$. The network outputs can be written as:

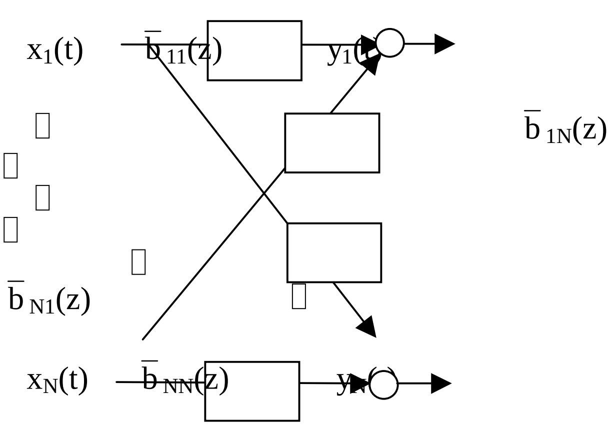


Fig. 1 Signal Separation Network

$$y_i(t) = \sum_{j=1}^{N}\sum_{k=0}^{M} b_{ij}(k)x_j(t-k), \qquad (5)$$

$$= \sum_{j=1}^{N} \bar{b}_{ij}(z)x_j(t), \qquad (i = 1,\ldots,N) \qquad (6)$$

where $\bar{b}_{ij}(z) = \sum_{k=0}^{M} b_{ij}(k)z^{-k}$ $(i, j = 1, \ldots,N)$ represent the transfer function between j-th input (observed signal) and i-th output signal. Eqn (6) can be rewritten in vector notation as

$$\mathbf{y}(t) = \bar{\mathbf{B}}(z)\mathbf{x}(t), \qquad (7)$$

where $\mathbf{y}(t) = [y_1(t),\ldots,y_N(t)]^T$ and $\bar{\mathbf{B}}(z) = [\bar{b}_{ij}(z)] = \sum_{k=0}^{M} \mathbf{B}(k)z^{-k}$, $(\mathbf{B}(k) = [b_{ij}(k)])$.

Substituting eqn (3) into eqn (7), we have

$$\mathbf{y}(t) = \bar{\mathbf{B}}(z)\bar{\mathbf{A}}(z)\mathbf{s}(t). \qquad (8)$$

When $\mathbf{C}(z) \equiv \bar{\mathbf{B}}_0(z)\bar{\mathbf{A}}(z) = \mathbf{D}(z)\mathbf{P}$, the outputs of the network become the filtered and permuted source signals, i.e., $\bar{\mathbf{s}}(t)=[\bar{s}_1(t),\ldots,\bar{s}_N(t)]^T = \mathbf{D}(z)\mathbf{P}\mathbf{s}(t)$. Here, $\mathbf{P}$ is an arbitrary permutation matrix, and $\mathbf{D}(z)$ is a diagonal matrix expressed as

$$\mathbf{D}(z) = \mathrm{diag}\{\sum_{k=-\infty}^{\infty} d_1(k)z^{-k},\ldots,\sum_{k=-\infty}^{\infty} d_N(k)z^{-k}\}.$$

$\bar{s}_i(t)$ $(i=1,\ldots,N)$ can also be regarded as source signals, because $\bar{s}_i(t)$ $(i = 1,\ldots,N)$ are mutually independent signals. Therefore, our goal is now to find the matrix $\bar{\mathbf{B}}_0(z)$ satisfying $\mathbf{C}(z) = \mathbf{D}(z)\mathbf{P}$.

4. Separation Method

When the mixing process is assumed to be an instantaneous mixture, that is, $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$, and $\mathbf{s}(t)$ is non-stationary signal, then Matsuoka, Ohya, and Kawamoto [15] showed that blind separation can be achieved by minimizing the following cost function:

$$Q = \frac{1}{2}\{\sum_{i=1}^{N} \log E[y_i(t)^2] - \log\det E[\mathbf{y}(t)\mathbf{y}(t)^T]\} \qquad (9)$$

In this paper, their method is extended to the case of convolutive mixture by considering the following cost function:

$$Q(t,\bar{\mathbf{B}}(z)) = \frac{1}{2}\{\sum_{i=1}^{N} \log E[y_i(t-L)^2] - \log\det E[\mathbf{y}(t-L)\mathbf{y}(t-L)^T]\} \qquad (10)$$

Note that the parameter L represents a delay. In our method, time t-L is regarded as t = 0. Therefore, our algorithm has access to both future and past values of the observed signals, that is, $\{\mathbf{x}(t),\ldots,\mathbf{x}(t-L+1)\}$ and $\{\mathbf{x}(t-L-1),\ldots,\mathbf{x}(t-M)\}$, respectively. Thanks to that, our proposed algorithm can be applied to non-minimum phase systems. The function given by (10) evaluates only one set of cross-correlation, $E[y_i(t-L)y_j(t-L)]$ $(i, j = 1,\ldots,N;\ i \neq j)$, and data outside that set, for example, $E[y_i(t)y_j(t-\tau)]$ $(i, j = 1,\ldots,N;\ i \neq j;\ \forall\tau)$ are not taken into account.

