A Neural Learning Algorithm of Blind Separation of Noisy Mixed Images Based on Independent Component Analysis

Hongyan Li  
College of Information Engineering of Taiyuan University of Technology, Taiyuan, China  
E-mail: tylihy@163.com

Xueying Zhang  
College of Information Engineering of Taiyuan University of Technology, Taiyuan, China  
E-mail: tyzhangxy@163.com

Abstract—Blind source separation problem has recently received a great deal of attention in signal processing and unsupervised neural learning. In the current approaches, the additive noise is negligible so that it can be omitted from the consideration. To be applicable in realistic scenarios, blind source separation approaches should deal evenly with the presence of noise. In this contribution, we proposed approaches to independent component analysis when the measured signals are contaminated by additive noise. A noisy multiple channels neural learning algorithm of blind separation is proposed based on independent component analysis. The data have no noise are used to whiten the noisy data, and the windage wipe off technique is used to correct the infection of noise, a neural network model having denoise capability is adopted to recover some original signals from their noisy mixtures observed by the same number of sensors. And a relaxation factor is introduced into the iteration algorithm, thus the new algorithm can implement convergence. Computer simulations and experiment results prove the feasibility and validity of the neural network modeling and control method based on independent component analysis, which can renew the original images effectively.

Index Terms—Independent component analysis; Neural network; Blind source separation; Noise cancellation; Relaxation factor

I. INTRODUCTION

Independent component analysis (ICA) is a statistical technique in signal processing and machine learning that aims at finding linear projections of the data that maximize their mutual independence. The process of ICA is extracting unknown independent source signals from sensor that are unknown combinations of the source signals. The work has been one of the most exciting topics in the fields of neural computation, advanced statistics, signal processing, and communication engineering which find independent components from observing multidimensional data based on higher order statistics[1–4].

Blind source separation is a technique that extracts the original signals from their mixtures observed by sensors. In order to solve the problem of blind signal separation, many techniques have been proposed [5–9]. Among them, the ICA methods are based on the assumption of mutual independence of the sources. Traditional ICA has some assumption such as assuming that there is no noise or noise is minute that can be ignored, so it can not solve the problem of the noisy blind signal separation. Most of ICA methods are developed in the case of noiseless data. Some fast and efficient algorithms have been proposed such as FASTICA [10–14]. However, all these algorithms perform poorly when noise affects the data. To overcome the ICA limitation, some work has been done to partially overcome this limitation [15–19]. An earlier research of this paper has been done in [20], which combined ICA and neural network to separate the noisy mixed speech. Moreover, in this paper, considering that the negative gradient direction is the fastest decline of the function value, thus, we select the gradient value as the relaxation factor, and the relaxation factor is introduced into the iteration algorithm to improve the convergence, and a improved noisy multiple channels neural learning algorithm of blind separation combining with windage wipe off technique is proposed to solve the problem of noisy multiple blind signal separation. As well as the proposed algorithm is applied in the noisy multiple channels blind images sources separation. Finally, more conclusions and analysis are presented in the simulations. Separation results obtained from test demonstrate the feasibility of our approach.

This paper is organized as follows: a recurrent neural network structure fundamental of ICA is exposed in section 2. In section 3, a noisy multiple channels blind signal separation algorithm based on neural network is introduced which performs noisy blind separation by combining a noisy multiple channels blind separation neural learning algorithm with windage wipe off technique, besides, a relaxation factor is introduced into the iteration algorithm. Section 4 is the simulation results for multiple noisy blind separation using the proposed algorithm, as well as the comparisons with other noisy...
blind separation algorithms. Finally, conclusions and analysis are presented in section 5.

II ICA RECURRENT NEURAL NETWORK STRUCTURE

ICA is one method for performing blind signal separation that aims to recover unknown sources from a set of their observations, in which they are mixed in an unknown manner. Suppose that signals $s_j(t)$ ($j=1,2,\ldots,n$) are generated by $n$ statistically independent sources, and their linear mixtures $x_i(t)$ ($i=1,2,\ldots,n$) are observed by $n$ sensors:

$$x_i(t) = \sum_{j=1}^{n} a_{ij}(t)s_j(t) \quad (1)$$

where $a_{ij}$ are constants independent of time $t$.

