Blind separation of sources using higher-order cumulants

Bin-Chul Ihm, Dong-Jo Park*

Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, 373-1 Kusong-dong, Yusong-gu, Taejon, 305–701, South Korea

Received 26 October 1997; received in revised form 12 October 1998

Abstract

In this paper, we propose a novel method of blind signal separation using a new performance criterion with nonlinear separating functions, and derive an adaptive algorithm. The criterion consists of nonlinear mappings of output signals and their cumulants for independence of the signals. We show that our method requires a weak condition for stable separation through the ODE analysis. Simulation results are given to verify the validity and advantage of the proposed algorithm.

Zusammenfassung

In diesem Artikel schlagen wir eine neuartige Methode zur blinden Signalseparation vor, die ein neues Leistungskriterium basierend auf nichtlinearen Separationsfunktionen verwendet. Daraus wird ein adaptiver Algorithmus abgeleitet. Das Kriterium geht aus nichtlinearen Abbildungen der Ausgangssignale und ihrer Kumultanten zur Unabhängigkeit der Signale hervor. Durch eine auf gewöhnlichen Differenzialgleichungen basierende Analyse zeigen wir, daß unsere Methode unter schwachen Voraussetzungen auf eine stabile Separation führt. Es werden Simulationsergebnisse vorgelegt, um die Gültigkeit und die Vorteile des vorgeschlagenen Algorithmus zu überprüfen. © 1999 Elsevier Science B.V. All rights reserved.

Résumé

Dans cet article, nous proposons une nouvelle méthode de séparation de signaux aveugle utilisant un nouveau critère de performance utilisant des fonctions de séparation non-linéaires, et nous en dérivons un algorithme adaptatif. Le critère consiste en des appariements non-linéaires des signaux de sortie et de leurs cumulants pour l’indépendance des signaux. Nous montrons que notre méthode demande une condition faible pour une séparation stable au travers d’une analyse par EDO. Des résultats de simulations sont fournis pour vérifier la validité et l’avantage de l’algorithme proposé. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Cumulant; Blind separation; Weight adaptation

*Corresponding author. Tel.: + 82 42 869 3438; fax: + 82 42 869 3410; e-mail: djpark@eekaist.kaist.ac.kr
1. Introduction

The problem of blind separation of sources arises in many signal processing applications like communications, array processing, speech analysis and speech recognition. In all these instances, the underlying assumption is that several linear mixtures of unknown, random, zero-mean, and statistically independent signals which are called sources are observed; the problem consists of recovering the original signals from their mixtures without a priori information of coefficients of the mixtures and knowledge of the sources. Mathematically, the situation is described by the simple and well-known data model:

\[ X(t) = A S(t), \quad t = 1, 2, \ldots, \quad (1) \]

where \( X(\cdot) \) and \( S(\cdot) \) are column vectors of sizes of \( n \) and \( m \), respectively, and \( A \) is an \( n \times m \) matrix.

The idea here is that the vector \( X(\cdot) \) results from measurements by \( n \) sensors receiving contributions from \( m \) sources. The matrix \( A \) is called as a mixing matrix. Our task is to obtain a good estimate \( C \) of \( A^{-1} \) (a pseudo-inverse when \( n \neq m \)) called as a separating matrix and to retrieve the sources by measurement outputs:

\[ Y(t) = C X(t). \quad (2) \]

However as shown in [18], there are indeterminacies in this problem, and \( A^{-1} \) itself cannot be identified. What we can do is to obtain a rescaled and permuted version of the \( m \) source signals. That is, \( C = DPA^{-1} \) where \( D \) and \( P \) are a diagonal matrix and a permutation matrix, respectively.

In the recent years, since the remarkable work was introduced by Jutten and Hérault [14], various approaches have been presented in the literature [1,4–13,15–18]. In [14], the separating matrix is updated by odd nonlinear functions, which are called separating functions later, in order to achieve the independence of the processed output signals \( Y(\cdot) \). Therein, the conditions of successful separation depend on characteristics of the source signals and the separating functions. For example, if a cubic (i.e., \( x^3 \)) and a linear (i.e., \( x \)) functions are used as the separating functions, sub-Gaussian signals (kurtosis < 3) are separable under assumptions with unit variances of the sources [16,17]. The Jutten–Hérault algorithm is inspired by a neuromimetic approach; this line is further followed by Karhunen et al. [15] and Cichocki et al. [8,12].

