J. Appl. Math. & Informatics Vol. 26(2008), No. 5 - 6, pp. 1149 - 1159
Website: http://www.kcam.biz

A TWO-STAGE SOURCE EXTRACTION ALGORITHM FOR TEMPORALLY CORRELATED SIGNALS BASED ON ICA-R

HONGJUAN ZHANG*, ZHENWEI SHI, CHONGHUI GUO AND ENMIN FENG

ABSTRACT. Blind source extraction (BSE) is a special class of blind source separation (BSS) methods, which only extracts one or a subset of the sources at a time. Based on the time delay of the desired signal, a simple but important extraction algorithm (simplified " BC algorithm")was presented by Barros and Cichocki. However, the performance of this method is not satisfying in some cases for which it only carries out the constrained minimization of the mean squared error. To overcome these drawbacks, ICA with reference (ICA-R) based approach, which considers the higher-order statistics of sources, is added as the second stage for further source extraction. Specifically, BC algorithm is exploited to roughly extract the desired signal. Then the extracted signal in the first stage, as the reference signal of ICA-R method, is further used to extract the desired sources as cleanly as possible. Simulations on synthetic data and real-world data show its validity and usefulness.

AMS Mathematics Subject Classification : TP391, 90C90, 82C43.
Key words and phrases : Blind signal extraction; Independent component analysis with reference; Blind source separation; Temporally correlated signal

1. Introduction

Independent component analysis (ICA) is a very general purpose statistical technique in which observed random data are linearly transformed into components that are maximally independent from each other, and simultaneously have interesting distributions. ICA has become one of the exciting new topics in the field of neural networks, especially unsupervised learning, and more generally in advanced statistics and signal processing. Its applicability includes blind source separation (BSS) [1, 2, 3], feature extraction [4, 5], etc.. Generally speaking, two kinds of promising techniques have been proposed. One is BSS (or ICA), which estimate same number of independent components (ICs) as the

---

Received August 15, 2007. Revised February 17, 2008. Accepted February 25, 2008.
*Corresponding author. This work was supported by Natural Science Foundation of China under grant No. 10571018, No. 60605002, No. 70431001.

observed mixtures [6, 7]. The classical ICA is only capable of separating the full space of ICs. However, extraction all the source signals from a large number observed sensor signals, for example, a magnetoencephalographic (MEG) measurement, which may output hundreds of recordings, could take a long time and in these signals only a very few are desired with given characteristics instead of all sources. Therefore it is necessary to develop another reliable, robust and effective technique to extract only the desired signals that are potentially interesting and contain useful information. Recently, Lu et al. [8, 9, 10] proposed a good candidate, that is ICA-R, for extracting several source signals from a large number of observed signals. It makes the best of the traces of the interesting sources, which are referred to as the reference signals. Reference signals carry some information to distinguish the desired components but are not identical to the corresponding sources. In order to design them, one must know lots of priori information. However it is not available in many cases usually. So this method is limited to the ones that the reference signals can be easily obtained. As to the reference selection of the desired signals, [11] has also introduced a second-order statistics based approach to reliably find suitable reference signals for weak temporally correlated source signals.

In many applications, such as ECG extraction, the desired source signal is periodic or quasi-periodic. So the period property can be used to extract the desired source signal. Barros and Cichocki [12] provided an algorithm that can quickly extract the desired source signal with a specific period. And this algorithm in all the cases can extract the desired source as long as they are decorrelated and show a temporal structure. However, this method only carries out the constrained minimization of the mean squared error, which dose not describe the probability distribution of the innovation of the signals well. It is a possible reason why this algorithm often contains noise contamination in the recovered signals [13]. To overcome these drawbacks, we propose a two-stage method for extracting the desired signals in this paper. At the first stage, the priori information about the autocorrelation function of the primary sources is exploited to roughly extract the desired signals from their linear mixtures. At the second stage, the extracted signals in the first stage, as the reference signals of ICA-R are further used to extract the source of interest as cleanly as possible. This method can be viewed from two aspects. On the one hand, when the signal extracted by BC algorithm [12] is used to be reference signal of ICA-R, we can obtain a very suitable reference signal for ICA-R algorithm, which solved the difficult problem of designing reference signal in ICA-R. Whereas adding the postprocessing of the ICA-R can also improve the quality of the extracted signals, which may be a rewarding complementarity of BC algorithm. Simulations on synthetic data demonstrated the validity of the proposed algorithm, in which the accuracy of the extracted signal improved greatly and experiments on real-world fetal electrocardiogram (FECG) data also show its usefulness.