The aim of ICA is to find the matrix $W$, The ICA recurrent neural network structure can be depicted by Fig.1 [21].

Improving this model with a whole connected recurrent neural network, each neuron is connected with the whole neurones, the model can be expressed as

$$y_i(t) = x_i(t) - \sum_{j=1}^{n} w_{ij}(t)y_j(t) \quad (2)$$

With the optimal weight $w_{ij}$ the output signals can be independent mutually. The independent vectors $f(x)$ and $g(y)$ have the formula

$$E[f(x)g(x)] = E[f(x)]E[g(x)] \quad (3)$$

$$R_s = E[f(y)g'(y)] - E[f(y)]E[g'(y)]$$

$$= \begin{bmatrix}
E[f(y)g(y)] - E[f(y)]E[g(y)] \\
\vdots \\
0 \\
\vdots \\
E[f(y)g(y)] - E[f(y)]E[g(y)]
\end{bmatrix}$$

where the covariance $E[f(y)g(y)] - E[f(y)]E[g(y)]$ is zero when $i$ is not equal to $j$, the variance $E[f(y)g(y)] - E[f(y)]E[g(y)]$ is not zero, and $f(y)$ and $g(y)$ are different nonlinear activation function.

According to the statistic independence and the normalize condition $E[f(y)g(y)] = \lambda_i$, the real-time learning algorithm can be deduced

$$\frac{dw_{ij}(t)}{dt} = \mu(t) [\lambda_i - f(y_i(t))g(y_j(t)) + \epsilon_{ij} - \epsilon_{ji}] \quad (5)$$

It can be expressed as matrix form

$$\frac{dW(t)}{dt} = \mu(t) [A - f(y(t))g^{T}(y(t))] \quad (6)$$

where $A = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$

III A NOISY MULTIPLE CHANNELS BLIND SIGNAL SEPARATION ALGORITHM BASED ON NEURAL NETWORK

For the noisy ICA model $x=As+n$, the hypothesis are showed as follows

1) The noise and the independent component are independent mutually
2) The covariance of noise is foregone
3) $n$ is Gaussian noise

In the presence of noise, ICA becomes quite difficulty, the biggest difficult problem is how to estimate the independent component without noise. Estimating the noise component becomes more difficult because the noise model is not reversible. We reduce the effects of noise from three aspects, first of all, in pre-processing, noise effects are considered in the whitening of the data, the data having no noise are used to whiten the noisy data, in addition, the windage wipe off technique is used to correct the infection of noise, at last, a ICA model based on clustering algorithm of neural network having denoise capability as well as a relaxation factor is introduced into the iteration algorithm to improve the convergence and realize the noisy multiple channels blind source separation.

A The Whitening of Data

The ICA problem is greatly simplified if the observed mixture vectors are first whitened or sphered. A zero-mean random vector $x=(x_1, \ldots, x_n)^T$ is said to be
white if its elements $x_i$ are uncorrelated and have unit variances. If the centralized signal is $x$, the whitened signal is $\tilde{X}$, in terms of the covariance matrix, that means

$$E[\tilde{X}\tilde{X}^T] = I$$

(7)

where $I$ is the unit matrix.

It is easy to see that the whitening signals have unit variance and uncorrelated, independently of the rotation angle. The whitening has a solution using eigenvalue decomposition for the covariance matrix of signals. Let $V$ be the orthogonal matrix consisted of covariance matrix $E[xx^T]$. Let $D = \text{diag}(d_1, \cdots, d_r)$ be the diagonal matrix of the eigenvalue of $E[xx^T]$, the whitening process can be described as

$$\tilde{x}(t) = VD^{-1/2}V^Tx(t)$$

(8)

Making that $U = VD^{-1/2}V^T$, where $U$ is the whitening matrix which is usually computed after singular or eigenvalue decomposition of the covariance matrix of $x(t)$, that is

$$\tilde{x}(t) = UD(t)$$

(9)

where $\tilde{x}(t)$ is the whitened mixing signal.

