Blind Signal Separation for MA Mixture Model

Xieting Ling, Wei Tian and Bin Liu
Department of Electronic Engineering, Fudan University, Shanghai 200433, China
Ruey-wen Liu
Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN.46556, USA

Abstract - This paper presents a linear feedback neural network and an adaptive algorithm to achieve the blind signal separation under near-field situation. The convergence property of algorithm and stability of equilibrium state are discussed. Some simulations are provided.

I. Introduction

Separating original signals of interest from observations is a fundamental problem in signal processing. In the most cases, sources as well as mixture (channel) are unknown. The blind signal separation is one that separates unknown sources from the observations without knowing the transmission channel. We can find its wide range of applications in array signal processing, sonar system, biomedical signal detecting, and so on. Generally there are two major approaches to this problem. One is achieved by orthogonalizing channel parameter matrix, and the orthogonalized parameter matrix can be obtained from eigen-decomposition (see [1][2]), this approach's weakness lies in the complexity of algorithm and difficulty to be applied in real-time processing. The other uses stochastic approximation method (see [3][4]), this approach can be used in real-time processing. In all of the above methods, only ideal transmission channel is considered, and its phase (delay) differences are skipped, or only narrow-band sinusoidal sources are considered (see [2][6]). In near-field situation, these approximations are inadequate.

In near-field situation, MA mixture model can be adopted. Under independence assumption of sources, applying Hebbian learning rules to build adaptive equations for system parameters, network and algorithm to realize blind signal separation in near-field situation are presented. The stability and convergence property of the approach are analyzed. And some computer simulations are provided to demonstrate the effectiveness.

II. Problem Statement and Separation Network

Problem statement.

Consider the following separation problem

$$x(z) = G(z)s(z) \tag{1}$$

where s(z) is the source vector, x(z) the observation vector, and G(z) the mixture matrix. Denote the transfer function from source $s_j(z)$ to ith receiver is $g_{ij}(z)$, using MA model, $g_{ij}(z)$ can be expressed as

$$g_{ij}(z) = \sum_{m=0}^{m_{ij}} g_{ij}(m) z^{-m} \tag{2}$$

where $g_{ij}(m)=0$ when $m>m_{ij}$, and

$$G(z) = \begin{bmatrix} g_{11}(z) & \dots & g_{1N}(z) \\ \dots & \dots & \dots \\ g_{N1}(z) & \dots & g_{NN}(z) \end{bmatrix} \tag{3}$$

The objective of blind signal separation is to determine s(z) from x(z) without knowing G(z). Note that for simplicity, we assume G(z) is a square matrix, and the transmission noise isn't considered in equation (1) (see simulations in section V).

Separation network.

Fig.1 shows a linear feedback neural network

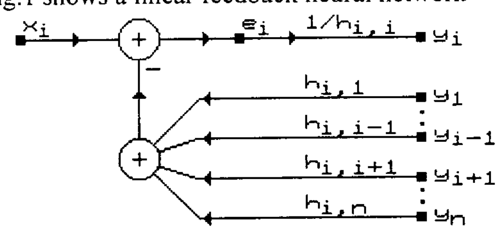


Fig. 1. Network's block diagram

Feed the received time series x(k) to the network, we have

$$e_i(z) = x_i(z) - \sum_{\substack{j=1 \\ j \neq i}}^{N} h_{ij}(z) y_j(z) \tag{4}$$

where $1 \le i \le N$, and

$$y_i(z) = h_{ii}^{-1}(z) e_i(z) \tag{5}$$

where $1 \le i \le N$. Combine (4) and (5), we get

$$h_{ii}(z) y_i(z) + \sum_{\substack{j=1 \\ j \neq i}}^{N} h_{ij}(z) y_j(z)$$

$$= x_i(z) = \sum_{j=1}^{N} g_{ij}(z) s_j(z) \tag{6}$$

Rewrite (6) in matrix form

$$H(z)y(z) = x(z) = G(z)s(z) \tag{7}$$

0-7803-2431-5/95 $4.00 © 1995 IEEE

From (7), once H(z)U=G(z), we can get y(z)=Us(z), where matrix U is a generalized permutation matrix (see [1]). So the architecture of the network allows to separate unknown sources. In blind signal separation, however, both G(z) and s(z) are unknown, y(z) cannot be computed directly. In the next section, we propose an adaptive algorithm to make H(z) approximate G(z).