This manuscript is organized as follows. The next section introduces BC algorithm and the technique of ICA-R, meanwhile derives the two-stage blind source extraction method based on BC algorithm and ICA-R algorithm. Section 3 demonstrates the proposed technique with experiments using artificial data and real-world FECG data. The final section provides discussions and conclusions.

2. Proposed Approach

2.1. BC algorithm. Let us denote the observed sensor signals $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^T$ described by matrix equation: $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$ where $\mathbf{A}$ is an $n \times n$ unknown nonsingular matrix, and $\mathbf{s}(t) = (s_1(t), \cdots, s_n(t))^T$ is a vector of unknown temporally correlated sources. We assume that desired source signal s_i has specific temporal structures, but they do not necessarily have to be statistically independent [15, 16]. They can be modelled by linear autoregressive models, which has just one predicting term as:

$$s_i(t) = as_i(t - \tau_i) + \delta s_i(t), \tag{1}$$

where $\delta s_i(t)$ is a zero-mean, independent identically distributed (i.i.d) time series called innovations and t is a specific time delay satisfying the relations $E\{s_i(t)s_i(t - \tau_i)\} \neq 0$ and $E\{s_i(t)s_j(t - \tau_i)\} = 0, \forall i \neq j$. Because we want to extract only a desired source signal, for this purpose we design a single neural processing unit described as:

$$\begin{aligned} y(t) &= \mathbf{w}^T\mathbf{x}(t), && (2)\\ \epsilon(t) &= y(t) - by(t - \tau), && (3) \end{aligned}$$

where $\mathbf{w} = (w_1, \cdots, w_n)^T$ is the weight vector, ϵ represents an innovation, b is a coefficient and $y(t)$ denotes a recovered source signal at time t.

Assuming that the measured sensor signals $\mathbf{x}$ have already been followed by an $n \times n$ whitening filter $\mathbf{V}$ such that the components of $\tilde{\mathbf{x}}(t) = \mathbf{V}\mathbf{x}(t)$ are unit variance and uncorrelated. Then the constrained minimization of the mean squared error $\xi(\mathbf{w}, b) = E\{\epsilon^2\}$ is used for two sets of parameters $\mathbf{w}$ and b. Minimizing $\xi(\mathbf{w}, b)$ obtained the BC algorithm when its gradient reaches zeros in relation to $\mathbf{w}$ and b. And, taking into account that $y = \mathbf{w}^T\tilde{\mathbf{x}}$ and $E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{I}$, which finally lead to a very simple learning rule:

$$\mathbf{w} = E\{\tilde{\mathbf{x}}y_\tau\}. \tag{4}$$

where $y_\tau = \mathbf{w}^T\tilde{\mathbf{x}}(t - \tau)$.

2.2. ICA-R algorithm. ICA-R algorithm was proposed by Lu et al. in [8, 9, 10]. The goal of ICA-R is to obtain a learning algorithm that satisfies the following two conditions simultaneously:

- every output is one of the ICs mixed in the input signals,
- the extracted ICs is the closest to the reference signal in some distance criterion.