Let $\tilde{X}$ be an arbitrary non-Gaussian random variable, $\eta$ is an independent Gaussian variable with noise variance $\sigma^2$, as long as the choice of $G$ is suitable, the relationship between $E[G(\eta)]$ and $E[G(\eta+n)]$ can be simply measured, this measurement method is also applicable to noisy ICA.

Selecting $G$ as a density function of a zero mean Gaussian random vector, the density function of vector $x$ with variance $c^2$ can then be expressed

$$\varphi_c(x) = \frac{1}{c} \frac{1}{\sqrt{2\pi c^2}} \exp\left(-\frac{x^2}{2c^2}\right)$$

(15)

Assuming that $\varphi^{(k)}_c$ is the $k$th order derivatives, $\varphi^{(-k)}_c$ is the $k$th order integral ($k$ is a positive integer), and $\varphi^{(k)}_c(x) = \int_0^x \varphi^{(-k)}_c(\xi) d\xi$, a theorem can be expressed

Theorem[21]: Supposing that $z$ is a non-Gaussian random vector, $n$ is an independent noise with variance $c^2$. Defining Gaussian function $\varphi$ with the formula (10), then for any constant $c>\sigma^2$, there is

$$E[\varphi_c(z)] = E[\varphi_c(z+n)]$$

(16)

where $d = \sqrt{c^2-\sigma^2}$, substituting $\varphi^{(k)}_c$ for $\varphi$, the formula still valid.

The theorem revealed that the noisy independent components can be estimated from noisy observation signals by maximization the contrast function.

C The Noisy ICA Algorithm Based on Neural Network

Using the improved ICA recurrent neural network algorithm, the algorithm can be expressed

$$\frac{dw(t)}{dt} = \mu(t)[A - f[y(t)]g^T[y(t)]]$$

(17)

The discrete iterative form is

$$w(n+1) = w(n) - \mu(n)[A - f[y(n)]g^T[y(n)]]$$

(18)

The algorithm has the advantage such as uncomplicated principle and handy realization. But in actual application, it has the limitations of stagnation and poor convergence, and can result in local minimum, thus leading to the bottlenecks of its wide application. Making further refinements to the algorithm, imposing momentum term, adopting relaxation factor, the new noisy blind signal separation based on neural network can be given by

$$w(n+1) = w(n) + \beta[w(n) - w(n-1)]$$

- $\mu(n)[1 + \Sigma][A - f[y(n)]g^T[y(n)]]$

(19)

where $0 \leq \beta \leq 1$ is a momentum factor, the affection of the weights variation of the last two times are passed to the present weights variation through momentum factor, so as to send weights variation along the mean direction of error nether camber surface, and reduce the sensitivity of the local details in error camber surface for
neural network, as well as restrain the defect of fall into local minimum. \( \mu \) is a relaxation factor, we select the biggish learning rate in learning elementary stage for faster speed, and reduce the learning rate in stage of closing to convergence for avoiding oscillation and non-convergence. By multivariable calculus, the negative gradient direction is the fastest decline of the function value, thus, we select the gradient value as the variable learning rate, that is

\[
\mu = \left[ \begin{array}{cc}
\frac{\partial E[\{y(n)g[y(n)]\}]}{\partial w_1} & 0 \\
0 & \frac{\partial E[\{y(n)g[y(n)]\}]}{\partial w_2}
\end{array} \right] \tag{20}
\]

According to the improved algorithm, the noisy ICA based on neural network principle block diagram can be shown in Fig.2.