III. Adaptive Separation Algorithm

Although sources s(k) are independent, the received signals x(k), through transmission channel (mixture), are cross-correlated. The goal of separation network H(z) is to make its outputs y(k) independent again, thus the purpose of signal separation is achieved. Using Hebbian learning rule, the adaptation of coefficients $h_{ij}(m)$ (where $1 \le i,j \le N$ and $0 \le m \le M_{ij}$), will reduce the correlation of network outputs. We propose following adaptation rules

$$\begin{cases} h_{ij}(m,k+1) = h_{ij}(m,k) + \varepsilon y_i(k) y_j(k-m) \\ \qquad \text{where } 1 \le i,j \le N \text{ and } 1 \le m \le M_{ij} \\ h_{ij}(0,k+1) = h_{ij}(0,k) + \varepsilon y_i^3(k) y_j(k) \\ \qquad \text{where } 1 \le i,j \le N \text{ and } i \ne j \\ h_{ii}(0) = 1 \\ \qquad \text{where } 1 \le i,j \le N \end{cases} \tag{8}$$

where ε is a positive adaptation gain.

From (8), we can deduce

$$\begin{cases} \Delta h_{ij}(m) = h_{ij}(m,k+1) - h_{ij}(m,k) = \varepsilon y_i(k) y_j(k-m) \\ \qquad \text{where } 1 \le i,j \le N \text{ and } 1 \le m \le M_{ij} \\ \Delta h_{ij}(0) = h_{ij}(0,k+1) - h_{ij}(0,k) = \varepsilon y_i^3(k) y_j(k) \\ \qquad \text{where } 1 \le i,j \le N \text{ and } i \ne j \end{cases} \tag{9}$$

The equilibrium states c^* of the network are solutions of the following equations

$$\begin{cases} < y_i^*(k) y_j^*(k-m) > = < y_i(c^*, s(k)) y_j(c^*, m, s(k)) > = 0 \\ \qquad \text{where } 1 \le i,j \le N \text{ and } 1 \le m \le M_{ij} \\ < y_i^{*3}(k) y_j^*(k) = < y_i^3(c^*, s(k)) y_j(c^*, s(k)) > = 0 \\ \qquad \text{where } 1 \le i,j \le N \text{ and } i \ne j \end{cases} \tag{10}$$

$$c = [h_{11}(0),.., h_{11}(M_{11}); h_{12}(0),.., h_{12}(M_{12});..; h_{NN}(0),.., h_{NN}(M_{NN})]^T$$

If signal sources and channel mixture satisfy the following two assumptions:

Assumption 1. (zero-mean and independence)

$$1. < s_i(k) > = 0 \qquad \text{where } 1 \le i \le N$$

$$2. < s_i(k) s_j(k-m) > = < s_i(k) > < s_j(k-m) > = 0$$
$$\text{where } 1 \le i,j \le N,\ i \ne j \text{ and } m=0,1,...$$

$$3. < s_i(k) s_i(k-m) > = \begin{cases} \delta_i^2 & 0 \le m \le \tau,\ 1 \le i \le N \\ 0 & m > \tau,\ 1 \le i \le N \end{cases}$$

Assumption 2. (realizability)

All the zero points of $g_{ii}(z)$, $1 \le i \le N$, are located within identity circle, $H^{-1}(z)$ is stable and realizable.

Theorem 1. The state c which achieves signal separation is an equilibrium state of equation (10).