Later, Cardoso [5] and Comon [10] considered the nonadaptive source separation problem, and proposed different solutions based on the fourth-order statistics of \( Y(\cdot) \). In [10], Comon defined the contrast, which was the square sum of the fourth-order cumulants, for the independence criterion. The contrast was used in other papers [11,6]. In these papers, the stability conditions can never be met if there are more than one Gaussian source signal. On the other hand, based on information theoretic ideas, Bell et al. proposed the nonlinear distortion of the output \( Y(\cdot) \), i.e., \( Y(t) = g(X(t)) \) [1]. When \( g(\cdot) \) has a hyperbolic tangent nonlinearity and the sources have super-Gaussian (kurtosis > 3) distributions, the sources can be separable by maximizing the joint entropy of \( Y(\cdot) \).

In this paper, we propose a new method for signal separation that is based on a new criterion for the case of the square separating matrix \( A \), namely, \( n = m \). Section 2 presents this new algorithm derived from the criterion. Section 3 analyses the stationary points of the proposed performance criterion. Finally, Section 4 gives the results of computer experiments, and Section 5 is devoted to conclusions.

2. New performance criterion and learning algorithm

2.1. Cumulants

In the blind source separation problem, we cannot know characteristics of source signals and channels a priori, therefore we use the only assumption that the source signals are statistically independent. So we need to discuss the statistical independence before we proceed. If we denote \( f_s(s_i) \) as a marginal probability density function of random variables \( s_1, s_2, \ldots, s_n \) and \( f_{s_1 s_2 \ldots s_n}(s_1, s_2, \ldots, s_n) \) as a joint probability density function. And if \( s_1, s_2, \ldots, s_n \) are independent, then the joint probability density function is represented by the product of
the marginal probability density function of each variable, i.e.,
\[ f_{s_1, s_2, \ldots, s_n}(s_1, s_2, \ldots, s_n) = \prod_{j=1}^{n} f_s(s_j). \]  

However, since we cannot estimate the probability density function of the signals, we use the cumulants as a measure of statistical independence. The cumulant is defined as follows [3]:
\[ \text{cum}(s_1, s_2, \ldots, s_n) = \sum (-1)^{p-1}(p-1)! \left( E[\prod_{j \in I_1} s_j] \cdots E[\prod_{j \in I_p} s_j] \right), \]  
where the summation extends over all partitions \((I_1, I_2, \ldots, I_p)\), \(p = 1, 2, \ldots, n\), of \(1, 2, \ldots, n\). The cumulant has a cumulant generating function as the moment. For random signals \(s_1, s_2, \ldots, s_n\), the cumulant generating function is defined as follows [3]:
\[ \psi(v_1, v_2, \ldots, v_n) = \log E(e^{\sum v_i}). \]  
\[ \text{cum}(s_1, s_2, \ldots, s_n) \]  
\[ \triangleq \sum_{i=0}^{\infty} a_i s_1^i. \]

Similarly, we can expand an infinitely differentiable function \(g(\cdot)\) with respect to \(s_2\) around the origin to obtain \(g(s_2) = \sum_{j=0}^{\infty} b_j s_2^j\). If we take the expectation of \(f(s_1)g(s_2)\), we have
\[ E\{f(s_1)g(s_2)\} = E\left\{ \sum_{i=0}^{\infty} a_i s_1^i \sum_{j=0}^{\infty} b_j s_2^j \right\} \]
\[ = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_i b_j E(s_1^i s_2^j). \]

Product of the expectation of \(f(s_1)\) and that of \(g(s_2)\) is obtained as
\[ E\{f(s_1)\}E\{g(s_2)\} = E\left\{ \sum_{i=0}^{\infty} a_i s_1^i \right\}E\left\{ \sum_{j=0}^{\infty} b_j s_2^j \right\} \]
\[ = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_i b_j E(s_1^i)E(s_2^j). \]

Therefore,
\[ \text{cum}(f(s_1), g(s_2)) = E\{f(s_1)g(s_2)\} - E\{f(s_1)\}E\{g(s_2)\} \]
\[ = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_i b_j \text{cum}(s_1^i, s_2^j). \]

If \(s_1\) and \(s_2\) are independent, \(\text{cum}(s_1^i, s_2^j)\) is zero for all \(i, j\) and these make \(\text{cum}(f(s_1), g(s_2))\) be zero. Therefore, cum\((f(s_1), g(s_2)) = 0\) is a necessary condition in order for those \(s_1\) and \(s_2\) to be independent. For independent random signals \(s_1, s_2, \ldots, s_n\), elements \(y_1, y_2, \ldots, y_n\) of \(Y = CX = CAS = WS\), where \(W = CA\), satisfy the following theorem [9].

