A Blind Separation Algorithm
for Convolutive Mixture of Nonstationary Sources

KIYOTOSHI MATSUOKA and AKIO YAMAZAKI
Department of Brain Science and Engineering
Kyushu Institute of Technology
Hibikino, Wakamatsu-ku, Kitakyushu
JAPAN

Abstract: - A blind separation algorithm utilizing nonstationarity of sources is proposed. It is suitable particularly for separation of such strongly nonstationary signals as voices. The original version of the algorithm was proposed by one of the authors. The present version has made two improvements. First, it is extended to be able to deal with not only instantaneous mixture but also convolutive mixture of sources. Second, a new multiplier is introduced to suppress instability inherent in the original algorithm. Some experiments show a remarkably high-speed convergence.

Key-Words: - blind source separation, independent component analysis, nonstationary signal, convolutive mixture, voice

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.

A BSS algorithm can be classified from various viewpoints. In view of the level of complexity of the mixing process, it can be classified into that for instantaneous mixture or that for convolutive mixture. Another classification can be made depending on whether it is an on-line algorithm or an off-line one. Although on-line type algorithms are suitable in the case that the mixing process varies with time, they have a common drawback that convergence is usually very slow. In this paper we propose an on-line BSS algorithm for convolutive mixture of sources.

The algorithm we propose here utilizes nonstationarity of the source signals. So, it is suitable particularly for such strongly nonstationary signals as voices. Its original version was proposed by one of the authors [1], which is for instantaneous mixtures of sources. As opposed to usual BSS algorithms, the method uses only second-order statistics and the separating matrix is found utilizing nonstationarity in the source signals.

The algorithm proposed in this paper has made two improvements over the original one. First, although the original version is for instantaneous mixture, the present algorithm is extended to cope with convolutive mixture. Second, to suppress instability inherent in the original
algorithm, a particular term is introduced. The most remarkable characteristic of the proposed algorithm is very high-speed convergence.

This paper is organized as follows. In the next section the conventional algorithm utilizing nonstationarity of the sources will be explained. In section 3 we shall point out a critical problem of the algorithm and show an idea for solving it. In section 4 an extension of the modified algorithm to convolutive BSS will be shown. In section 5 we shall add some description about the actual implementation. Section 6 will show a couple of examples that demonstrate the effectiveness of the present method. Section 7 concludes the paper.

2 A BSS Algorithm for Instantaneous Mixture of Nonstationary Signals

Let \( s(t) = [s_1(t),\ldots,s_N(t)]^T \) and \( x(t) = [x_1(t),\ldots,x_N(t)]^T \) be source and observation vectors, respectively, whose entries can be complex-valued in general. For simplicity the number of the observations is assumed to be the same as that of the sources. For the mixing process we start with a simple case, i.e., instantaneous mixture of the sources:

\[
A s(t) = x(t). \tag{1}
\]

Since sources \( s_1(t),\ldots,s_N(t) \) are assumed to be statistically independent of each other and zero mean, their correlation matrix is diagonal; \( R(t) = E[s(t)s'(t)] \)

\[
= \text{diag}\{r_1(t),\ldots,r_N(t)\}, \quad \text{where } E[] \text{ stands for the statistical expectation of random variable }[]\text{ and diag}\{\ldots\} \text{ denotes a diagonal matrix with diagonal entries }\{\ldots\}. \text{ The most important point in this paper is that the source signals } s_i(t) \text{ are not stationary processes and their variances } r_i(t) \text{ are assumed to vary with time.}
\]

The demixing is performed by a separator with matrix \( W(t) \), i.e., the output

\[
y(t) = W(t)x(t). \tag{2}
\]

Matrix \( W(t) \) is to be updated in an on-line manner. Define the overall process from the source to the separator’s output as \( V(t) = W(t)A \), leading to

\[
y(t) = V(t)s(t). \tag{3}
\]

Although matrix \( V(t) \) is ‘invisible’, we shall mainly treat this matrix rather than the ‘visible’ matrix \( W(t) \) in the stability analysis of the algorithm.