Proof: denote y(c,z)=D(c,z)s(z), we have

$$\begin{cases} y_i(c,k) = \sum_{\alpha=1}^{N} \sum_{n=0}^{\infty} d_{i\alpha}(c,n) s_\alpha(k-n) \\ y_j(c,k-m) = \sum_{\beta=1}^{N} \sum_{l=0}^{\infty} d_{j\beta}(c,l) s_\beta(k-l-m) \end{cases} \tag{11}$$

where $1 \le i, j \le N$. From assumption 1, we get

$$< y_i(c,k) y_j(c,k-m) > = \sum_{\alpha=1}^{N} \sum_{n=m}^{\infty} d_{i\alpha}(c,n) d_{j\alpha}(c,n-m) \delta_\alpha^2$$
$$\text{where } 1 \le i,j \le N,\ 1 \le m \le M_{ij} \text{ and } i \ne j \tag{12}$$

$$< y_i(c,k) y_i(c,k-m) > = \sum_{\alpha=1}^{N} \sum_{\substack{n=m-\tau \\ n \ge 0}}^{\infty} \sum_{t=0}^{\tau} d_{i\alpha}(c,n) d_{j\alpha}(c,n-m+t) \delta_\alpha^2$$
$$\text{where } 1 \le i \le N \text{ and } 1 \le m \le M_{ij} \tag{13}$$

$$< y_i^3(c,k) y_j(c,k) > =$$

$$\sum_{\alpha=1}^{N} \sum_{n=0}^{\infty} \{ d_{i\alpha}^3(c,n) d_{j\alpha}(c,n) + 3 \sum_{\substack{r=0 \\ r \ne n}}^{\infty} d_{i\alpha}^2(c,n) d_{i\alpha}(c,r) \} \delta_\alpha^4 +$$

$$3 \sum_{\alpha=1}^{N} \sum_{\beta=1}^{N} \sum_{n=0}^{\infty} \{ d_{i\alpha}^2(c,n) d_{i\beta}(c,n) d_{j\beta}(c,n) + \sum_{\substack{r=0 \\ r \ne n}}^{\infty} d_{i\alpha}^2(c,n) d_{i\beta}(c,r) d_{j\beta}(c,r) \} \delta_\alpha^2 \delta_\beta^2$$
$$\text{where } 1 \le i,j \le N,\ i \ne j \tag{14}$$

Because state c can fulfill signal separation, so D(c, z) is a generalized permutation matrix. From this fact we can deduce all coefficients of δ_α^2, δ_α^4 and $\delta_\alpha^2 \delta_\beta^2$ in equations (12-14) are equal to zero. Hence the state c which achieves signal separation is an equilibrium state of equation (10).

Theorem 2. Assume $M_{ij}=M^*$, $1 \le i, j \le N$, when M^* is large enough, the equilibrium states c^* of equation (10) always achieves signal separation.

Proof: From assumption 2, $H^{-1}(z)$ is stable, so $d_{i\alpha}(c^*,n)$ (where $n>M^*$)equals to zero when M^* is large enough. We have

$$y_i(c^*,k)=\sum_{\alpha=1}^{N}\sum_{n=0}^{M^*} d_{i\alpha}(c^*,n)s_\alpha(k-n) \quad (15)$$

where $1 \le i \le N$, c^* is the equilibrium state of (10), so

$$\begin{cases} <y_i(c^*,k)y_j(c^*,k-m)>=0 & 1\le i,j\le N \text{ and } 1\le m\le M^* \\ <y_i^3(c^*,k)y_j(c^*,k)>=0 & 1\le i,j\le N \text{ and } i\ne j \end{cases} \quad (16)$$

From (12) and (16), we can get lemma 1.

Lemma 1. If there exists a component $d_{i\alpha}(c^*,n_\alpha)\neq 0$ in vector $d_\alpha(c^*,n_\alpha)$(where $1\le i\le N$ and $0\le n_\alpha \le M^*$), then the other vector $d_\alpha(c^*,n)=0$(where $0\le n\le M^*$ and $n\ne n_\alpha$).

From lemma 1, we know $N_0\le N$(N_0 is the number of non-zero vector $d_\alpha(c^*,n_\alpha)$). If $N_0<N$, in practice, some sources will not appear in network output. This is a little possibility state, generally we always have $N_0=N$.