Let us consider $\mathbf{x}(t) = \left(x_1(t), \ldots, x_n(t)\right)^T$ is an observed signal, and $\mathbf{y}(t) = \left(y_1(t), \ldots, y_n(t)\right)^T$ is the recovered signal of ICA-R algorithm. The output $\mathbf{y}(t)$

is given by $\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t)$, where $\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_n)^T$ is the matrix containing n weight vector $\mathbf{w}_j = (w_{j1}, \cdots, w_{jn})^T$ which need to be learned. In the single source extraction, our purpose is to find from the mixed vector $\mathbf{x}$ one given component s_i of the source signal $\mathbf{s}$. For carrying this out, we use all the components of the input vector $\mathbf{x}$ to obtain the output signal, just like above (2), which denoted by $y(t) = \mathbf{w}^T\mathbf{x}(t)$, where $\mathbf{w}$ is a vector. And $r(t)$ denotes the reference signal which includes some prior information of the desired source signal. A measure of the closeness between the the desired output $y(t)$ and the reference signal $r(t)$ is given by $\varepsilon(y(t), r(t))$, in general which defines as $\varepsilon(y(t), r(t)) = E\{(y(t) - r(t))^2\}$. For simplicity of the equations the time index t is omitted in the following.

Suppose the contrast function is given by negentropy, defined by

$$J(\mathbf{w}) \approx \rho[E\{G(y)\} - E\{G(v)\}]^2, \tag{5}$$

where ρ is a positive constant, v is a Gaussian variable with zero mean and unit variance, and $G(\cdot)$ denotes any non-quadratic function [17]. Assuming that one of the ICs is the one and the only one closest to reference r, i.e.

$$\varepsilon\big(\mathbf{w}^{*T}\mathbf{x}, r\big) < \varepsilon\big(\mathbf{w}_1^T\mathbf{x}, r\big) \leq \ldots \leq \varepsilon\big(\mathbf{w}_{n-1}^T\mathbf{x}, r\big), \tag{6}$$

where $\mathbf{w}^*$ corresponds to the desired output. Hence, a threshold $\xi \in \big[\varepsilon\big(\mathbf{w}^{*T}\mathbf{x}, r\big), \varepsilon\big(\mathbf{w}_1^T\mathbf{x}, r\big)\big]$ can be used to distinguish the desired signal from others. So there exists a threshold parameter ξ such that $g(\mathbf{w}) = \varepsilon(y, r) - \xi \leq 0$ is satisfied only when $y = \mathbf{w}^{*T}\mathbf{x}$, but not with any other vectors $\mathbf{w}_i (\neq \mathbf{w}^*)$.

Therefore, defining the following constrained contrast function:

$$\begin{cases} \max J(\mathbf{w}) \approx \rho[E\{G(y)\} - E\{G(v)\}]^2 \\ \text{s.t.} \quad g(\mathbf{w}) = \varepsilon(y, r) - \xi \leq 0, h(\mathbf{w}) = E\{y^2\} - 1 = 0. \end{cases} \tag{7}$$

Based on above contrast function and the approximate Newton learning, Lu et al. provided the one-unit ICA-R learning algorithm [8, 9, 10]:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta R_{\mathbf{xx}}^{-1}\mathcal{L}'(\mathbf{w}_k)/s(\mathbf{w}_k), \tag{8}$$

where

$$\mathcal{L}'(\mathbf{w}_k) = \bar{\rho}E\{\mathbf{x}G'(y)\} - \frac{1}{2}\mu E\{\mathbf{x}g'(y)\} - \lambda E\{\mathbf{x}y\}, \tag{9}$$

$$s(\mathbf{w}_k) = \bar{\rho}E\{G''(y)\} - \frac{1}{2}\mu E\{g''(y)\} - \lambda, \tag{10}$$

and k denotes the iteration index, η is the learning rate, $R_{\mathbf{xx}} = E\{\mathbf{x}\mathbf{x}^T\}$, $\bar{\rho} = \pm\rho$ whose positive or negative sign coincident with $[E\{G(y)\} - E\{G(v)\}]$, $G'(y)$ and $g'(y)$ are the first derivatives of $G(y)$ and $g(y)$ with respect to y, $G''(y)$ and $g''(y)$ corresponding to their second derivatives.