If the mixing matrix \( A \) happens to be known beforehand, the source signals can be recovered by setting \( W(t) = A^{-1} \), of course. Essential difficulty in BSS is that \( A \) or \( A^{-1} \) must be estimated from the observed data \( x(t) \) only. As known very well, the task of BSS has some inevitable indeterminacy. That is, any matrix of the form \( DP^TA^{-1} \) can be considered a valid separator, where \( P \) is an arbitrary permutation matrix and \( D \) is an arbitrary diagonal matrix. Henceforth, however, we only consider the case of \( P = I \) because the indeterminacy due to permutation is not so essential.

The BSS algorithm proposed in [1] is

\[
W(t+1) = W(t) - \alpha \left( \hat{\rho}_1^{-1}(t),\ldots,\hat{\rho}_N^{-1}(t) \right) \quad \text{off-diag} \left( y(t)y'(t) \right) W(t), \tag{4}
\]

where \( \alpha \) is a small positive constant, \( \text{off-diag}(\cdot) \) sets every diagonal entry of matrix (\( \cdot \)) to be zero, and \( (\cdot)' \) denotes the conjugate transpose of vector or matrix \( (\cdot) \). When \( (\cdot) \) is a scalar, \( (\cdot)' \) will indicate just a complex conjugate. Variable \( \hat{\rho}_i(t) \) is an estimate of \( \rho_i(t) = E[|y_i(t)|^2] \); how to obtain the estimate will be described in section 6. Matrix \( V(t) = W(t)A \) obeys exactly the same dynamics as eqn (4):

\[
V(t+1) = V(t) - \alpha \left( \hat{\rho}_1^{-1}(t),\ldots,\hat{\rho}_N^{-1}(t) \right) \quad \text{off-diag} \left( y(t)y'(t) \right) V(t). \tag{5}
\]

What we want to show below is that \( V(t) \) converges to a
diagonal matrix along with iterative modification.

We assume that the estimation of \( \rho_i(t) = \hat{\rho}_i(t) - \rho_i(t) \) and \( \alpha \) is sufficiently small. Then, \( V(t) \) obeys the following equation approximately:

\[
V(t+1) = V(t) - \alpha \text{diag} \left\{ \rho_i^{-1}(t), ..., \rho_N^{-1}(t) \right\} \cdot \text{off-diag} \left\{ E[y(t)y'(t)] \right\} V(t).
\]

Since \( \rho_i(t) = v_i(t)R(t)v_i^*(t) \), where \( v_i(t) \) is the \( i \)-th row of \( V(t) \) and \( E[y(t)y'(t)] = V(t)R(t)V'(t) \), eqn (6) reads

\[
V(t+1) = V(t) - \alpha \cdot \text{diag} \left\{ (v_i(t)R(t)v_i^*(t))^{-1}, ..., (v_N(t)R(t)v_N^*(t))^{-1} \right\} \cdot \text{off-diag} \left\{ V(t)R(t)V'(t) \right\} V(t).
\]

Obviously any diagonal matrix \( D = \text{diag} \{ d_1, ..., d_N \} \) (\( d_i \neq 0 \)) is an equilibrium solution of the above equation, because \( \text{off-diag} \left\{ DR(t)D' \right\} \) becomes the zero matrix for every \( t \).

Now we investigate local stability of the equilibrium solution, \( V(t) \equiv D \). Consider a small perturbation in the vicinity of \( D \) as \( V(t) = D + \delta V(t) \). Then, we have

\[
\delta V(t+1) = \delta V(t) - \alpha \text{diag} \left\{ \overline{T}_{-1}(t), ..., \overline{T}_{-1}(t) \right\} \cdot \text{off-diag} \left\{ \delta V(t)R(t)D' + DR(t)\delta V'(t) \right\} D,
\]

where \( \overline{T}(t) = \{ d_i \}^T r(t) \), which is the variance of \( y_i(t) \) for the solution \( V(t) \equiv D \). Extracting the diagonal elements in eqn (8), we find

\[
\delta v_i(t+1) = \delta v_i(t).
\]

This corresponds to the fact that any diagonal matrix is an equilibrium solution of eqn (7), implying that the equilibrium is semi-stable in the direction of \( \delta v_i \). To eliminate this indeterminacy a particular normalization will be introduced in section 5 (and Appendix).