Fig.2 The noisy ICA based on neural network

According to the iterative formula of the algorithm, the steps of the algorithm can be shown as follows:

1) Making that \( i = 1 \), where \( i \) is the number of the sources;
2) Making that whitening signal

\[
\tilde{x}(t) = Ux(t) = A^T V^T x(t)
\]

where \( U \) is the whitening matrix, \( V = [v_1, \ldots, v_n]^T \), and \( V \) is the norm feature vectors matrix of the covariance matrix, \( A = [A_1, \ldots, A_n] \) is a diagonal matrix, and the diagonal elements are the eigenvalues of the covariance matrix;
3) Choosing an initial matrix \( w(0) \) and making that \( k = 1 \);
4) Calculating the relaxation factor \( \mu \) by formula (20)
5) Making that

\[
w(n+1) = w(n) + \beta \left[ w(n) - w(n-1) \right] - \mu(n) \left[ 1 + \sum |A - f[y(n)]g[y(n)]| \right]
\]

6) Making that \( w(k) = w(k) \)
7) If \( \left| w_k(k) - w_{k-1} \right| \) is equal or close to 1, the iteration finished, otherwise make \( k = k + 1 \) and go back to step 4)
8) Making that \( i = i + 1 \), when \( i < \) number of original signals, go back to step 3).

IV SIMULATIONS

We select two original images ic and lenna as the sources which can be shown in Fig.3. The mixed images depicted by Fig.4 are mixed by matrix \( A \) with original images, where

\[
A = \begin{bmatrix} 0.2011 & 0.8109 \\ 0.2314 & 0.3541 \end{bmatrix}.
\]

adding white Gaussian noise to mixed images for three times, the variances of noise are 10, 50 and 100, the noisy mixed images can be depicted by Fig.5.
The mixed images with noise once

The mixed images with much noise

The mixed images with much more noise

Fig. 5 The noisy mixed images

The separated images by NOISYFASTICA can be depicted by Fig.6.

Using the gradient value as the relaxation factor $\mu$ and selecting non-linear function $f(y) = y \cdot g(y) = \tanh(y)$, the separated images by the proposed algorithm can be depicted by Fig.7.

The separated images by NOISYFASTICA for the first time

The separated images by NOISYFASTICA for the second time

The separated images by NOISYFASTICA for the third time

Fig. 6 The separated images by NOISYFASTICA

The separated images by the proposed algorithm for the first time
We examine the signal-noise ratio of mixing images, noisy mixing images, separation images by NOISYFASTICA and separation images by the proposed algorithm which are shown by Table 1.

<table>
<thead>
<tr>
<th>images</th>
<th>The images SNR with noise once (dB)</th>
<th>The images SNR with much noise (dB)</th>
<th>The images SNR with much more noise (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mixed images</td>
<td>ic</td>
<td>-2.3159</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lenna</td>
<td>3.2346</td>
<td></td>
</tr>
<tr>
<td>noisy mixing images</td>
<td>ic</td>
<td>-4.3298</td>
<td>-8.5241</td>
</tr>
<tr>
<td></td>
<td>lenna</td>
<td>-0.3245</td>
<td>-4.7512</td>
</tr>
<tr>
<td>Separation images by</td>
<td>ic</td>
<td>8.9452</td>
<td>8.7584</td>
</tr>
<tr>
<td>NOISYFASTICA</td>
<td>lenna</td>
<td>11.4687</td>
<td>11.1425</td>
</tr>
<tr>
<td>separation</td>
<td>ic</td>
<td>9.3652</td>
<td>9.1465</td>
</tr>
<tr>
<td></td>
<td>lenna</td>
<td>12.3512</td>
<td>11.3578</td>
</tr>
</tbody>
</table>

The performance index comparison of NOISYFASTICA and the proposed algorithm can be depicted by Table 2.

<table>
<thead>
<tr>
<th>algorithm</th>
<th>simulation</th>
<th>iteration time (s)</th>
<th>PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOISYFASTICA</td>
<td>first time</td>
<td>2.6432</td>
<td>0.2987</td>
</tr>
<tr>
<td></td>
<td>second time</td>
<td>2.5398</td>
<td>0.3786</td>
</tr>
<tr>
<td></td>
<td>third time</td>
<td>2.9568</td>
<td>0.5119</td>
</tr>
<tr>
<td>The proposed algorithm</td>
<td>first time</td>
<td>3.7461</td>
<td>0.2376</td>
</tr>
<tr>
<td></td>
<td>second time</td>
<td>3.3475</td>
<td>0.3545</td>
</tr>
<tr>
<td></td>
<td>third time</td>
<td>3.4854</td>
<td>0.4435</td>
</tr>
</tbody>
</table>