Notice that, when observed signals $\mathbf{x}$ are whited as $\tilde{\mathbf{x}}$, $R_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}} = E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{I}$. Hence (8) can be simplified:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta\mathcal{L}'(\mathbf{w}_k)/s(\mathbf{w}_k). \tag{11}$$

And we can omit the equality constraint $h(\mathbf{w})$ and its corresponding Lagrange multiplier λ because the vector $\mathbf{w}$ can be normalized after the learning algorithm (11). Therefore:

$$\mathcal{L}'(\mathbf{w}_k) = \bar{\rho}E\{\mathbf{x}G'(y)\} - \frac{1}{2}\mu E\{\mathbf{x}g'(y)\}, \quad (12)$$

$$s(\mathbf{w}_k) = \bar{\rho}E\{G''(y)\} - \frac{1}{2}\mu E\{g''(y)\}. \quad (13)$$

The multiplier μ is updated by gradient-ascent method:

$$\mu_{k+1} = \max\{0, \mu_k + \gamma g(\mathbf{w}_k)\}. \quad (14)$$

2.3. Proposed algorithm. In the first stage, the BC algorithm in Section 2.1 is considered as the main procedure. The goal of this stage is to roughly extract the desired source signal through exploiting its autocorrelation structure. Then this output signal as the reference signal to further extract the desired signal based on ICA-R (11-14) [18].

Specifically, the two-stage algorithm can be described as follows:

- Step 1: Center the observed signals $\mathbf{x}$ and white them to $\tilde{\mathbf{x}}$ by an $n \times n$ whitening filter $\mathbf{V}$.
- Step 2: Random initialize $\mathbf{w}'_0$ and update it by $\mathbf{w}'_{k+1} = E\{\tilde{\mathbf{x}}y_p\}$.
- Step 3: Normalize $\mathbf{w}'$ by $\mathbf{w}'_{k+1} = \mathbf{w}'_{k+1}/\|\mathbf{w}'_{k+1}\|$, until $\|\mathbf{w}'_k\mathbf{w}'_{k+1}\|$ approaches 1 (up to a small error), then let $r = \mathbf{w}'^T_{k+1}\tilde{\mathbf{x}}$, go to Step 4. Otherwise, go back to Step 2.
- Step 4: Choose penalty parameter γ and initialize $\mathbf{w}_0$, the Lagrange multiplier μ.
- Step 5: Update the Lagrange multiplier μ by $\mu_{k+1} = \max\{0, \mu_k + \gamma g(\mathbf{w}_k)\}$.
- Step 6: Choose a suitable learning rate η and update $\mathbf{w}$ by $\mathbf{w}_{k+1} = \mathbf{w}_k - \eta\mathcal{L}'(\mathbf{w}_k)/s(\mathbf{w}_k)$.
- Step 7: Normalize $\mathbf{w}$ by $\mathbf{w}_{k+1}/\|\mathbf{w}_{k+1}\|$.
- Step 8: If not converged, go back to Step 4.

3. Simulations

3.1. Experiments on artificial data. To verify the validity of our algorithms, extensive computer simulations are carried out. The performance of algorithms to estimate the desired signal is measured by performance index (PI), which is defined as follows

$$\mathrm{PI} = \sum_{j=1}^{n} \frac{|\, p_j \,|}{\max_k |\, p_k \,|} - 1, \ k = 1, \ldots, n \quad (15)$$

where p_j denotes the j element of the global vector $\mathbf{p} = \mathbf{w}^T\mathbf{V}\mathbf{A}$. PI is zero when the desired signals are perfectly extracted. Besides, the accuracy of the recovered source signals compared to the sources is expressed using the signal-to-noise ratio (SNR) in dB given by

$$\mathrm{SNR} = 10\log_{10}(s^2/\mathrm{MSE}), \tag{16}$$

where s^2 denotes the variance of the source signal, MSE denotes the mean square error between the original signal and the recovered signal. The higher SNR is, the better performance is. We generated three zero-mean and unit-variance