Matrix $\bar{\mathbf{B}}_0(z)$ (satisfying $\mathbf{C}(z) = \mathbf{D}(z)\mathbf{P}$) is found by minimizing the function $Q(t,\bar{\mathbf{B}}(z))$. In order to minimize the cost function (10) the natural gradient algorithm [1] is used:

$$\Delta\mathbf{B}(k) = -\alpha \frac{\partial Q(t,\overline{\mathbf{B}}(z))}{\partial \mathbf{B}(k)} \overline{\mathbf{B}}(z^{-1})^T \overline{\mathbf{B}}(z),$$

$$= -\alpha \left[\frac{\partial Q(t,\overline{\mathbf{B}}(z))}{\partial b_{ij}(k)}\right] \overline{\mathbf{B}}(z^{-1})^T \overline{\mathbf{B}}(z),$$

$$(k = 0,\ldots,M) \qquad (11)$$

where α is a small positive constant.

Calculating the right-hand side of eqn (11), we have

$$\Delta\mathbf{B}(k) = \alpha z^{-k}\{\mathbf{I} - (\mathrm{diag}E[\mathbf{y}(t-L)\mathbf{y}(t-L)^T])^{-1} \times E[\mathbf{y}(t-L)\mathbf{y}(t-L)^T]\}\overline{\mathbf{B}}(z)$$

$$(k = 0,\ldots,M) \qquad (12)$$

where diag$\mathbf{X}$ represents a diagonal matrix with the diagonal elements of matrix $\mathbf{X}$.

In practice, $E[\mathbf{y}(t-L)\mathbf{y}(t-L)^T]$ is replaced by its instantaneous value $\mathbf{y}(t-L)\mathbf{y}(t-L)^T$:

$$\Delta\mathbf{B}(k) = \alpha z^{-k}\{\mathbf{I} - (\mathrm{diag}E[\mathbf{y}(t-L)\mathbf{y}(t-L)^T])^{-1} \times \mathbf{y}(t-L)\mathbf{y}(t-L)^T\}\overline{\mathbf{B}}(z)$$

$$(k = 0,\ldots,M) \qquad (13)$$

To estimate $\mathrm{diag}E[\mathbf{y}(t-L)\mathbf{y}(t-L)^T]$, we use the following moving average:

$$\phi_i(t) = \beta\phi_i(t-1) + (1-\beta)y_i(t-L)^2$$

$$(i = 1,\ldots,N;\ 0 < \beta < 1) \qquad (14)$$

Then, eqn (13) becomes

$$\Delta\mathbf{B}(k) = \alpha z^{-k}\{\mathbf{I} - \mathbf{\Phi}(t)^{-1}\mathbf{y}(t-L)\mathbf{y}(t-L)^T\}\overline{\mathbf{B}}(z),$$

$$(k = 0,\ldots,M) \qquad (15)$$

where $\mathbf{\Phi}(t) = \mathrm{diag}\{\phi_1(t),\ldots,\phi_N(t)\}$. Eqns (14) and (15) is used to update $\mathbf{B}(k)$ $(k = 0,\ldots,M)$.

Here, let us consider how to remove ambiguity of parameters $d_i(k)$ $(i = 1,\ldots,N;\ -\infty \le k \le \infty)$ in $\mathbf{D}(z)$. In our previous method [12,13], the ambiguity of parameter $d_i(k)$ in $\mathbf{D}(z)$ was removed by fixing the diagonal element $\overline{b}_{ii}(z)$ of matrix $\overline{\mathbf{B}}(z)$ to z^{-L}. In this paper, in order to remove the ambiguity of parameter $d_i(k)$ in $\mathbf{D}(z)$, the normalization of each row of $\overline{\mathbf{B}}(z)$ is implemented:

$$\overline{b}_{ij}(z)/b_{ii}(L) \quad (i, j = 1,\ldots,N)$$

This normalization dose not influence the learning algorithm (12) because even if the output signal $\mathbf{y}(t)$ is multiplied by any nonzero constant, we can prove that the value of the cost function $Q(t,\overline{\mathbf{B}}(z))$ does not change.