For the off-diagonal elements of \( \delta V(t) \), we have the following equation with respect to a pair of variables:

\[
\delta v_{ij}(t+1) = \begin{bmatrix}
1 - \alpha \frac{\overline{T}_{ij}(t)}{\overline{T}_{ij}(t)} & -\alpha \frac{d_j}{d_j} \\
-\alpha \frac{d_i}{d_j} & 1 - \alpha \frac{\overline{T}_{ij}(t)}{\overline{T}_{ij}(t)}
\end{bmatrix} \delta v_{ij}(t). \quad (10)
\]

where \( \delta v_{ij}(t) = [\delta v_{ij}(t) \delta v_{ji}(t)]^T \). Further we transform the variable as \( \delta \nu(t) = [\delta v_{ij}(t) \delta v_{ji}(t)]^T \) for the coefficient matrix to be symmetric:

\[
\delta \nu(t+1) = \begin{bmatrix}
1 - \alpha \frac{\overline{T}_{ij}(t)}{\overline{T}_{ij}(t)} & -\alpha \\
-\alpha & 1 - \alpha \frac{\overline{T}_{ij}(t)}{\overline{T}_{ij}(t)}
\end{bmatrix} \delta \nu(t). \quad (11)
\]

The eigenvalues and the corresponding eigenvectors of the coefficient matrix are

\[
\lambda_1(t) = 1, \quad \lambda_2(t) = 1 - \alpha \frac{\overline{T}_{ij}(t)}{\overline{T}_{ij}(t)} + \frac{\overline{T}_{ij}(t)}{\overline{T}_{ij}(t)}, \quad (12)
\]

\[
\phi_1(t) = \begin{bmatrix} \overline{T}_{ij}(t) \\ -\overline{T}_{ij}(t) \end{bmatrix}, \quad \phi_2(t) = \begin{bmatrix} \overline{T}_{ij}(t) \\ \overline{T}_{ij}(t) \end{bmatrix}. \quad (13)
\]

It should be noted that the two eigenvectors are orthogonal to each other. We can show that, if \( |\lambda_2(t)| < 1 \), then

\[
\| \delta \nu(t+1) \| \leq \| \delta \nu(t) \| \quad \text{holds; the equality holds only when } \delta \nu(t) \text{ lies on the line represented by vector } \phi_2(t).\]

Thus we find that if \( |\lambda_2(t)| < 1 \) or

\[
0 < \alpha < 2 \sqrt{\frac{\overline{T}_{ij}(t)}{\overline{T}_{ij}(t)} + \frac{\overline{T}_{ij}(t)}{\overline{T}_{ij}(t)}} \quad (14)
\]

holds for every time \( t \), then \( \| \delta \nu(t) \| \) never increases with time. Moreover, if the direction of \( \phi_1(t) \) continues to change with time (that is, \( r(t)/r_j(t) \) continues to fluctuate with time), then we can expect that \( \delta \nu(t) \) (or \( \delta \nu(t) \) ) converges to the zero vector. A more specific condition for the convergence is given in [1].
3 A Modification of the Algorithm

According to the last discussion, the value of the learning coefficient \( \alpha \) should be chosen so as to satisfy eqn (14) at every time \( t \). Actually, however, it is rarely possible to attain the condition. When, for example, source \( i \) is silent \((r_i(t) = 0)\) and source \( j \) is active \((r_j(t) \neq 0)\), the value of \( 2/\left[ \tau(t)/\tau(t) + \tau(t)/\tau(t) \right] \) becomes zero. So, any small value of \( \alpha \) cannot satisfy the condition.

In order to solve this problem we modify eqn (4) as

\[
\{11\}^{*} \hat{\theta}(1) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} - \alpha \theta (t) \{11\}^{*} \hat{\theta}(1) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} + \alpha \theta (t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} = \alpha \theta (t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix}
\]

or

\[
\{11\}^{*} \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} - \alpha \theta (t) \{11\}^{*} \hat{\theta}(1) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} + \alpha \theta (t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} = \alpha \theta (t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix}
\]

where \( \hat{\theta}(t) = \min_i \hat{\rho}_i(t) \). The key point here is that every element of \( \hat{\theta}(t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} \) is not greater than unity even if \( \hat{\rho}_i(t) \) takes any large value.

