Journal of Computational and Applied Mathematics 229 (2009) 240–247

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Blind source separation with nonlinear autocorrelation and non-Gaussianity ☆

Zhenwei Shi *, Zhiguo Jiang, Fugen Zhou, Jihao Yin

Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing 100083, PR China

ARTICLE INFO

Article history:
Received 19 March 2008
Received in revised form 14 July 2008

Keywords:
Blind source separation (BSS)
Independent component analysis (ICA)
Nonlinear autocorrelation
Non-Gaussianity

ABSTRACT

Blind source separation (BSS) is a problem that is often encountered in many applications, such as biomedical signal processing and analysis, speech and image processing, wireless telecommunication systems, data mining, sonar, radar enhancement, etc. One often solves the BSS problem by using the statistical properties of original sources, e.g., non-Gaussianity or time-structure information. Nevertheless, real-life mixtures are likely to contain both non-Gaussianity and time-structure information sources, rendering the algorithms using only one statistical property fail. In this paper, we address the BSS problem when source signals have non-Gaussianity and temporal structure with nonlinear autocorrelation. Based on the two statistical characteristics of sources, we develop an objective function. Maximizing the objective function, we propose a gradient ascent source separation algorithm. Furthermore, We give some mathematical properties for the algorithm. Computer simulations for sources with square temporal autocorrelation and non-Gaussianity illustrate the efficiency of the proposed approach.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Recently, due to its wide application in the areas of biomedical signal processing and analysis, speech and image processing, wireless telecommunication systems, data mining, sonar, radar enhancement, and so on, blind source separation (BSS) [6,12] has been studied extensively. It is a technique to recover the original sources from their mixtures without the knowledge of the mixing process, only using some statistical properties of original sources.

The blind source separation problem has been studied by researchers in applied mathematics, neural networks and statistical signal processing. Several methods for BSS using the statistical properties of original sources have been proposed, such as non-Gaussianity (or equivalently, independent component analysis, ICA) [1,3,5–7,9,12–14,19,27], or time-structure information, such as linear predictability or smoothness [2,6], linear autocorrelation [4,16,26], coding complexity [10,20,21,23], temporal predictability [18], nonstationarity [11,15,17], energy predictability [22], nonlinear innovation [24], nonlinear autocorrelation [25], etc. One often solves the BSS problem by using only one statistical property of original sources, e.g., only non-Gaussianity or only time-structure information. Nevertheless, real-life mixtures are likely to contain both non-Gaussianity and time-structure information sources, rendering the algorithms using only one statistical property fail. In this paper, we address the BSS problem when source signals have non-Gaussianity and temporal structure with nonlinear autocorrelation, which extends the previous situations. Based on the two statistical characteristics of sources, we develop a gradient ascent source separation algorithm. Computer simulations for sources with square temporal autocorrelation and non-Gaussianity illustrate the efficiency of the proposed approach.

☆ The work was supported by the National Science Foundation of China (60605002, 60776093 and 10571018).
* Corresponding author. Tel.: +86 10 823 16 502; fax: +86 10 627 86 911.
E-mail address: shizhenwei@mail.tsinghua.edu.cn (Z. Shi).

0377-0427/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2008.10.031

The structure of the paper is as follows. The objective function based on nonlinear autocorrelation and non-Gaussianity of the desired source signals, and a gradient ascent algorithm for optimizing the objective function are proposed in Section 2. Furthermore, we analyze the stability conditions of the algorithm in Section 3. In Section 4, some experiments are presented. Conclusions are drawn in Section 5.

2. Proposed algorithm

2.1. Objective function

The mixed sensor signals $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^{\mathrm{T}}$ is given by the matrix equation:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{1}$$

where $\mathbf{A}$ is an $n \times n$ unknown mixing matrix, $\mathbf{s}(t) = (s_1(t), \ldots, s_n(t))^{\mathrm{T}}$ is a vector of unknown zero-mean and unit-variance original sources. We assume that original sources are mutually independent with non-Gaussian and nonlinear autocorrelation.

The basic problem of BSS is then to estimate both the mixing matrix $\mathbf{A}$ and the source signals $s_i(t)$ using observations of the mixtures $x_i(t)$ $(i = 1, \ldots, n)$ and some statistical properties of original sources, to find an $n \times n$ separating matrix $\mathbf{W} = (\mathbf{w}_1, \ldots, \mathbf{w}_n)^{\mathrm{T}}$ such that the unmixed signals $\mathbf{y}(t) = (y_1(t), \ldots, y_N(t))^{\mathrm{T}}$,

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) \tag{2}$$

are the estimated source signals [6,12]. The sources are recovered up to scaling and permutation.