5. Simulation Result

We show an example to check the validity of the proposed method.

Example: In this example, there are three sources (N = 3). Source signals $s_1(t)$, $s_2(t)$, and $s_3(t)$ are the following stationary Gaussian signal and two non-stationary signals:

$$s_1(t) = u_1(t),\ s_2(t) = \eta_2(t)u_2(t),\ s_3(t) = \eta_3(t)u_3(t)$$

where $u_i(t)$ $(i = 1, 2, 3)$ are Gaussian white signals with zero mean and unity variance, and $\eta_2(t)$, $\eta_3(t)$ are given by $\eta_2(t) = 2\sin(\pi/200)t$, $\eta_3(t) = 2\sin(\pi/500)t$, respectively. The channel matrix $\overline{\mathbf{A}}(z)$ was given as

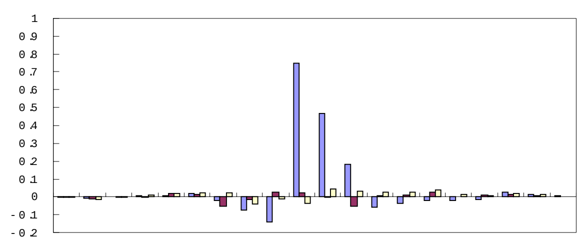
$$\overline{\mathbf{A}}(z) = \begin{bmatrix} z^{-1}+0.4z^{-2} & 0.4+0.4z^{-1} & 0.25z^{-2} \\ 0.4+0.3z^{-1} & 1+0.4z^{-1}+0.2z^{-2} & 0.4+0.4z^{-1} \\ 0.25z^{-3} & 0.5z^{-2} & z^{-1}+0.5z^{-2} \end{bmatrix}$$

The poles of $\overline{\mathbf{A}}(z)^{-1}$ are -0.32-0.52I, -0.32+0.52I, 0.055-0.24I, 0.055+0.24I, -0.67, and 5.19. Therefore, $\overline{\mathbf{A}}(z)^{-1}$ has one pole outside the unit circle $|z| = 1$, from which we deduce that $\overline{\mathbf{A}}(z)$ is a non-minimum phase system. Parameters M of eqn (5) and L of eqn (15) were set to 16 and 8, respectively. The parameters of the learning algorithm were chosen as $\alpha = 0.00025$ and $\beta = 0.9$. The initial values of $b_{ij}(k)$ $(k = 0,\ldots,16;\ i, j = 1, 2, 3;\ i \ne j)$, $b_{ii}(8)$ $(i = 1, 2, 3)$, $b_{ii}(k)$ $(k = 0,\ldots,16;\ k \ne 8;\ i = 1, 2, 3)$, and $\phi_i(t)$ were set to 0, 1, 0, and 1, respectively.

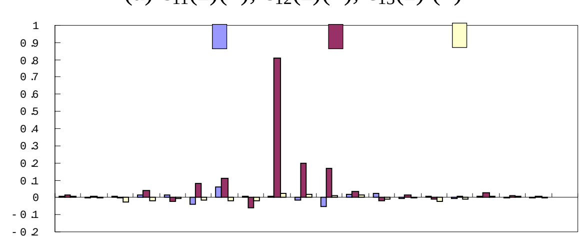
Figure 2 shows the elements $c_{ij}(z)$ $(i, j = 1, 2, 3)$ of the matrix $\mathbf{C}(z) = \overline{\mathbf{B}}_0(z)\overline{\mathbf{A}}(z)$. Each $c_{ij}(z)$ is composed of 20 elements, that is,

$$c_{ij}(z) = \sum_{k=0}^{19} c_{ij}(k)z^{-k}, \quad (i, j = 1, 2, 3).$$

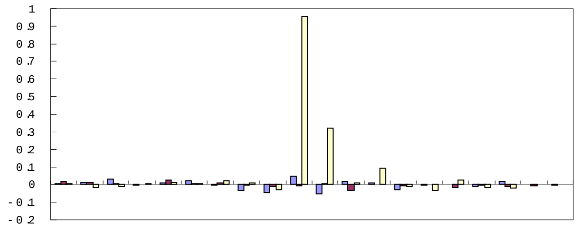
$\overline{\mathbf{B}}_0(z)$ was found after 10000 iterations using our learning algorithm. Fig. 2 shows that the non-diagonal elements $c_{ij}(z)$ $(i, j = 1, 2, 3;\ i \ne j)$ are nearly equal to zero. Therefore, one can see that the proposed algorithm could separate the source signals from their convolutive mixtures.