Again we assume that \( \alpha \) is small and the estimation of \( E[y(t) y(t)^*] \) is complete. Then, eqn (16) can be approximated by the following equation:

\[
\{11\}^{*} \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} - \alpha \theta (t) \{11\}^{*} \hat{\theta}(1) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} + \alpha \theta (t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} = \alpha \theta (t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix}
\]

where \( \theta(t) = \min_i \rho(t) \). As before, any diagonal matrix \( D = \{d_1, \ldots, d_N\} \) can be an equilibrium solution of the above equation.

Now we investigate local stability of the equilibrium. Consider a perturbation \( \delta V(t) \) in the vicinity of \( D \) as \( V(t) = D + \delta V(t) \). Then, we have

\[
\begin{align*}
\delta V(t+1) &= \delta V(t) - \alpha \hat{\theta}(t) \{11\}^{*} \hat{\theta}(1) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} + \alpha \theta (t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} \\
&= \delta V(t) - \alpha \hat{\theta}(t) \{11\}^{*} \hat{\theta}(1) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix} + \alpha \theta (t) \{11\}^{*} \begin{bmatrix} \theta \rho \omega \rho \omega \\ \end{bmatrix}
\end{align*}
\]

where \( \hat{\theta}(t) = \min_i \hat{\rho}_i(t) \) and \( \bar{\theta}(t) = \min_i \bar{\rho}_i(t) \). Extracting the diagonal term, we find \( \delta \bar{\rho}_i(t+1) = \delta \bar{\rho}_i(t) \) again. For the off-diagonal elements we have the following equation, corresponding to eqn (11):

\[
\begin{align*}
\delta \bar{\rho}_i(t+1) &= \begin{bmatrix} 1 - \alpha \bar{\theta}(t) & -\alpha \bar{\theta}(t) \\
-\alpha \bar{\theta}(t) & 1 - \alpha \bar{\theta}(t) \end{bmatrix} \begin{bmatrix} \bar{\rho}_i(t) \\
\bar{\rho}_i(t) \end{bmatrix} \delta \bar{\rho}_i(t).
\end{align*}
\]

The eigenvalues of the coefficient matrix of this system are found to be

\[
\lambda_1(t) = 1, \quad \lambda_2(t) = 1 - \alpha \bar{\theta}(t) \left( \frac{\bar{\rho}_i(t)}{\bar{\rho}_i(t)} + \frac{\bar{\rho}_i(t)}{\bar{\rho}_i(t)} \right).
\]

The corresponding eigenvectors are equivalent to those in eqn (13). A condition for \( \|\delta V_i(t+1)\|_2 \leq \|\delta V_i(t)\|_2 \) proves to be

\[
0 < \alpha < 2 \left( \frac{\bar{\theta}(t)}{\bar{\rho}_i(t)} + \frac{\bar{\rho}_i(t)}{\bar{\rho}_i(t)} \right)
\]

Since \( \bar{\theta}(t) \left( \frac{\bar{\rho}_i(t)}{\bar{\rho}_i(t)} + \frac{\bar{\rho}_i(t)}{\bar{\rho}_i(t)} \right) \leq \bar{\rho}_i(t) + \bar{\rho}_i(t) \), if

\[
\alpha < 1 \left( \max_i \bar{\rho}_i(t) \right),
\]

then \( \|\delta V_i(t)\|_2 \) never increase with time. Under the condition described in the last section we can expect that \( \|\delta V_i(t)\|_2 \) converges to zero. Important is that, as long as
$\mathcal{T}(t)$ is bounded to the upper, there always exists such $\alpha$ that satisfies eqn (21).

4 A BSS Algorithm for Convolutive Mixture of Nonstationary Signals

Now we consider the case we are mainly interested in, i.e., the case that the mixing process is a convolutive one:

$$x(t) = \sum_k A_k s(t-k) = A(z)s(t),$$  \hspace{1cm} (22)

where $A(z) = \sum_k A_k z^{-k}$. Due to statistical independence among $s_1(t),...,s_N(t)$, the auto-correlation matrix of $s(t)$ is diagonal; $R(t,k) \triangleq E[ss^\top] = \text{diag}\{r_1(t,k),...,r_N(t,k)\}$. Note that the source signals are nonstationary and the value of $R(t,k)$ varies with time $t$.