Assume that we want to estimate a desired source signal, for this purpose we design a single processing unit described as:

$$y_i(t) = \mathbf{w}_i^T \mathbf{x}(t), \tag{3}$$

$$y_i(t-\tau) = \mathbf{w}_i^T \mathbf{x}(t-\tau), \tag{4}$$

where $\mathbf{w}_i = (w_{i1}, \ldots, w_{in})^{\mathrm{T}}$ is the weight vector, and τ is some lag constant, often equal to one.

We first whiten the measured sensor signals $\mathbf{x}$, which is a useful preprocessing strategy in BSS [6,12]. For example, an $n \times n$ whitening matrix $\mathbf{V}$ is used, by $\tilde{\mathbf{x}}(t) = \mathbf{V}\mathbf{x}(t) = (E\{\mathbf{x}(t)\mathbf{x}(t)^{\mathrm{T}}\})^{-\frac{1}{2}}\mathbf{x}(t)$ such that the components of $\tilde{\mathbf{x}}(t) = \mathbf{V}\mathbf{x}(t)$ are unit variance and uncorrelated, and the matrix $\mathbf{W}$ is constrained to be orthogonal [6,12].

We present the following constrained maximization problem based on both non-Gaussian and nonlinear autocorrelation of the desired source:

$$\begin{aligned}\max_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) &= \lambda E\{G(\tilde{y}_i(t))\} + (1-\lambda)E\{G(\tilde{y}_i(t))G(\tilde{y}_i(t-\tau))\} \\ &= \lambda E\{G(\mathbf{w}_i^T\tilde{\mathbf{x}}(t))\} + (1-\lambda)E\{G(\mathbf{w}_i^T\tilde{\mathbf{x}}(t))G(\mathbf{w}_i^T\tilde{\mathbf{x}}(t-\tau))\},\end{aligned} \tag{5}$$

where $\tilde{y}_i(t) = \mathbf{w}_i^T\tilde{\mathbf{x}}(t)$, λ $(0 \le \lambda \le 1)$ defines the relative weighting afforded to non-Gaussian sources and (1-λ) defines the relative weighting afforded to nonlinear autocorrelation sources. G is a differentiable function, which measures the non-Gaussian and the nonlinear autocorrelation degree of the desired source. An example of choice is $G(u) = -\log\cosh(u)$.

The objective function combines both of non-Gaussianity and nonlinear autocorrelation of the desired source signal. In fact, assume that the source signal has no time dependency, the objective function reduces to:

$$\max_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) = E\{G(\tilde{y}_i(t))\} = E\{G(\mathbf{w}_i^T\tilde{\mathbf{x}}(t))\}. \tag{6}$$

This is the well-known ICA objective function based on the non-Gaussianity of the source signal [12]. In particular, if the function G in (6) is chosen as the log density of the component $\tilde{y}_i$, we can find one independent component by maximizing the objective function in (6) [12]. Thus, the function G should be chosen based on the probability distribution of the desired source, which is a well-known fact in BSS [12]. In BSS, if the signals to be reconstructed satisfy certain properties, an exact form of the function G is not required in order to achieve the desired estimation results [6,12]. We may therefore optimistically assume that the exact form of the function G is not very important here either, as long as it is qualitatively similar enough. As a special case, assume that we choose $G(u) = u^2$, the optimization problem which leads to an algorithm is formally identical to the one encountered when the value of kurtosis is maximized to find the most non-Gaussian directions in ICA. However, the algorithm using $G(u) = u^2$ is sensitive to the outliers as the kurtosis-based FastICA [12]. However, the nonlinear function $G(u) = -\log(\cosh(u))$ is better than the kurtosis-based function, due to its better analysis property and robustness to outliers [12]. Thus, we choose the nonlinear function $G(u) = -\log(\cosh(u))$ in the paper.

2.2. Learning algorithm

Maximizing the objective function in (5), we derive a gradient ascent blind source separation (BSS) algorithm. The gradient of $\Psi(\mathbf{w}_i)$ with respect to $\mathbf{w}_i$ can be obtained as

$$\frac{\partial \Psi(\mathbf{w}_i)}{\partial \mathbf{w}_i} = \lambda E\{g(\tilde{y}_i(t))\tilde{\mathbf{x}}(t)\} + (1-\lambda)E\{g(\tilde{y}_i(t))G(\tilde{y}_i(t-\tau))\tilde{\mathbf{x}}(t) + G(\tilde{y}_i(t))g(\tilde{y}_i(t-\tau))\tilde{\mathbf{x}}(t-\tau)\}. \tag{7}$$

The function g is the derivative of G. Thus, the gradient ascent BSS algorithm for non-Gaussian and nonlinear autocorrelation sources is obtained as follows:

(1) Center the data to make its mean zero and whiten the data to give $\tilde{\mathbf{x}}(t)$. Choose random initial values for the parameters.
(2) Update the weight vector by

$$\begin{aligned}&\mathbf{w}_i \leftarrow \mathbf{w}_i + \mu(\lambda E\{g(\tilde{y}_i(t))\tilde{\mathbf{x}}(t)\} + (1-\lambda)E\{g(\tilde{y}_i(t))G(\tilde{y}_i(t-\tau))\tilde{\mathbf{x}}(t) + G(\tilde{y}_i(t))g(\tilde{y}_i(t-\tau))\tilde{\mathbf{x}}(t-\tau)\}),\\ &\mathbf{w}_i \leftarrow \mathbf{w}_i/\|\mathbf{w}_i\|,\end{aligned} \tag{8}$$

where μ is a learning rate.
(3) If not converged, go back to step (2).

Note that convergence of the algorithm means that the distance between the old and new values of $\mathbf{w}_i$ is very near (such as 0.00001 in the paper).

The above algorithm is only used to estimate one source, we can estimate all the source signals or the separating matrix $\mathbf{W} = (\mathbf{w}_1, \ldots, \mathbf{w}_n)^{\mathrm{T}}$ by the symmetric orthogonalization or a deflation scheme (one-by-one estimation) [12].

3. Theoretical properties

We analyze the stability conditions about the presented algorithm along similar lines as the ICA algorithm for maximizing the non-Gaussianity [8,12].[1]

Theorem 1. *Assume that the input data follows the model (1) with whitened data $\tilde{\mathbf{x}} = \mathbf{VAs}$ where $\mathbf{V}$ is the whitening matrix, and $\Psi(\mathbf{w})$ is a sufficiently smooth even function. Furthermore, assume that $\{s_i, s_{i,\tau}\}$ and $\{s_j, s_{j,\tau}\}$ $(\forall j \neq i)$ are mutually independent. Then the local maxima of $\Psi(\mathbf{w})$ under the constraint $\|\mathbf{w}\| = 1$ include one row of the inverse of the mixing matrix $\mathbf{VA}$ such that the corresponding desired source signal s_i satisfy*

$$\begin{aligned}&E\{\lambda(g'(s_i) - s_i g(s_i)) + (1-\lambda)(g'(s_i)G(s_{i,\tau}) + g'(s_{i,\tau})G(s_i)\\ &\quad + 2s_j s_{j,\tau} g(s_i)g(s_{i,\tau}) - s_i g(s_i)G(s_{i,\tau}) - s_{i\tau} g(s_{i,\tau})G(s_i))\} < 0, \quad (\forall j \neq i),\end{aligned} \tag{9}$$

where g is the derivative of G, and g' is the derivative of g.

Proof. Assume that $\{s_i, s_{i,\tau}\}$ and $\{s_j, s_{j,\tau}\}$ $(\forall j \neq i)$ are mutually independent, and the vector $\tilde{\mathbf{x}}$ is white, we have:

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{VA}E\{\mathbf{ss}^T\}\mathbf{A}^T\mathbf{V}^T = (\mathbf{VA})(\mathbf{VA})^{\mathrm{T}} = \mathbf{I}, \tag{10}$$

which means the matrix $\mathbf{VA}$ is orthogonal. Make the change of coordinates $\mathbf{p} = (p_1, \ldots, p_n)^{\mathrm{T}} = \mathbf{A}^T\mathbf{V}^T\mathbf{w}$, we have:

$$\Psi(\mathbf{p}) = E\{\lambda G(\mathbf{p}^T\mathbf{s}) + (1-\lambda)G(\mathbf{p}^T\mathbf{s})G(\mathbf{p}^T\mathbf{s}_\tau)\}. \tag{11}$$

Analyzing the stability of the point $\mathbf{p} = \mathbf{e}_1$ is enough, where $\mathbf{e}_1 = (1, 0, 0, \ldots)^{\mathrm{T}}$ (because $\Psi(\mathbf{w})$ is even function, nothing changes for $(-1, 0, 0, \ldots)^{\mathrm{T}}$) [8,12]. At the point $\mathbf{p} = \mathbf{e}_1$, we compute its gradient and Hessian and use the assumptions of independency, and make a small perturbation $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots)^{\mathrm{T}}$, and we have:

$$\begin{aligned}\Psi(\mathbf{e}_1 + \varepsilon) &= \Psi(\mathbf{e}_1) + \varepsilon^T\frac{\partial \Psi(\mathbf{e}_1)}{\partial \mathbf{p}} + \frac{1}{2}\varepsilon^T\frac{\partial^2 \Psi(\mathbf{e}_1)}{\partial \mathbf{p}^2}\varepsilon + o(\|\varepsilon\|^2)\\ &= \Psi(\mathbf{e}_1) + E\{\lambda(s_1 g(s_1)) + (1-\lambda)(s_1 g(s_1)G(s_{1,\tau}) + s_{1\tau} g(s_{1,\tau})G(s_1))\}\varepsilon_1\\ &\quad + \frac{1}{2}[E\{\lambda(s_1^2 g'(s_1)) + (1-\lambda)(s_1^2 g'(s_1)G(s_{1,\tau}) + s_{1,\tau}^2 g'(s_{1,\tau})G(s_1) + 2s_1 s_{1,\tau} g(s_1)g(s_{1,\tau}))\}]\varepsilon_1^2\\ &\quad + \frac{1}{2}\sum_{j>1}[E\{\lambda g'(s_1) + (1-\lambda)(g'(s_1)G(s_{1,\tau}) + g'(s_{1,\tau})G(s_1) + 2s_j s_{j,\tau} g(s_1)g(s_{1,\tau}))\}\varepsilon_j^2] + o(\|\varepsilon\|^2).\end{aligned} \tag{12}$$

[1] We will drop the sampling index t for simplicity, i.e., $y_\tau = y(t-\tau)$, $s_{i,\tau} = s_i(t-\tau)$, and $\mathbf{w} = \mathbf{w}_i$.

Using the constraint $\|\mathbf{w}\| = 1$ and the matrix $\mathbf{VA}$ is orthogonal, we obtain $\|\mathbf{p}\| = \|\mathbf{A}^T\mathbf{V}^T\mathbf{w}\| = 1$. Thus, we get $\varepsilon_1 = \sqrt{1-\varepsilon_2^2-\varepsilon_3^2-\cdots} - 1$. Using the property that $\sqrt{1-\alpha} = 1-\frac{\alpha}{2}+o(\alpha)$, the term of order ε_1^2 is $o(\|\varepsilon\|^2)$, i.e., of higher order, and can be neglected [8,12]. Computing the first-order approximation for ε_1, we have $\varepsilon_1 = -\sum_{j>1}\frac{\varepsilon_j^2}{2} + o(\|\varepsilon\|^2)$ [8,12], which finally gives

$$\Psi(\mathbf{e}_1+\varepsilon) = \Psi(\mathbf{e}_1) + \frac{1}{2}\sum_{j>1} E\{\lambda(g'(s_1) - s_1 g(s_1)) + (1-\lambda)(g'(s_1)G(s_{1,\tau}) + g'(s_{1,\tau})G(s_1) + 2s_j s_{j,\tau} g(s_1)g(s_{1,\tau}) - s_1 g(s_1)G(s_{1,\tau}) - s_{1,\tau} g(s_{1,\tau})G(s_1))\}\varepsilon_j^2 + o(\|\varepsilon\|^2). \tag{13}$$

The formula shows that $\mathbf{p} = \mathbf{e}_1$ is an extremum, thus, we finish the theorem. □

If we choose only one time delay, we can similarly prove the following corollary.

Corollary 1. *Assume that the signal has no time dependencies and $G(s_{i,\tau}) = 1$, from the stability condition (9), we obtain:*

$$E\{g'(s_i) - s_i g(s_i)\} < 0. \tag{14}$$

This is in fact the well-known stability condition of independent component analysis (ICA) based on non-Gaussianity [12].

Corollary 2. *Assume that the signal is Gaussian, but has nonlinear time dependency, from the stability condition (9), we obtain:*

$$E\{g'(s_i)G(s_{i,\tau}) + g'(s_{i,\tau})G(s_i) + 2s_j s_{j,\tau} g(s_i)g(s_{i,\tau}) - s_i g(s_i)G(s_{i,\tau}) - s_{i\tau} g(s_{i,\tau})G(s_i)\} < 0, \quad (\forall j \neq i). \tag{15}$$

4. Experimental results

We demonstrate when we choose $G(u) = -\log(\cosh(u))$, the proposed algorithm can separate the source signals with square temporal autocorrelation (meaning that $E\{s_i(t)^2 s_i(t-\tau)^2\} > 0$) and non-Gaussianity.

We create 20 artificial signals which have square temporal autocorrelation and non-Gaussian as follows: 10 sources with sparse non-Gaussian (ACvsparse10.mat, taken from http://www.bsp.brain.riken.jp/ICALAB/ICALABSignalProc/benchmarks/) and 10 sources with square temporal autocorrelation (with Gaussian marginal distributions, zero linear autocorrelations). We create 10 square temporal autocorrelation as follows: first, we create 10 signals using a first-order autoregressive model with constant variances of the innovations, with 1000 time points. The signals are created with Gaussian innovations and have identical autoregressive coefficients (0.9). All these innovations have constant unit variance. Then, the signs of the signals are completely randomized by multiplying each signal by a binary i.i.d. signal that takes the values ± 1 with equal probabilities. The waveforms of 20 original signals are shown in Fig. 1.