where performance index

\[
PI = \frac{1}{2N} \left( \sum_{i=1}^{N} \left( \frac{\|G_{i1}\|}{\max_{j} |g_{ij}|} - 1 \right) + \sum_{j=1}^{N} \left( \frac{\|G_{j1}\|}{\max_{i} |g_{ij}|} - 1 \right) \right)
\]

G is the transmission matrix and it is \(N \times N\) matrix, and \(G=WA\), \(g_{ij}\) is the element of \(G\), \(\max_{j} |g_{ij}|\) is the maximum absolute value of line \(i\), \(\max_{i} |g_{ij}|\) is the maximum absolute value of row \(j\).

In order to further verify the stability of the proposed algorithm, Adding white Gaussian noise to mixed images for 10 times, the variances of noise are 10, 20, 30, 40, 50, 60, 70, 80, 90, we examine the SNR of mixed images, noisy mixed images, the separation images by NOISYFASTICA and the separation images by the proposed algorithm, the SNR comparison of image ic can be depicted by Fig.7 and the SNR comparison of image lenna can be depicted by Fig.8. where the x axis N0,N1,N2,N3,N4,N5,N6,N7,N8,N9 are the number of the simulations.
Furtherly, the performance index comparison of NOISYFASTICA and the proposed algorithm for the 10 times simulation can be depicted by Fig. 9. The iteration time comparison of NOISYFASTICA and the proposed algorithm can be depicted by Fig. 10, where the x axis N0,N1,N2,N3,N4,N5,N6,N7,N8,N9 are the number of the simulations.

From the simulation results and the data in the table can be seen, with the increase of the noise strength, the antinoise ability of NOISYFASTICA and the proposed algorithm are also very strong even in the case of the images covered by noise, both of them can improve the SNR of the noisy images effectively. Comparing with NOISYFASTICA algorithm, the proposed algorithm has the better performance and separation effect, but the iterative for long time.

V CONCLUSION

1) A method for performing multiple channels blind signal separation in the presence of additive noise is described. The method is proposed of combining neural network and independent component analysis to separate noisy mixed images. The windage wipe off technique was used to correct the infection of noise, a neural network model having denoise capability was adopted to realize the multiple channels blind source separation method for mixing images corrupted with white noise, and a relaxation factor are introduced into the iteration algorithm to improve the convergence.

2) The proposed algorithm reduces the noise influences, overcomes the shortcoming of the traditional ICA cannot be used in the presence of noise model, and makes the online ICA model more easily applied.

3) The simulation and experiment results prove the feasibility and validity of the modeling and control method based on neural network and independent component analysis, which can renew the original image effectively. The results indicate that it has better performances than the classical of NOISYFASTICA and it satisfies the noisy multiple channels blind signal separation problem better. Furthermore, with this algorithm, ICA can be used to solve the problem of more realistic blind signal separation.

ACKNOWLEDGEMENTS

This work was supported by Doctoral Fund of Ministry of Education of China (No. 2011081047) and Shanx National Science Foundation (No. 2013011016-1) and Shanxi Young Science Foundation (No.2013021016-1).

REFERENCES


Hongyan Li was born in Shanxi, China, in 1973. She is an associate professor at Taiyuan University of Technology. Since 2009, she has been a Ph.D. degree candidate in circuit and system from Taiyuan University of Technology. She is currently pursuing the postdoctoral at Taiyuan University of Technology. Her research interests include blind signal processing and pattern identification.

Xueying Zhang was born in Shanxi, China, in 1964. She received the Ph.D. degree in underwater acoustic engineering from Harbin Engineering University. Since 1999, she has been a professor at Taiyuan University of Technology. Her research interests include speech signal processing, multimedia communication, embedded system and application and the internet of things technology, etc.