The demixing process in the convolutive case is given by

$$y(t) = \sum_k W_k(t)x(t-k) = W(t,z)x(t),$$  \hspace{1cm} (23)

where $W(t,z) = \sum_k W_k(t)z^{-k}$. Define the overall process from the sources to the separator’s output as $V_k(t) \triangleq \sum_i W_{k-i}(t)A_i$ or $V(t,z) \triangleq W(t,z)A(z)$ . Then, the above equation can be written as

$$y(t) = \sum_k V_k(t)s(t-k) = V(t,z)s(t).$$  \hspace{1cm} (24)

In the case of convolutive mixture, the indeterminacy is extended up to ‘filtering indeterminacy’. Namely, any separator given by the following form can be considered a valid separator.

$$W(z) = D(z)P^TA^{-1}(z),$$  \hspace{1cm} (25)

where $D(z)$ is a diagonal matrix whose diagonal elements are arbitrary analytic functions of $z$. For the same reason as before we simply consider the case of $P = I$.

In addition to filtering indeterminacy, it should be noted that the impulse response $\{W_k(t)\}$ might need to take a non-causal form in general, i.e., $W_k(t) \neq 0$ ($k<0$). It is because the mixing process is not necessarily a minimum-phase system. The problem of non-causality is solved by designing $W(t,z)$ such that the source signals may be reproduced with a time lag. However, for the sake of clarity of description, we omit the time lag in the algorithms below, except for section 5.

The BSS algorithm for convolutive mixture is

$$W_k(t+1) = W_k(t) - \alpha \hat{\theta}(t) \text{diag}\{\hat{\rho}_1^{-1}(t),...,\hat{\rho}_N^{-1}(t)\} \sum_i \text{off-diag}(y(t)y^\top(t-k+l))W_i(t),$$  \hspace{1cm} (26)

or

$$V_k(t+1) = V_k(t) - \alpha \hat{\theta}(t) \text{diag}\{\hat{\rho}_1^{-1}(t),...,\hat{\rho}_N^{-1}(t)\} \sum_i \text{off-diag}(y(t)y^\top(t-k+l))V_i(t).$$  \hspace{1cm} (27)

As before, the dynamics of the above equation can be approximated by the following equation under a certain condition:

$$V_k(t+1) = V_k(t) - \alpha \theta(t) \text{diag}\{\rho_1^{-1}(t),...,\rho_N^{-1}(t)\} \sum_i \text{off-diag}(E[y(t)y^\top(t-k+l)]V_i(t).$$  \hspace{1cm} (28)

Moreover, we assume that $R(t,k)$ changes slowly with time and hence $R(t+\tau,k) = R(t,k)$ holds for small $|\tau|$.

Then, eqn (28) can be rewritten as

$$V_k(t+1) = V_k(t) - \alpha \theta(t) \text{diag}\{\rho_1^{-1}(t),...,\rho_N^{-1}(t)\} \sum_i \text{off-diag}(V_i(t)R(t,k-l)V_{i,m}(t)\right)\}V_m(t).$$  \hspace{1cm} (29)

Introducing the Fourier transforms of $V_k(t)$ and $R(t,k)$ (from the $k$-domain to the $f$-domain, not from the $t$-domain) as

$$\tilde{V}(f) \triangleq \sum_k V_k(t)e^{-2\pi ikf} \quad (-1/2 \leq f \leq 1/2),$$

$$\tilde{R}(f) = \text{diag}\{\tilde{r}_1(f),...,\tilde{r}_N(f)\} \triangleq \sum_k R(t,k)e^{-2\pi ikf} ,$$

eqn (28) is transformed to
\[
\tilde{V}(t+1,f) = \tilde{V}(t,f) - \alpha \theta(t) \text{diag}\left\{ \rho_n^{-1}(t), \ldots, \rho_n^{-1}(t) \right\} \cdot \\
\text{off-diag}\left( \tilde{V}(t,f) \tilde{R}(t,f) \tilde{V}^*(t,f) \right) \tilde{V}(t,f)
\]  
(30)

Obviously any diagonal matrix \( \mathbf{D}(z) \) or \( \tilde{\mathbf{D}}(f) = \text{diag}\left\{ \tilde{d}_1(\cdot), \ldots, \tilde{d}_N(\cdot) \right\} = \mathbf{D}(e^{z\gamma}) \) can be an equilibrium solution of the above equation.