From (13-14) and (16), lemma 2 can be deduced.

Lemma 2. In every vector $d_\alpha(c^*,n_\alpha)$, there exists only one non-zero component $d_{i(\alpha)\alpha}(c^*,n_\alpha)$. Moreover, $i(\alpha_1)\ne i(\alpha_2)$ when $\alpha_1\ne\alpha_2$.

Combine lemma 1 and lemma 2, we know $D(c^*, z)$ is a generalized permutation matrix, hence the theorem has been proved.

IV. Analysis for Stability and Convergence

Stability of equilibrium state.

Denote $\delta c=c-c^*$, adopt linear approximation, we have

$$\begin{cases} \dfrac{1}{\varepsilon}\dfrac{d<\delta h_{ij}(m)>}{dt}=\sum_{p=1}^{N}\sum_{q=1}^{N}\sum_{r=1}^{M^*}\left(\dfrac{\partial<y_i(k)y_j(k-m)>}{\partial h_{pq}(r)}\right)_{c=c^*}<\delta h_{pq}(r)> \\ \dfrac{1}{\varepsilon}\dfrac{d<\delta h_{ij}(0)>}{dt}=\sum_{p=1}^{N}\sum_{q=1}^{N}\sum_{r=1}^{M^*}\left(\dfrac{\partial<y_i^3(k)y_j(k)>}{\partial h_{pq}(r)}\right)_{c=c^*}<\delta h_{pq}(r)> \end{cases} \quad (17)$$

We can get (the deduction procedure is omitted)

$$\frac{1}{\varepsilon}\frac{d<\delta h_{ii}(m)>}{dt}=-\begin{bmatrix} b_i(ii,0) & 0 & \dots & 0 \\ b_i(ii,1) & b_i(ii,0) & \dots & 0 \\ \dots & \dots & \dots & 0 \\ b_i(ii,M^*-1) & b_i(ii,M^*-2) & \dots & b_i(ii,0) \end{bmatrix}<\delta h_{ii}(m)> \quad (18)$$

where $1 \le i \le N$, and

$$\sum_{n=0}^{\infty} b_i(pq,n)z^{-n}=\frac{\Delta^*_{pq}(z)}{\Delta^*(z)} \quad (19)$$

where $\Delta^*(z)$ is the determinant of $H(c^*,z)$, and $\Delta^*_{pq}(z)$ is the complementary minor of $H_{pq}(c^*,z)$.

From (12), if $\varepsilon>0$ and $b_i(ii,0)>0$ ($b_i(ii,0)$ is decided by G(z)), equilibrium state c^* must be stable.

Convergence of algorithm.

Theorem 3. Under assumption 1 and 2, the convergence of adaptation algorithm (8) can be guaranteed by properly choosing gain ε.

Proof: System's energy function is defined as

$$E(k)=<y^T(k)y(k)>=\sum_{i=1}^{N}<y_i^2(k)>=\sum_{i=1}^{N}E_i(k) \quad (20)$$

From network's architecture, we have

$$\Delta E_i(k)=2\sum_{p=1}^{N}\sum_{r=0}^{M^*}<y_i(k)\frac{\partial y_i(k)}{\partial h_{ip}(r)}\Delta h_{ip}(r)> \quad (21)$$

where $1 \le i \le N$, and

$$\frac{\partial y_i(k)}{\partial h_{ip}(r)}=-y_p(k-r)-\sum_{\substack{j=1\\ j\ne i}}^{N}h_{ij}(0)\frac{\partial y_j(k)}{\partial h_{ip}(r)} \quad (22)$$

where $1 \le i \le N$ and $1 \le r \le M^*$

If network's feedback effect is omitted, the 2th part in (22)'s right side equals to zero, using adaptation algorithm described in (8), we have

$$\Delta E_i(k)=-2\varepsilon\{\sum_{j=1}^{N}\sum_{r=0}^{M^*}<y_i^2(k)y_j^2(k-m)>+\sum_{\substack{j=1\\ j\ne i}}^{N}<y_i^4(k)y_j^2(k)>\}\le 0 \quad (23)$$

where $1 \le i \le N$

So the adaptation algorithm is asymptotic convergent.