Thus, source signals of this kind could not be separated by ordinary source separation methods based on non-Gaussianity, such as FastICA [12], or only time-structure information, such as linear correlation method AMUSE [26] and SOBI [4], coding complexity [10,20,21,23], temporal predictability [18], nonstationarity [11,15,17], energy predictability [22], nonlinear innovation [24], and nonlinear autocorrelation [25].

The source signals are mixed with 20×20 random matrices. The waveforms of 20 mixed signals are shown in Fig. 2. The proposed algorithm with the symmetric orthogonalization is used to estimate the separating matrix $\mathbf{W}$. In order to measure the accuracy of separation, we calculate the performance index [1,22,24,25]

$$\mathbf{PI} = \frac{1}{n^2}\left\{\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{|p_{ij}|}{\max\limits_k |p_{ik}|} - 1\right) + \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\frac{|p_{ij}|}{\max\limits_k |p_{kj}|} - 1\right)\right\}, \tag{16}$$

where p_{ij} is the ijth element of $n \times n$ matrix $\mathbf{P} = \mathbf{WVA}$. The larger the value $\mathbf{PI}$ is, the poorer the statistical performance of a BSS algorithm.

Furthermore, the following conditions are included in the comparison of the algorithm:

(i) The presented algorithm with the different learning rates, such as $\mu = 1$, $\mu = 0.1$, $\mu = 0.01$ and $\mu = 0.001$ (We choose the initial values $\tau = 1$ and $\lambda = 0.5$ in the cases);
(ii) The presented algorithm with the different time lags, such as $\tau = 1$, $\tau = 2$, $\tau = 3$ and $\tau = 4$ (We choose the initial values $\mu = 1$ and $\lambda = 0.5$ in the cases);
(iii) The presented algorithm with the different relative weightings, such as $\lambda = 0.5$, $\lambda = 0.4$, $\lambda = 0.3$ and $\lambda = 0.2$ (We choose the initial values $\mu = 1$ and $\tau = 1$ in the cases).

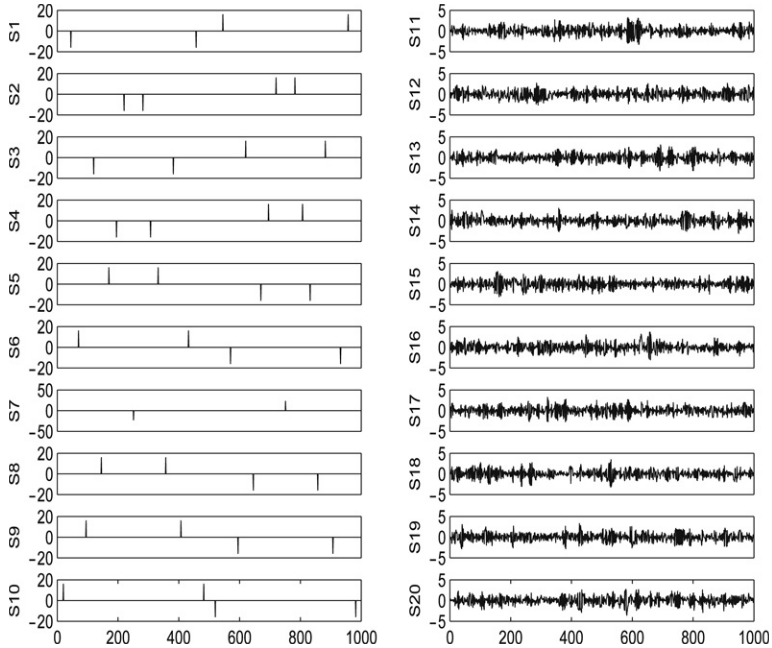

Fig. 1. 20 source signals.