Next we investigate stability of the equilibrium. For eqn (30), consider a small perturbation in the vicinity of \( \tilde{f}, \tilde{D} \) as \( \tilde{V}(t,f) = \tilde{f}(t,f) + \delta \tilde{V}(t,f) \). Then, eqn (31) becomes
\[
\delta \tilde{V}(t+1,f) = \delta \tilde{V}(t,f) - \alpha \tilde{\sigma}(t) \text{diag}\left\{ \tilde{\rho}_1^{-1}(t), \ldots, \tilde{\rho}_N^{-1}(t) \right\} \cdot \\
\text{off-diag}\left( \delta \tilde{V}(t,f) \tilde{R}(t,f) \tilde{V}^*(t,f) \right) \tilde{V}(t,f)
\]  
(31)

Here, as before, \( \tilde{\sigma}(t) = E\left[ |y(t)|^2 \right] \) for the equilibrium solution of the separator and \( \tilde{\beta}(t) = \min_i \tilde{\sigma}_i(t) \). Extracting the diagonal elements, we have again
\[
\delta \tilde{v}_i(t+1,f) = \delta \tilde{v}_i(t,f) .
\]  
(32)

For the off-diagonal elements, we have the following equation:
\[
\delta \tilde{\psi}_y(t+1,f) = \delta \tilde{\psi}_y(t,f) - \alpha \cdot \\
\left[ \begin{array}{c}
\tilde{\beta}(t) \left[ \tilde{d}_i^*(f) \tilde{\check{r}}(f,t) \right] \\
\tilde{\beta}(t) \left[ \tilde{d}_j^*(f) \tilde{\check{r}}(f,t) \right]
\end{array} \right] - \alpha
\]  
(33)

where \( \tilde{\psi}_y(t,f) \triangleq \left[ \delta \tilde{v}_y(t,f) \delta \tilde{v}_y^*(t,f) \right]^T \). As opposed to the analysis in the last section we cannot say \( \delta \tilde{\psi}_y(t,f) \rightarrow 0 \ (t \rightarrow \infty) \) in general, but there are two particular cases where we can expect the convergence to a desired separator.

One case is the one where \( \tilde{r}_i(t,f) \) can be decomposed roughly as
\[
\tilde{r}_i(t,f) \approx \tilde{r}_i(f) r_i(t) .
\]  
(34)

This implies that the ‘shape’ of the power spectrum does not vary so much with time. Then the power spectrum of \( y_i(t) \) at time \( t \) for the desired separator is
\[
\left[ \tilde{d}_i(f) \right]^2 \tilde{r}_i(t,f) \approx \left[ \tilde{d}_i(f) \right]^2 \tilde{r}_i(f) r_i(t) ,
\]  
and hence \( \tilde{\sigma}_i(t) \) becomes
\[
\tilde{\sigma}_i(t) \approx \int_{-\pi/2}^{\pi/2} \left[ \tilde{d}_i(f) \right]^2 \tilde{r}_i(f) df \cdot \tilde{\sigma}(t) \triangleq c_i r_i(t) .
\]  
(35)

Then, eqn (33) leads to
\[
\delta \tilde{\psi}_y(t+1,f) \approx \left[ \begin{array}{c}
f_{11} f_{21} \\
f_{21} f_{22}
\end{array} \right] \delta \tilde{\psi}_y(t,f) ,
\]  
(36)

\[
f_{11} = 1 - \alpha \tilde{\sigma}(t) \left[ \left[ \tilde{d}_i(f) \right]^2 \tilde{r}_i(f) r_i(t) \right] c_i ,
\]

\[
f_{22} = 1 - \alpha \tilde{\sigma}(t) \left[ \left[ \tilde{d}_j(f) \right]^2 \tilde{r}_j(f) r_j(t) \right] c_j ,
\]

\[
f_{21} = -\alpha \tilde{\sigma}(t) \left[ \tilde{d}_j^*(f) \tilde{d}_j^*(f) \right] c_j / c_i .
\]