2.2. New performance criterion

Since, in practice, it is not possible to estimate the cross-cumulants of the whole orders, we use nonlin-
Theorem 1. The following three statements are equivalent:
1. \( y_1, y_2, \ldots, y_n \) are pair-wisely independent.
2. \( y_1, y_2, \ldots, y_n \) are mutually independent.
3. \( W = DP \), where \( D \) is a diagonal matrix and \( P \) is a permutation matrix.

If we make all pairs of \( n \) signals be independent, then the signals become mutually independent. Based on this and the above necessary condition, we define a new performance criterion [13]. From now on we abbreviate \( f(y_i) \) and \( g(y_j) \) as \( f_i \) and \( g_p \), respectively, for convenience’s sake.

Definition 1. A new performance criterion for separation of sources is defined as
\[
J = \| M - \text{diag} M \|_F^2 ,
\]
where
\[
M = \begin{bmatrix}
\text{cum}(f_1,g_1) & \text{cum}(f_1,g_2) & \cdots & \text{cum}(f_1,g_n) \\
\text{cum}(f_2,g_1) & \text{cum}(f_2,g_2) & \cdots & \text{cum}(f_2,g_n) \\
\vdots & \vdots & \ddots & \vdots \\
\text{cum}(f_n,g_1) & \text{cum}(f_n,g_2) & \cdots & \text{cum}(f_n,g_n)
\end{bmatrix} ,
\]
\[
\text{diag} M = \begin{bmatrix}
M_{11} & 0 & \cdots & 0 \\
0 & M_{22} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & M_{nn}
\end{bmatrix},
\]
\[
\| M \|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij}^2} ,
\]
where \( \| \cdot \|_F \) is the Frobenius norm.

2.3. Learning algorithm

We reformulate the problem as that of minimizing the performance criterion and will discuss the learning of a separator matrix in order to get a desirable value. Let the separator matrix \( C \) be as follows:
\[
Y = CX ,
\]
where
\[
C = \begin{bmatrix}
1 & c_{12} & \cdots & c_{1n} \\
c_{21} & 1 & \cdots & c_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
c_{n1} & c_{n2} & \cdots & 1
\end{bmatrix} .
\]

In order to minimize the performance criterion, the separator matrix \( C \) is to be updated by using the steepest descent method as follows:
\[
C(k+1) = C(k) + \mu \Delta C(k) ,
\]
where
\[
\Delta C(k) = \frac{\partial J}{\partial C(k)} .
\]

The performance criterion \( J \) is represented by the sum of squares of elements in \( M \) except the diagonal terms as follows:
\[
J = \sum_{i=1}^{n} \sum_{j \neq i} \left( \text{cum}(f_i,g_j) \right)^2 - \sum_{i=1}^{n} \sum_{j \neq i} \left( E(f_i) - E(f_i)E(g_j) \right)^2 .
\]

By differentiating \( J \) partially about an element \( c_{ab} \), we can obtain the following equation:
\[
\frac{\partial J}{\partial c_{ab}} = 2 \sum_{i \neq a} (E(f_a g_i) - E(f_a)E(g_i)) \cdot (E(f_a^{(1)} x_b g_i) - E(f_a^{(1)} x_b)E(g_i))
\]
\[
+ 2 \sum_{i \neq a} (E(g_a f_i) - E(g_a)E(f_i)) \cdot (E(g_a^{(1)} x_b f_i) - E(g_a^{(1)} x_b)E(f_i))
\]
\[
= 2 \sum_{i \neq a} \left[ \text{cum}(f_a,g_i) \cdot \text{cum}(f_a^{(1)} x_b g_i) \right]
\]
\[
+ \text{cum}(g_a,f_i) \cdot \text{cum}(g_a^{(1)} x_b f_i) \right] .
\]

That is, the partial derivatives can be defined as
\[
\frac{\partial J}{\partial c_{ab}} = H_{ab}(C,X,Y) .
\]

Example 1. Let the matrices and functions be
\[
A = \begin{bmatrix}
1 & 0.5 \\
0.5 & 1
\end{bmatrix},
\]
\[
C = \begin{bmatrix}
1 & c_{12} \\
c_{21} & 1
\end{bmatrix} .
\]
\( f(x) = \tanh(x) \) and \( g(x) = \tanh(0.5x) \). \( s_1 \) and \( s_2 \) are random signals whose probability density functions are uniformly distributed on \([-1,1]\).