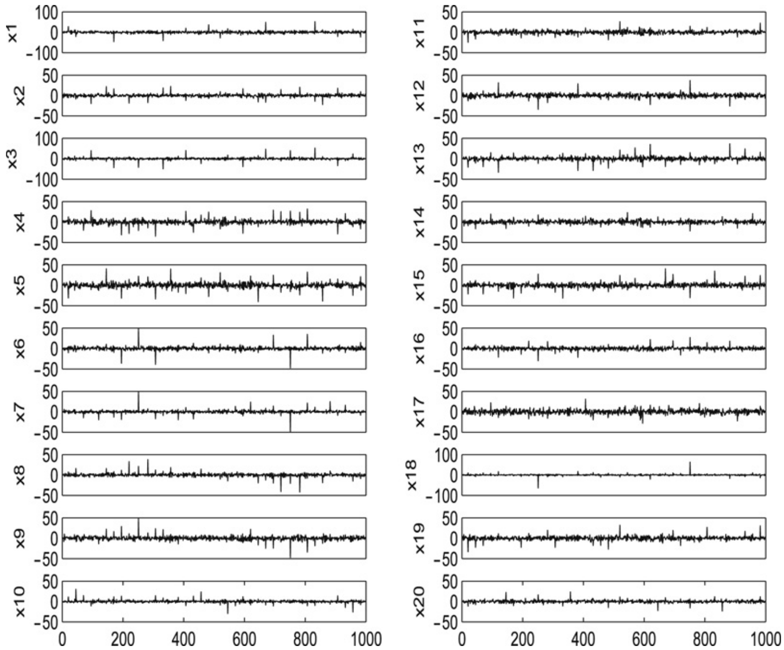

Fig. 2. 20 mixed signals.

Fig. 3 shows 20 separated signals by the proposed algorithm when the initial values are $\mu = 1$, $\tau = 1$ and $\lambda = 0.5$. From the results, we see that all the source signals are recovered very well. Fig. 4 shows the average performance indexes over 100 independent trials against iteration numbers by the proposed algorithm with the different learning rates μ in the case (i). We can see that the performance with $\mu = 1$ is fastest and best. However, when we choose the learning rate $\mu = 0.0001$, in fact, the algorithm is fail. We can seen that the convergence of the gradient ascent algorithm is dependent on the learning rate μ. Fig. 5 shows the average performance indexes over 100 independent trials against iteration numbers by the proposed algorithm with the different time lags τ in the case (ii). We can see that the performance with $\tau = 1$ is fastest and best, however, the performances with other lags τ are also similar to $\tau = 1$, except the slower convergence speed. Also, Fig. 6 shows the average performance indexes over 100 independent trials against iteration numbers by the proposed algorithm

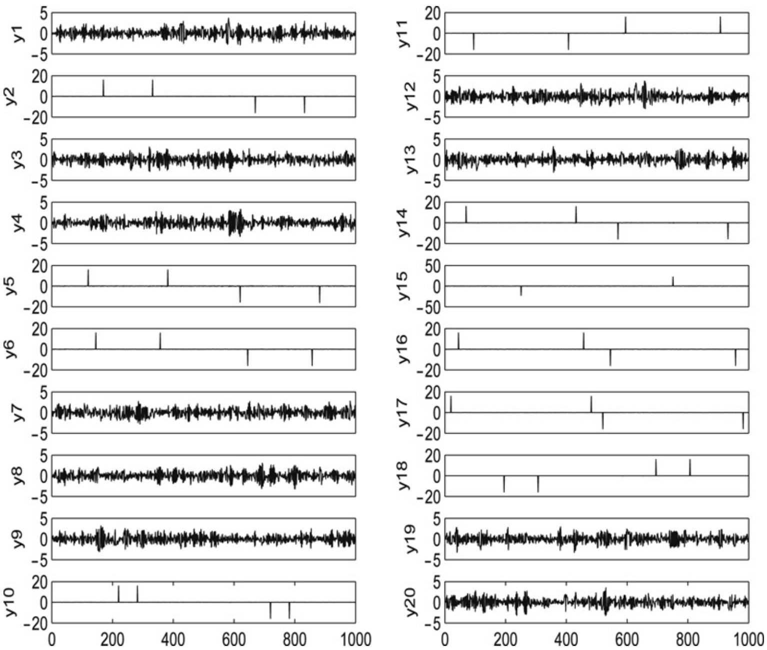

Fig. 3. 20 separated signals.

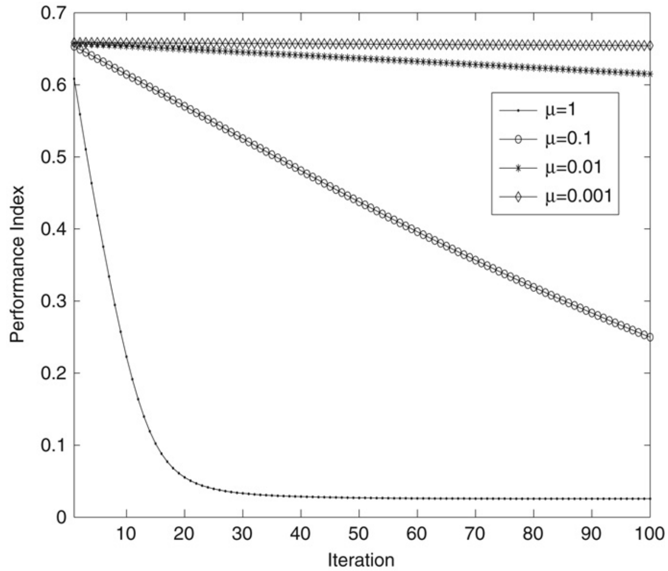

Fig. 4. The average performance indexes over 100 independent trials against iteration numbers by the proposed algorithm with the different learning rates μ.

with the different relative weightings λ in the case (iii). We can see that the performance with $\lambda = 0.5$ is fastest and best, which is similar to the case (ii).