The eigenvalues of the coefficient matrix of eqn (36) are
\[
\lambda_1(t) = 1 ,
\]

\[
\lambda_2(t) = 1 - \alpha \tilde{\sigma}(t) \left[ \left[ \tilde{d}_i(f) \right]^2 \tilde{r}_i(f) r_i(t) + \left[ \tilde{d}_j(f) \right]^2 \tilde{r}_j(f) r_j(t) \right] / \left[ \left[ \tilde{d}_j(f) \right]^2 \tilde{r}_j(f) r_j(t) \right] .
\]  
In a similar way to before we can show a condition for parameter \( \alpha \) to guarantee the stable convergence to the desired separator:
\[
\alpha < \frac{1}{\lambda_2(t)} \left( \max_{i,j} \left[ \tilde{d}_i(f) \right]^2 \tilde{r}_i(f) r_i(t) \right) .
\]  
(37)

Another interesting case is the one where source signals are sparse in the sense that at almost any time only one source is active and other sources inactive. Consider
that in a time interval, source \( j \) is active and source \( i \) is inactive (silent). Then, eqn (33) reduces to

\[
\delta \hat{v}_j(t+1, f) = \left[ 1 - \alpha \delta_j(t) \frac{\hat{\delta}_j(f)}{\hat{p}_j(t)} \right] \delta \hat{v}_j(t, f) + \frac{\hat{\delta}_j(f)}{\hat{p}_j(t)} \delta \hat{v}_i(t, f),
\]

implying that \( \delta v_j(t, f) \to 0 \) in this interval. If this kind of sparseness is present for almost every instant of time, we can expect that all \( \delta v_j(t, f) \) will converge to zero.

5 Actual Implementation

Here we add some description about the actual implementation. First of all, since the mixing process is not necessarily a minimum-phase system, the separator must be non-causal in general. The problem of non-causality can be solved by designing \( \hat{W}(t, z) \) such that the source signals are reproduced with a time lag;

\[
y(t-L) = \sum_{k=0}^{L} W_i(t) s(t-L-k).
\]

In connection with this, a larger time lag must be introduced for the updating rule of the separator.

Direct calculation of eqn (26) is not efficient. Instead we employ

\[
\mathbf{u}(t-L_0) = \sum_{i=L_0}^{L_0+L} W^*_i(t) y(t-L_0+r) \quad (L_0 = L_0 + L_2),
\]

\[
\mathbf{U}(t-L_0) = \sum_{i=L_0}^{L_0+L} \text{diag} \mathbf{y}^*(t-L_0+r) \cdot \mathbf{W}_i(t),
\]

\[
\Delta \mathbf{W}_i(t) = -\alpha \cdot \hat{\delta}(t-L) \cdot \text{diag} \left\{ \hat{\rho}_1(t-L_0), \ldots, \hat{\rho}_n(t-L_0) \right\} \cdot \mathbf{y}(t-L_0) \cdot \mathbf{u}^*(t-L_0-k) - \text{diag}(y(t-L_0) \cdot U(t-L_0-k))
\]

\[
(L_0 = 2L_0 + L_2)
\]

where \( \text{diag} \mathbf{y} \) denotes a diagonal matrix whose diagonal entries are given by the entries of vector \( \mathbf{y} \).

The estimate of \( E[|y_i(t)|^2] \) is given by

\[
\hat{\rho}_i(t-L_0) = \frac{1}{2L_0 + 1} \sum_{l=0}^{L_0} |y_i(t-L_0+l)|
\]

To enhance the stability, the learning coefficient \( \alpha_k \) is given different values for different \( k \) as

\[
\alpha_k = \begin{cases} 
\alpha(1+k/L_0) & \text{for } -L_0 \leq k < 0 \\
\alpha(1-k/L_2) & \text{for } 0 \leq k \leq L_2
\end{cases}
\]

This idea was introduced in [4].