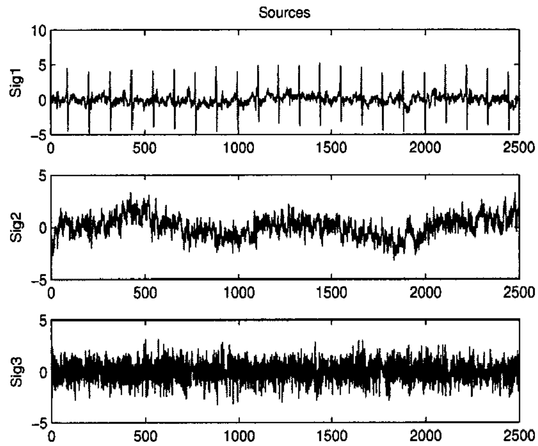

Fig.1 Three artificial signals. (Sigl): FECG; (Sig2): Breathing artifact; (Sig3): One Gaussian signal.

source signals (2500 samples), shown in Fig.1. From the top down, they are, a FECG whose sampling period is 112 (i.e. $\tau = 112$), a breathing artifact and a Gaussian signal, respectively. The observed signals are generated by a 3×3 mixing matrix (randomly chosen), which is given as follows:

$$\mathbf{A} = \begin{pmatrix} 0.61164 & -0.217 & -1.5127 \\ -1.3606 & -1.5758 & -1.0967 \\ 0.71325 & -0.81149 & -0.76034 \end{pmatrix}. \tag{17}$$

Fig.2 provides the mixed signals.

We ran BC algorithm [12], algorithm in [11] and our two-stage algorithm simultaneously. In these simulations, the time delay τ is all chosen to be 112 and parameter P = 1 in algorithm in [11]. In the ICA-R stage of the latter two algorithms the learning rate η is set to be 5×10^{-5}, $\rho = 0.01$, $\gamma = 0.1$ and the multiplier μ is initialized as 0.05. In this experimentsthe SNRs of the extracted signals are 10.2777(dB) (BC algorithm), 16.0793(dB) (Algorithm in [11]) and 17.0385(dB) (our algorithm) respectively. The superiority of our algorithm over the BC algorithm and Algorithm in [11] is straightforward. Fig.3 presents the extracted FECGs by three algorithms.

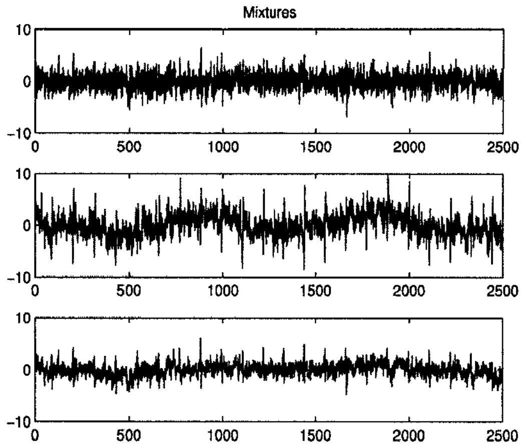

Fig.2 The mixtures by three artificial signals.

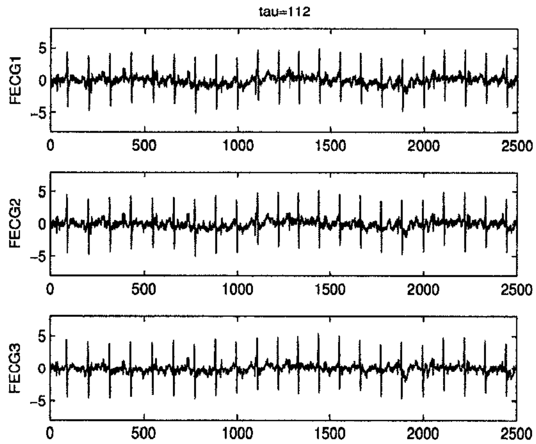

Fig.3 Comparisons of extracted FECGs at $\tau = 112$. FECG1 is the extracted by BC algorithm; FECG2 is the extracted by Algorithm in [11]; FECG3 is the extracted by our algorithm.