If the feedback effect cannot be omitted, the adaptation algorithm must be changed to

$$h_{ij}(m,k+1)=h_{ij}(m,k)+\varepsilon\frac{\Delta_{ii}(H(0))}{\Delta H(0)}y_i(k)y_j(k-m) \quad (24)$$

where $1 \le i \le N$ and $1 \le m \le M^*$

$$H(0)=\begin{bmatrix} 1 & h_{12}(0) & \dots & h_{1n}(0) \\ h_{21}(0) & 1 & \dots & h_{2n}(0) \\ \dots & \dots & \dots & \dots \\ h_{n1}(0) & h_{n2}(0) & \dots & 1 \end{bmatrix} \quad (25)$$

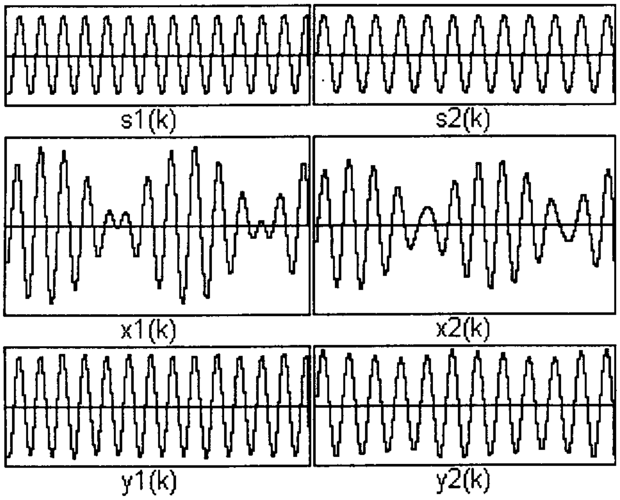
where $\Delta H(0)$ is the determinant of H(0), $\Delta_{ii}(H(0))$ is the complementary minor of $H_{ii}(0)$. We also have $\Delta E_i\le 0$. So the modified algorithm is still asymptotic convergent.

V. Simulations

Simulation 1. Separation of two sources

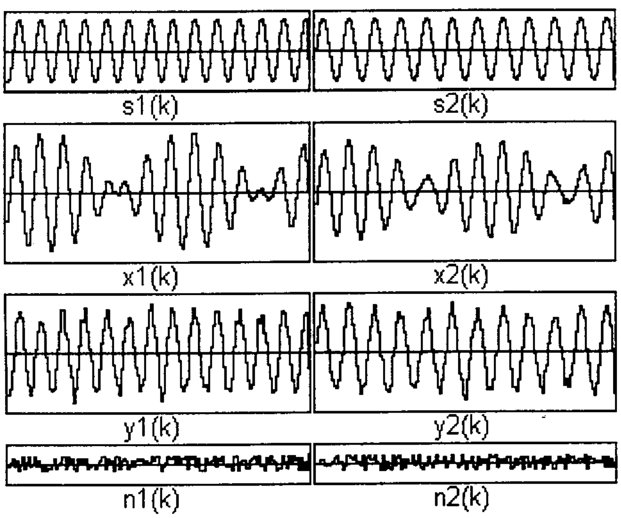
$$g_{11}(z)=1+0.35z^{-1}+0.01z^{-2}+0.001z^{-3} \quad g_{12}(z)=0.8+0.5z^{-1}-0.1z^{-2}+0.01z^{-3}$$
$$g_{21}(z)=0.49+0.24z^{-1}+0.098z^{-2}+0.008z^{-3} \quad g_{22}(z)=1+0.25z^{-1}+0.1z^{-2}+0.01z^{-3}$$

Gains are chosen as: $\varepsilon_{ii}=10^{-5}$, $\varepsilon_{ij}=10^{-4}$, $i \neq j$. Simulation results are:



Simulation 2. Separation of sources with noises

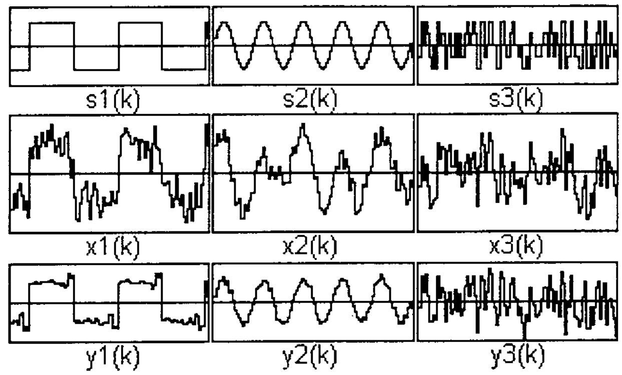
Add stochastic noise $n_i(k)$ to $x_i(k)$ in simulation 1, i=1,2. Simulation results are:



Simulation 3. Separation of three sources

$$g_{11}(z)=1+0.2z^{-1}+0.03z^{-2}+0.001z^{-3} \quad g_{12}(z)=-0.178-0.083z^{-1}+0.226z^{-2}+0.282z^{-3}$$
$$g_{13}(z)=0.35-0.414z^{-1}-0.275z^{-2}-0.09z^{-3} \quad g_{21}(z)=0.005-0.054z^{-1}-0.476z^{-2}-0.061z^{-3}$$
$$g_{22}(z)=1+0.1z^{-1}+0.02z^{-2}-0.003z^{-3} \quad g_{23}(z)=0.137-0.11z^{-1}+0.057z^{-2}+0.347z^{-3}$$
$$g_{31}(z)=0.167+0.282z^{-1}-0.461z^{-2}+0.237z^{-3}$$
$$g_{32}(z)=-0.268-0.343z^{-1}-0.331z^{-2}+0.434z^{-3} \quad g_{33}(z)=1+0.25z^{-1}+0.13z^{-2}+0.002z^{-3}$$

Gains are chosen as: $\varepsilon_{ii}=1.2\times10^{-6}$, $\varepsilon_{ij}=0.5\times10^{-4}$, $i \neq j$. Simulation results are:



VI. Conclusion

Due to the asymptotic convergence of the adaptation algorithm based on Hebbian learning rules, as well as the network's outputs at equilibrium state are orthogonal, general sources' blind separation can be achieved by properly choosing network initial state. Computer simulations show that even in the case that there are some additive noises in received signals, the separation can still be achieved. Only there will emerge some proportional noises in network outputs. Generally, signal's equalization process is slower than signal's separation process, so gain ε_{ii} must be chosen smaller than ε_{ij}.

References

[1] L.Tong *et al.*, "Indeterminacy and identifiability of blind identification", *IEEE Trans. on Circuits and Systems*, vol.38, pp.499-509, 1991

[2] S.Haykin, *Advance in spectrum analysis and array processing*, vol.II, Prentice-Hall Inc., 1991

[3] J.Herault *et al.*, "Detection de grandeurs primitives dans un message composite par une architecture de calcul neuromimetique en apprentissage non supervise", *Proc. Xeme colloque GRETSI*, pp.1017-1022, Nice, France, May 20-24, 1985

[4] X.T.Ling and R.Liu, "A stability theory of blind signal separation", *1993 International Symposium on Nonlinear Theory and Its Applications*, Hawaii, Dec. 5-9, 1993

[5] B.Hu and X.T.Ling, "Blind equalization based on Hebbian unsupervised learning rule: (I) minimum phase channel, (II) nonminimum phase channel", Acta Communica Sinica, to appear

[6] X.T.ling, "Self-learning blind separation of narrow-band signals with delay", Acta Electronica Sinica, to appear

[7] G.Mirchandani *et al.*, "A new adaptive noise cancellation scheme in the presence of crosstalk", *IEEE Trans. Circuits and Systems (II)*, vol.39, pp.681-694, 1992