The proposed algorithm (26) is a nonholonomic type algorithm, i.e., the scaling or filtering indeterminacy exists and the result depends on the initial value of \( \hat{W}(t, z) \). To eliminate the indeterminacy we add a following modification to the separator at each iteration:

\[
\hat{W}(t, z) \leftarrow (1-\varepsilon)\hat{W}(t, z) + \varepsilon \left\{ \left(1 - \frac{1}{N} \mathbf{e} \mathbf{e}^T \right) \hat{W}(t, z) + \frac{1}{N} \mathbf{e} \mathbf{f}^T \right\}
\]

where \( \mathbf{e} \triangleq [1, 1, \ldots, 1]^T \) and \( \mathbf{f} \triangleq [1, 0, \ldots, 0]^T \). The meaning of this particular normalization is explained in Appendix.

6 Examples

To demonstrate the effectiveness of the proposed method, we first show a simulation. In this simulation, although the sources are real voices (speeches of the same woman), the mixing was artificial as

\[
A(z) = \begin{bmatrix} 1 & z^{-1} \\
                              & 1 \end{bmatrix}.
\]

Since \( z^{-1} \) means a time delay of 1 [ms], \( x_1(t) \) and \( x_2(t) \) are almost the same signals. Fig.1 (a) and (b) show the source signals and the observations, respectively. Fig.1 (c) shows the separation result. Although the task seems to be very difficult, the separation was successfully attained in some seconds after the start. It should be noted that a conventional method using non-gaussianity [3] resulted in a total failure.

Also we tested the algorithm to the data that was obtained in a ‘real’ situation. The experimental setup is shown in Fig.2. Also in this case the desired separator wa
obtained in a some seconds, as shown in Fig.3.

7 Conclusion
We have proposed a BSS algorithm for convolutive mixture of nonstationary signals. The most attractive feature of the present algorithm is a high-speed convergence. However, the algorithm still reveals an instability when the length of the separating filter was long. In the separation of sound signals the filter length must usually be made very long if one wants a high accuracy of separation. How to maintain the stability in such a case is an important future subject.

References:

Appendix: A Particular Normalization of W(z)
Inherently BSS has two kinds of indeterminacy. One is indeterminacy in the numbering or labeling of the sources and the other is that in the scaling or normalization. The latter is more essential particularly for convolutive mixture of sources. Any linear transform of a source signal can also be considered a source signal, implying that there exist an infinite number of valid separators that extract the source signals. An idea for normalizing the separator is
\[ \text{diag}(W^{-1}(z)) = 1, \]  
which is named Minimal Distortion Principle (MDP) for a reason [5]. This constraint leads to
\[ W(z) = \text{diag}\{a_{1}(z) \cdots a_{N}(z)\}A^{-1}(z) \]
or \[ y_i(t) = a_i(z)s_i(t). \]

The constraint used in the simulation is different
\[ e^T W(z) = f^T, \]  
where \( e \triangleq [1,1,\ldots,1]^T \) and \( f \triangleq [1,0,\ldots,0]^T \). The valid separator that satisfies this constraint is
\[ W(z) = \text{diag}\{a_{1}(z),\ldots,a_{N}(z)\}A^{-1}(z), \]  
and \[ y_i(t) = a_i(z)s_i(t). \] It can be proved as follows. A valid separator takes the form \( D(z)A^{-1}(z) \), where \( D(z) \) is a diagonal matrix. Therefore, the constraint (A2) gives
\[ e^T D(z)A^{-1}(z) = f^T \quad \text{or} \quad e^T D(z) = f^T A(z). \] From this we find that \( d_i(z) = a_i(z) \), leading to eqn (A3).

Let \( \Pi \) be the space of \( W(z) \) satisfying
\[ e^T W(z) = f^T. \] Given \( W(t,z) \), the nearest point on \( \Pi \) from \( W(t,z) \) is
\[ \left( I - \frac{1}{N} ee^T \right) W(t,z) + \frac{1}{N} ee^T. \] The operation in eqn (44) is to force \( W(t,z) \) to approach this point.
Fig. 1 The result of the simulation

(a) Sources

(b) Observations

(c) Output of the separator

Fig. 2 Experimental setup

(a) Sources

(b) Observations

(c) Output of the separator

Fig. 3 The result of the real experiment