In order to compare the performance of separations, we tested three algorithms using the performance index (15) when τ equals 112. The performance was estimated as the average PI values of 100 independent trials. At every trial, three algorithms were run with 100 iterations respectively, which seemed to be always enough for convergence. Here $\mathbf{A}$ and $\mathbf{W}$ were initialized randomly. Fig.4 provided comparison results.

It can be seen that our algorithm has more stable convergence than other algorithms. Moreover, the average PIs are 0.7371, 0.68181 and 0.31859 for BC algorithm, Algorithm in [11] and our algorithm respectively, from which we could

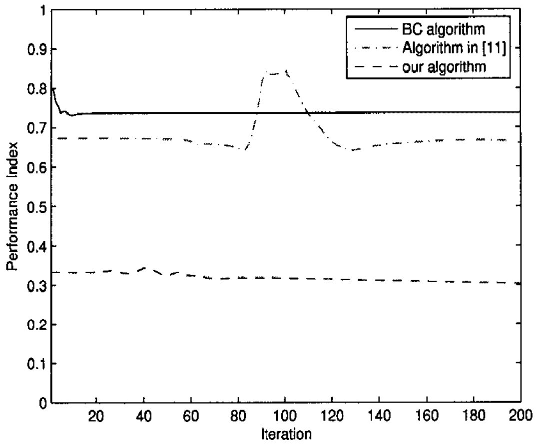

Fig.4 Averaged PIs over 100 independent runs to three algorithms at $\tau = 112$.

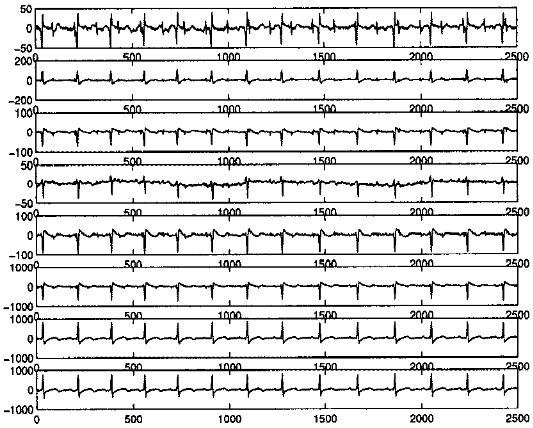

Fig.5 The 8-channel of ECG recording obtained from a pregnant woman.

see that the performance of our algorithm is superior to the others, especially to BC algorithm.

3.2. Experiments on the real-world data. To check the validity of the proposed algorithm, we have performed experiments on real-world ECG data which is distributed by De Moor [19]. This data is a famous ECG measured from a pregnant woman (in Fig.5). The ECG measurements are recorded over 10s and sampled at 250Hz (although in De Moor's homepage he claims the sampling frequency is 500Hz, Barros et al. [12] assure it is 250Hz). By using prior information about FECG frequency, we can estimate

the optimal time delay $\tau = 112$. Note that the parameter P $= 1$ in algorithm in [11] and in the ICA-R stage of algorithm in [11] and our algorithm, the learning rate η is set to be 5×10^{-4}, $\rho = 0.01$, $\gamma = 0.1$ and the multiplier μ is initialized as 0.05. Three extracted FECGs are shown in Fig.6. From the figure we can see that FECG3 is the clearest because of removing the respiratory noise completely, which is contained in FECG1 and FECG2 distinctly.