Fig. 1 shows the shape of the performance criterion \( J \) about the axes \( c_{12} \) and \( c_{21} \) and trajectories of those elements which have different initial points are given in Fig. 2. In this example, equilibria of \((c_{12}, c_{21})\) are \((-0.5, -0.5)\) and \((-2, -2)\) as shown in the figure. We can know that after sufficient learning, \((c_{12}, c_{21})\) goes to one of the equilibria. Fig. 3 illustrates the learning result of \( CA \). In this experiment, we initialize \( C \) by the identity matrix \( I \).

2.4. Phantom solutions

If the separator matrix \( C \) converges and stays in the steady state, then \( \Delta C(k) \) becomes 0 to make \( \text{cum}(f_i, g_j) = 0 \) for \( i \neq j \). However since \( \text{cum}(f_i, g_j) = 0 \) is a necessary condition for independence of \( y_i \) and \( y_j \), the matrix \( C \) may have phantom solutions. Phantom solutions can make the performance criterion \( J \) be zero but cannot separate the measured signals to recover the source signals. For example, assume \( n = 2 \) and \( f(x) = g(x) = e^x - 1 \). After sufficient learning, if \( \text{cum}(f_1, g_2) = 0 \) is made, then \( E(e^{y_1} e^{y_2}) = E(e^{y_1}) E(e^{y_2}) \) is established.

If we denote \( W = CA \), then outputs \( y_1 \) and \( y_2 \) become \( w_{11}s_1 + w_{12}s_2 \) and \( w_{21}s_1 + w_{22}s_2 \), respectively, which leads to \( E(e^{w_{11}s_1 + w_{12}s_2}) E(e^{w_{21}s_1 + w_{22}s_2}) \) to result in the following solutions:

\[
W = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \text{ or } \begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix}. \tag{25}
\]

The first two of the above solutions are desirable to recover the source signals, but the last one cannot separate the measured signals, though it makes the performance criterion \( J \) have a minimum value. The latter case happens in case of source signals whose probability density functions are impulsive.
When more than one Gaussian source exist, we may obtain extra solutions that make the cost function be minimum and the outputs be independent. The proposed cost function has the following physical meaning for independence:

\[ \text{cum}(y_i^p, y_j^q) = 0, \quad i \neq j, \ i,j = 1,2, \ldots, n, \ p,q = 1,2,3,\ldots. \]  

(26)

In case of Gaussian sources, it is sufficient that \[ \text{cum}(y_i, y_j) = 0, \ i \neq j. \]  

\( Y = \{y_1, y_2, \ldots, y_n\} \) has elements which consist of linear combinations of Gaussian signals, therefore the components are also Gaussian. In this case, if \( R_y = E(YY') \) is a diagonal matrix, \( y_i, i = 1,2,\ldots,m \) are independent. \( R_y = CAR_yA^{\text{T}} \) where, without loss of generality, we can assume that \( R_y = I \) is a diagonal matrix and this leads to \( CA = DG \) where \( G \) is a general orthogonal matrix [4]. This means that by decorrelating the output, we can obtain the independent components from the mixture of Gaussian sources, however they may not be the same as the original sources. Only if \( G \) is a permutation matrix, the sources can be separated.

3. Stability analysis

The ordinary differential equation (ODE) method is useful to study the asymptotic behaviour of an algorithm such as

\[ \theta_{n+1} = \theta_n - \mu_n \psi(\theta_{n}, \xi_n), \]  

(27)

where \( \xi_n \) is a stationary sequence of random variables, and \( \mu_n \) is a sequence of positive numbers [2]. Since our algorithm has the above form, that is, we can replace \( \theta, \psi \) and \( \xi \) by \( C, H \) and \( \{X,Y\} \), respectively, we investigate the stability using the ODE method in the case of \( n = m = 2 \).

A stationary point \( \theta_\ast \) verifies \( E\psi(\theta_\ast, \xi) = 0 \) and is said to be stable if all the eigenvalues of a matrix \( \Gamma \) defined as

\[ \Gamma = \left. \frac{\partial E\psi(\theta, \xi)}{\partial \theta} \right|_{\theta = \theta_\ast} \]  

(28)

have positive real parts.

Since the separating points are equilibria of algorithm (20), \( \theta_\ast \) is one of the separating points of \( C \).