(a) $c_{11}(z)$( ), $c_{12}(z)$( ), $c_{13}(z)$ ( )



(b) $c_{21}(z)$( ), $c_{22}(z)$( ), $c_{23}(z)$ ( )



(c) $c_{31}(z)$( ), $c_{32}(z)$( ), $c_{33}(z)$( )

Fig. 2 The elements $c_{ij}(z)$ (i, j = 1, 2, 3)

6. Conclusion

We have proposed a method of blind separation for convolved non-stationary signals. This method is an extension of our previous one in [15] for the case of convolutive mixture.

The simulation result has shown that our algorithm of blind separation works well, even if the channel matrix $\overline{\mathbf{A}}$ (z) is a non-minimum phase filter.

References

[1] S. Amari, S. C. Douglas, and A. Cichocki, "Multichannel blind deconvolution and source separation using the natural gradient," submitted to IEEE Trans. SP.

[2] D. C. B. Chan, S. J. Godsill, and P. J. W. Rayner, "Blind signal separation by output decorrelation," NIPS'96 Workshop on Blind Signal Processing and Their Applications (URL:http://www.bip.riken.go.jp/absl/back/nips96ws/nips96ws.html).

[3] A. Cichoki, S. Amari, and J. Cao, "Blind separation of delayed and convolutive signals with self-adaptive learning rate," Proceedings of 1996 International Symposium on Nonlinear Theory and its Applications, pp. 229-232, 1996.

[4] A. Cichoki, I. Sabala, S. Choi, B. Orsier, and R. Szupiluk, "Self-adaptive independent compone-nt analysis for sub-Gaussian and super-Gaussian mixtures with an unknown number of sources and additive noise," Proc. International Symposium on Nonlinear Theory and its Applications, Hawaii, pp.731-734, 1997.

[5] P. Comon, "Independent component analysis, a new concept?," Signal Processing, Vol. 36, pp. 287-314, 1994.

[6] S. Van Gerven, D. Van Compernolle, H. L. Nguyen Thi, and C. Jutten, "Blind separation of sources: a comparative study of a 2-nd and a 4-th order solution," Proceedings of Signal Processing VII, Theories and Applications, pp. 1153-1156, 1994.

[7] S. Van Gerven and D. Van Compernolle, "Signal separation by symmetric adaptive decorrelation; Stability, convergence, and uniqueness," IEEE Trans. on SP, 43(7); pp. 1602-1612, July 1995.

[8] M. Girolami, "The binaural cocktail party effect and blind signal processing," Proc. International ICSC Workshop on Independence & Artificial Neural Networks 98, Tenerife, Spain, pp.1387-1395, 1998.

[9] J. Herault, C. Jutten, and B. Ans, " Detection de grandeurs primitives dans un message composite par une architecture de calcul neuromimetique en apprentissage non supervise," Actes du Xeme colloque GRETSI, 1985.

[10] J. Herault and C. Jutten, "Space and time adaptive signal processing by neural network models," Neural Network for Computing, AIP Conference Proceedings 151, pp. 207-211, 1986.

[11] M. Kawamoto, K. Matsuoka, and M. Oya, "Blind separation of sources using temporal correlation of the observed signals," IEICE Trans. Fundamentals, Vol. E80-A, No. 4, April 1997.

[12] M. Kawamoto, K. Matsuoka, and N. Ohnishi, "Blind signal separation of convolved non-stationary signals," Proc. International Symposium on Nonlinear Theory and its Applications, Hawaii, pp.1001-1004, 1997.

[13] M. Kawamoto, A. K. Barros, and N. Ohnishi, "A neural network for blind separation of convolved non-stationary signals," Proc. International ICSC Workshop on Independence and Artificial Neural Networks 98, Tenerife, Spain, pp.1374-1379, 1998.

[14] T. W. Lee, A. J. Bell, and R. H. Lambert, "Blind separation of delayed and convolved sources," Advances in Neural Information Processing Systems 9, pp. 758-764, 1996.

[15] K. Matsuoka, M. Ohya, and M. Kawamoto, "A Neural Net for Blind separation of Nonstationary Signals," Neural Networks, Vol. 8, No. 3, pp. 411-419, 1995.

[16] A. Mansour, C. Jutten, and N. Ohnishi, "Kurtosis: Definition and Properties," To appear at Fusion98, 1998.

[17] K. Torkkola, "Blind separation of convolutive sources based on information maximization," Neural Networks for Signal Processing VI, pp. 423-432, 1996.