5. Conclusions

We have proposed an algorithm for blind source separation based on non-Gaussianity and nonlinear autocorrelation of source signals and have analyzed its stability conditions. We show that the BSS problem can be solved by maximizing the non-Gaussianity and nonlinear autocorrelation of sources, which extends the previous situations. When sources have square temporal autocorrelation and non-Gaussianity, we demonstrate that the efficient implementation of the method. Thus, the proposed algorithm, which is based on the two statistical properties of sources, provides a novel way to perform BSS.

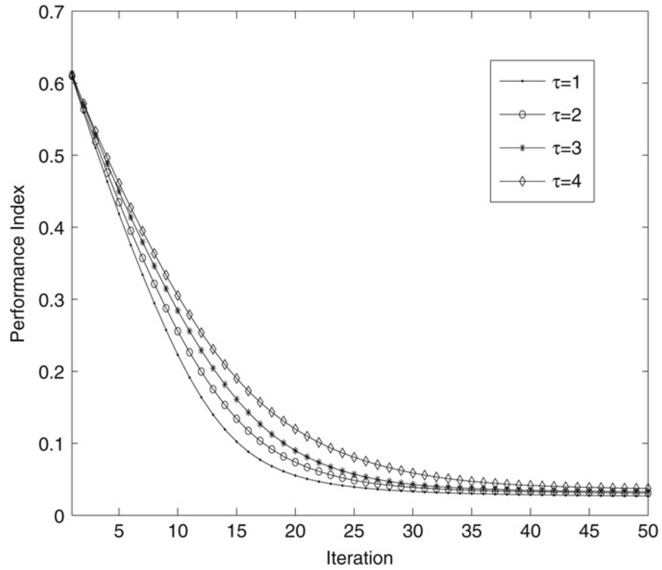

Fig. 5. The average performance indexes over 100 independent trials against iteration numbers by the proposed algorithm with the different time lags τ.

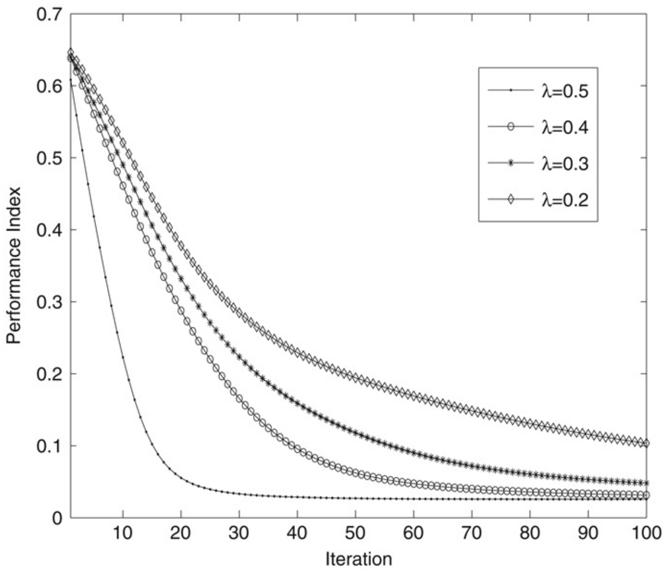

Fig. 6. The average performance indexes over 100 independent trials against iteration numbers by the proposed algorithm with the different relative weightings λ.

The objective function presented here is closely connected to several other BSS objective functions. First, assume that the signal has no time dependency, then our objective function reduces to the well-known one-unit objective function of ICA. When $\lambda = 0$ and $G = u^2$, the objective function can lead to the nonstationary BSS algorithm [11], the fourth-order cross-cumulant in the algorithm [11] could be considered as a normalized version of the autocorrelation of square (energy). In particular, if the signals have no linear time-correlations, and $E\{s_l^2\} = E\{s_{l,\tau}^2\} = 1$ $(\forall l)$, the cross-cumulant is equal to the autocorrelation of the square (energy) as the objective function of the presented algorithm when the nonlinear $G(u) = u^2$ is chosen.