For \( n = m = 2 \), there are two separating points of \( C \). The first point is called a natural separating point \( C_1 \) because it restores \( s_1 \) and \( s_2 \) in the natural order. The second is a reverse separating point \( C_2 \) implying the reverse order.

If we denote \( \theta = [c_{12}, c_{21}]^T \), then from Eqs. (20) and (23) \( \Gamma \) at \( C_i \) can be described by

\[ \Gamma(C_i) = \begin{bmatrix} \frac{\partial^2 J}{\partial c_{12}^2} & \frac{\partial^2 J}{\partial c_{12} \partial c_{21}} \\ \frac{\partial^2 J}{\partial c_{12} \partial c_{21}} & \frac{\partial^2 J}{\partial c_{21}^2} \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}, \quad i = 1,2 \]  

(29)

(30)

For stability, \( \Gamma(C_i) \) needs to have eigenvalues which have positive real parts, and this can be checked easily. \( \Delta \Gamma(C_i) \)'s are roots of the following polynomial equation:

\[ \lambda^2 - (\alpha + \gamma)\lambda + \alpha\gamma - \beta^2 = 0. \]  

(31)

We obtain two eigenvalues

\[ \lambda_{1,2} = \frac{(\alpha + \gamma) \pm \sqrt{(\alpha - \gamma)^2 + 4\beta^2}}{2}. \]  

(32)

To satisfy the positivity of real parts, the following two conditions are sufficient:

\[ \alpha + \gamma > 0, \]  

(33)

\[ \alpha\gamma - \beta^2 > 0. \]  

(34)

Fig. 4. Variations of \( W = CA \) in case of mixture of sub- and super-Gaussian signals.
Fig. 5. Separation result of train horn sound and bird chirp sound.
The first condition is satisfied since $C_i$s are local minima. In the natural separating point $C_1$, the outputs $y_1$ and $y_2$ are scaled $s_1$ and $s_2$, respectively, i.e., $y_i = \phi_{1,i} s_i$ for $i = 1, 2$ for some scaling factors $\phi_{1,i}$. From this fact and the zero-mean assumptions, $\alpha$, $\beta$ and $\gamma$ are written as
\begin{align}
\alpha &= 2a_{22}^2 \left[ E^2(f_{1}^{(1)}) E^2(g_{2}s_2) + E^2(g_{1}^{(1)}) E^2(f_{2}s_2) \right], \quad (34) \\
\beta &= 2a_{11}a_{22} \left[ E(f_{1}s_1) E(g_{2}s_2) E(f_{1}^{(1)}) E(g_{2}^{(1)}) \right. \\
&\quad \left. + E(g_{1}s_1) E(f_{2}s_2) E(g_{1}^{(1)}) E(f_{2}^{(1)}) \right], \quad (35) \\
\gamma &= 2a_{11}^2 \left[ E^2(f_{1}^{(1)}) E^2(g_{1}s_1) + E^2(g_{2}^{(1)}) E^2(f_{2}s_2) \right]. \quad (36)
\end{align}
where $f_i$ and $g_i$ denote $f(\phi_{1,i} s_i)$ and $g(\phi_{1,i} s_i)$, respectively, for $i = 1, 2$. Through some manipulations, we can obtain the following relation:
\begin{equation}
\alpha \gamma - \beta^2 = a_{11}^2 a_{22}^2 (U_1 - V_1)^2 \geq 0, \quad (37)
\end{equation}
where $U_1 = E(f_{1}^{(1)} f_{2}^{(1)}) E(g_{1}s_1 g_{2}s_2)$ and $V_1 = E(g_{1}^{(1)} g_{2}^{(1)}) E(f_{1}s_1 f_{2}s_2)$. Similarly, for the reverse separating point $C_2$, we can obtain with $y_1 = \phi_{2,1} s_2, y_2 = \phi_{2,2} s_1$,
\begin{equation}
\alpha \gamma - \beta^2 = a_{12}^2 a_{21}^2 (U_2 - V_2)^2 \geq 0, \quad (38)
\end{equation}
where $U_2 = E(f_{1}^{(1)} f_{2}^{(1)}) E(g_{2}s_1 g_{1}s_2)$ and $V_2 = E(g_{1}^{(1)} g_{2}^{(1)}) E(f_{2}s_1 f_{1}s_2)$. Therefore, the algorithm is locally stable in the vicinity of the separating points unless $U_i = V_i$ such as $f(\cdot) = \eta g(\cdot)$ where $\eta$ is constant.