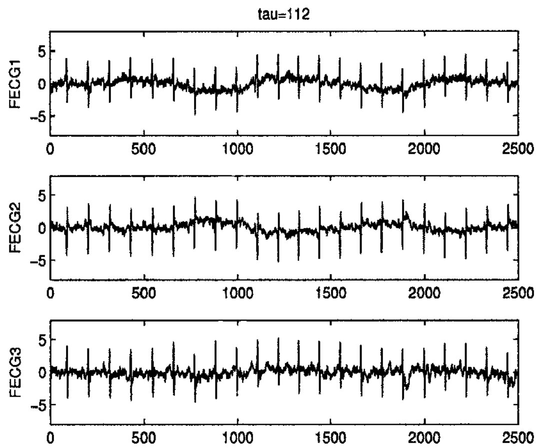

Fig.6 Comparisons of extracted FECGs at $\tau = 112$. FECG1 is the extracted by BC algorithm; FECG2 is the extracted by Algorithm in [11]; FECG3 is the extracted by our algorithm.

4. Conclusions

In the proposed algorithm, the simple but famous algorithm– BC algorithm is used as the first stage of the blind extraction of the temporally correlated sources. It must be pointed out that BC algorithm in all the cases can extract the desired source as long as they are decorrelated and show a temporal structure. However, this method only carries out the constrained minimization of the mean squared error, which dose not describe the probability distribution of the innovation of the signals well. In order to overcome these drawbacks, we added ICA-R base approach as the second stage for further source extraction in this work, that is, at the first stage, BC algorithm is applied to roughly extract the desired signal from their linear mixtures, then, the signal, which was extracted in above stage as the reference signal of ICA-R is further used to extract the desired source as cleanly as possible. Simulations on artificial data and real-world FECG data have shown the better performance of the proposed algorithm. Meanwhile, some comparisons with two existing algorithms, such as BC algorithm and algorithm in [11], further confirmed the validity and usefulness of the proposed algorithm.

Moreover, the constrained ICA (or ICA-R) algorithm has been widely applied to many applications recently. But a crucial problem to the algorithm is how to design a reference signal in advance, which should be closely related to the desired source signal. If there is no enough a priori information about it, the reference signal is difficult to design. With some detailed discussions on the algorithm of ICA-R, we proposes this two-stage source extraction approach to reliably find suitable reference signals for desired source signals. As pointed by literature [11], this kind of methods do not need to know the occurrence time of such responses and their shapes, which is the other advantage of the proposed method. In fact it only requires that the desired source signals are temporally correlated signals, which is satisfied in most cases. Therefore the proposed method is convenient when the desired source signals are very weak or extra a priori information (e.g. the input stimuli) is not available. Although the proposed approach needs to first estimate the time delays of the desired signals, it is not a difficult problem since there are numerous algorithms to estimate such time delays. Therefore, the proposed method is easier and more reliable in practical application.

Acknowledgments

This work is supported by Natural Science Foundation of China under grant No. 10571018, No. 60605002, No. 70431001. The authors would like to thank the referees and the editorial board for their insightful comments and suggestions.

REFERENCES

1. J.F.Cardoso, A. Souloumiac *Blind beamforming for non-Gaussian signals*, in Proc. Inst. Electr. Eng. Radar and Signal Processing, **6** (1993), 362–370.
2. J. Karhunen, A. Hyväinen, R. Vigario, J. Hurri, and E. Oja, *Applications of neural blind separation to signal and image processing*, IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'97), Munich, Germany, 1997.
3. S. Makeig, T. Jung, A. Bell, D. Ghahremani, T. Sejnowski, *lind separation of auditory event-related brain responses into independent components*, Proc. Nat. Acad. Sci., **94**(1997), 10979–10984.
4. A. D. Back, *A first application of independent component analysis to extracting structure from stock returns*, Int. J. Neural Syst., Vol. **8**, No. **4**, (1997), 473–484.
5. J. H. Lee, H. J. Jung, T. W. Lee and S. Y. Lee, *Speech feature extraction using independent component analysis*, in Proc. ICASSP,**2** (1998), 1249–1252.
6. P. Comon, *Independent component analysis: A new concept ?*, Signal Processing, **36** (1994), 287–314.
7. A. Hyvärinen, J. Karhunen and E. Oja, *Independent Component Analysis*, Wiley, New York, 2001.
8. W. Lu and J. C. Rajapakse, *ICA with reference*, Proceeding the 3rd International Conference on Independent Component Analysis and Blind Source Separation (ICA2001), (2001) 120–125.
9. W. Lu, and J. C. Rajapakse, *ICA with reference*, Neurocomputing, Vol. **69** (2006), 2244–2257.
10. W. Lu, and J. C. Rajapakse, *Approach and applications of constrained ICA*, IEEE Trans. on Neural Networks, Vol. **16**, No. **1** (2005), 203–212.
11. Z. L. Zhang, *Morphologically constrained ICA for extraction weak temporally correlated signals*, Neurocomputing, **DOI:**10.1016/j.neucom.2007.04.004.
12. A. K. Barros and A. Cichocki, *Extraction of specific signals with temporal structure*, Neural Computation, Vol. **13**, (2001), 1995–2003.
13. Z. Shi and C. Zhang, *Semi-blind source extraction for fetal electrocardiogram extraction by combining non-Gaussianity and time-correclation*, Neurocomputing, Vol. **70**, (2007), 1574–1581.