Due to the use of the gradient ascent, the convergence of the algorithm presented here is dependent on the learning rate μ, and the theoretical proof about the convergence of the gradient ascent BSS algorithm is somewhat difficult. Further theoretical endeavors of the gradient ascent BSS algorithm and the improved faster algorithm are subjects for future study.

Acknowledgments

The authors wish to thank all anonymous reviewers who provided insightful and helpful comments.

References

[1] S.I. Amari, A. Cichocki, H. Yang, A new learning algorithm for blind source separation, Advances in Neural Information Processing Systems 8 (1996) 757–763.
[2] A.K. Barros, A. Cichocki, Extraction of specific signals with temporal structure, Neural Computation 13 (9) (2001) 1995–2003.
[3] A. Bell, T. Sejnowski, An information-maximization approach to blind separation and blind deconvolution, Neural Computation 7 (6) (1995) 1129–1159.
[4] A. Belouchrani, K.A. Meraim, J.-F. Cardoso, E. Moulines, A blind source separation technique based on second order statistics, IEEE Transactions on Signal Processing 45 (2) (1997) 434–444.
[5] J.-F. Cardoso, B.H. Laheld, Equivariant adaptive source separation, IEEE Transactions on Signal Processing 44 (12) (1996) 3017–3030.
[6] A. Cichocki, S.-I. Amari, Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications, Wiley, New York, 2002.
[7] P. Comon, Independent component analysis—a new concept? Signal Processing 36 (1994) 287–314.
[8] A. Hyvärinen, E. Oja, Independent component analysis by general nonlinear Hebbian-like learning rules, Signal Processing 64 (3) (1998) 301–313.
[9] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Transactions on Neural Networks 10 (3) (1999) 626–634.
[10] A. Hyvärinen, Complexity pursuit: Separating interesting components from time-series, Neural Computation 13 (4) (2001) 883–898.
[11] A. Hyvärinen, Blind source separation by nonstationarity of variance: A cumulant-based approach, IEEE Transactions on Neural Networks 12 (6) (2001) 1471–1474.
[12] A. Hyvärinen, J. Karhunen, E. Oja, Independent component analysis, Wiley, New York, 2001.
[13] C. Jutten, J. Herault, Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture, Signal Processing 24 (1991) 1–10.
[14] T.-W. Lee, M. Girolami, T. Sejnowski, Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources, Neural Computation 11 (2) (1999) 417–441.
[15] K. Matsuoka, M. Ohya, M. Kawamoto, A neural net for blind separation of nonstationary signals, Neural Networks 8 (3) (1995) 411–419.
[16] L. Molgedey, H.G. Schuster, Separation of a mixture of independent signals using time delayed correlations, Physical Review Letters 72 (23) (1994) 3634–3637.
[17] D.-T. Pham, J.-F. Cardoso, Blind separation of instantaneous mixtures of non stationary sources, IEEE Transactions on Signal Processing 49 (9) (2001) 1837–1848.
[18] J.V. Stone, Blind source separation using temporal predictability, Neural Computation 13 (2001) 1559–1574.
[19] Z. Shi, H. Tang, Y. Tang, A new fixed-point algorithm for independent component analysis, Neurocomputing 56 (2004) 467–473.
[20] Z. Shi, H. Tang, Y. Tang, A fast fixed-point algorithm for complexity pursuit, Neurocomputing 64 (2005) 529–536.
[21] Z. Shi, C. Zhang, Gaussian moments for noisy complexity pursuit, Neurocomputing 69 (7–9) (2006) 917–921.
[22] Z. Shi, C. Zhang, Energy Predictability to Blind Source Separation, Electronics Letters 42 (17) (2006) 1006–1008.
[23] Z. Shi, C. Zhang, Semi-blind source extraction for fetal electrocardiogram extraction by combining non-Gaussianity and time-correlation, Neurocomputing 70 (2007) 1574–1581.
[24] Z. Shi, C. Zhang, Nonlinear innovation to blind source separation, Neurocomputing 71 (2007) 406–410.
[25] Z. Shi, Z. Jiang, F. Zhou, A fixed-point algorithm for blind source separation with nonlinear autocorrelation, Journal of Computational and Applied Mathematics (2008) doi:10.1016/j.cam.2008.03.009.
[26] L. Tong, R.-W. Liu, V. Soon, Y.-F. Huang, Indeterminacy and identifiability of blind identification, IEEE Transactions on Circuits and Systems 38 (5) (1991) 499–509.
[27] H. Zhang, C. Guo, Z. Shi, E. Feng, A new constrained fixed-point algorithm for ordering independent components, Journal of Computational and Applied Mathematics (2008) doi:10.1016/j.cam.2007.09.010.