Some separating functions deserve comments. If $f(x) = x^3$ and $g(x) = x$ as many algorithms, $U_1 = 9\phi_{1,1}^2 \phi_{1,2}^2 E^2(s_1^2) E^2(s_2^2)$ and $V_1 = \phi_{1,1}^2 \phi_{1,2}^2 E(s_1^2) E(s_2^2)$. This leads to the condition $E(s_1^2) E(s_2^2) \neq 9 E^2(s_1^2) E^2(s_2^2)$ for stability. This relation holds for the reverse separating point $C_2$. Therefore, when we use a cubic and a linear function for the separating functions, stable separation can never be met if there are more than one Gaussian sources. Note that our condition is weaker than requirements of other researches [1,6,16,17], where the stable conditions are described in terms of inequality. If we employ the different nonlinear functions as the separating functions, Eq. (33) holds in the neighbourhood of $C_1$ and $C_2$. Therefore, the algorithm with the initial points near $C_1$ or $C_2$ converges to $C_1$ or $C_2$, respectively. We left the stability analysis of the higher-order cases for $n, m > 2$ for further works to the readers.

4. Computer simulation results

In the first computer simulation, we applied the algorithm to mixtures of a sub-Gaussian signal and a super-Gaussian signal with the mixing matrix
\begin{equation}
A = \begin{bmatrix}
1 & 0.31 \\
0.72 & 1
\end{bmatrix}. \quad (39)
\end{equation}
The sub-Gaussian signal $(\triangle s_1)$ is uniformly distributed on $[-1, 1]$ ($\gamma_{a,s_1} = E(s_1^2) - 3 E^2(s_1^2) = -0.8 < 0$) and the super-Gaussian signal $(\triangle s_2)$ is exponentially distributed with a parameter $1$ ($\gamma_{a,s_2} = E(s_2^2) - 3 E^2(s_2^2) = 6 > 0$). The separating function $f(x)$ and $g(x)$ are $\tanh(x)$, $\tanh(0.5x)$, respectively. The step-size $\mu$ used for the algorithm is $0.08$ and $C(0)$, the initial value of $C$ is $I$.

In Fig. 4, we present variations of the product of $C$ and $A$. For successful separation, $CA$ has a form of product of a diagonal matrix $D$ and a permutation matrix $P$ and from the result, we can know that $CA$ becomes that kind of form. This example is not solvable by the previous algorithms since they have the restriction on these signals, e.g., some techniques are applicable to either the sub- or super-Gaussian signals and some methods are suitable to only signals whose sum of kurtosis have negative

![Fig. 6. Variations of $W = CA$.](image-url)
value. Fig. 4 proves the validity of the proposed algorithm.

In the second and the third experiments, the algorithm is employed to the mixtures of two sampled signals and three sampled signals, respectively. The mixing matrix $A$ has 1 as diagonal elements and 0.5 as off-diagonal components and the other settings ($f(x)$, $g(x)$, $\mu$ and $C(0)$) are the same as in the first simulation.

Fig. 5 shows the successful separation result of a train horn sound and a bird chirp sound. Fig. 6 represents that the algorithm makes the separating matrix $C$ be trained such that $CA$ has the negligible off-diagonal elements in comparison with the diagonal components.

Figs. 7 and 8 depict the separation results of three arbitrary source signals and learning trajectories of elements of the matrix $CA$. 

---

Fig. 7. Separation result of three source signals.
5. Conclusions

In this paper, we proposed the new blind separation algorithm in order to apply it to the various source signals. It seems to be so difficult to solve this problem since the characteristics of the sources and channels are not known. The assumption we have made is only that the source signals are mutually independent and the channel modeling matrix is nonsingular. So far, many researchers have shown the remarkable results on this problem, but their results had limitations that they were not applicable to some sorts of signals.

For example, either sub- or super-Gaussian sources can be separable. However, we used the cumulants and the nonlinear functions so that the algorithm based on the newly proposed cost function had little restriction to various kinds of source signals. As shown in the computer simulations, the mixtures of the arbitrary signals could be separated successfully and especially the case which could not be solved by the previous researches have been resolved.

In the future research, the study on the bound of the step size and the steady-state error between $CA$ and $DP$ will be performed. Improvement of the learning speed are currently under investigation and we are expecting further research results on the cases of dependent source signals, nonlinear channels and additive noises.

References