14. Z. L. Zhang and Y. Zhang, *Robust extraction of specific signals with temporal structure*, Neurocomputing, Vol. **69**, No. **7-9**(2006), 888–893.
15. A. Belouchrani, K. Meraim, J. F.Cardoso and E. Moulines, *A blind source separation technique based on second order statistics*, IEEE Trans. on Signal Processing, Vol. **45** (1997), 434–444.
16. S. Amari, *ICA of temporally correlated signals learning algorithm*,In Proc. ICA'99, Aussois, France (1999), 13-18.
17. Hyvärinen and E. Oja, *A fast fixed-point algorithm for independent component analysis*, Neural computation, Vol. **9** No. **7**(1997),1483–1492.
18. Q. H. Lin, Y. R. Zheng, F. L. Yin, H. L. Liang and V. D. Calhoun, *A fast algorithm for one-unit ICA-R*, Information Sciences, Vol. **177**(2007),1265–1275.
19. D. De Moor(Ed.), *Daisy: database for the identification of systems, available online at*: http://www.esat.kuleuven.ac.be/sista/daisy.

Hongjuan Zhang received her B.S. degree in Mathematics from Ludong University, China, in 2003, M.S. degree from Department of Applied Mathematics of Dalian University of technology in 2005. She is now a docterate student in Department of Applied Mathematics of Dalian University of Technology. Her research interests cover blind signal processing and its applications.

Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, P. R. China
e-mail: `zhhj2108@163.com`

Zhenwei Shi received his B.S. and M.S. degrees in Mathematics from the Inner Mongolia University, Huhhot, China, in 1999 and 2002, respectively. He received his Ph.D. degree in Mathematics from the Dalian University of Technology, Dalian, China, in 2005. He was a Postdoctoral Researcher in the Department of Automation, Tsinghua University from 2005 to 2007. Now he is currently a lectuer of the Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics. His current research interests include blind signal processing, pattern recognition, machine learning and neuroinformatics.

Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing 100083, P. R. China.
e-mail: `shizhenwei@mail.tsinghua.edu.cn`

Chonghui Guo received the B.S. degree in Mathematics from Liaoning University, China, in 1995, M.S. degree in Operational Research and Control Theory, and Ph.D. degree in Management Science and Engineering from Dalian University of Technology, China, in 2002. He was a postdoctoral research fellow in the Department of Computer Science, Tsinghua University from 2002 to 2004. Now he is an associate professor of the Institute of Systems Engineering, Dalian University of Technology. His current research interests include machine learning, data mining, pattern recognition, and systems optimization.

Institute of Systems Engineering, Dalian University of Technology, Dalian 116024, P. R. China.
e-mail: `guochonghui@tsinghua.org.cn`

Enmin Feng is a professor and Ph.D. advisor in Department of Applied Mathematics of Dalian University of Technology. His research interests focus on control theory and optimization.
Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, P. R. China.
e-mail: `emfeng@dlut.edu.cn`