Two-Stage Temporally Correlated Source Extraction Algorithm with Its Application in Extraction of Event-Related Potentials

Zhi-Lin Zhang[1,2], Liqing Zhang[1], Xiu-Ling Wu[1], Jie Li[1], and Qibin Zhao[1]

[1] Department of Computer Science and Engineering,
Shanghai Jiao Tong University, Shanghai 200240, China
[2] School of Computer Science and Engineering,
University of Electronic Science and Technology of China,
Chengdu 610054, China
zlzhang@uestc.edu.cn, zhang-lq@cs.sjtu.edu.cn

Abstract. To extract source signals with certain temporal structures, such as periodicity, we propose a two-stage extraction algorithm. Its first stage uses the autocorrelation property of the desired source signal, and the second stage exploits the independence assumption. The algorithm is suitable to extract periodic or quasi-periodic source signals, without requiring that they have distinct periods. It outperforms many existing algorithms in many aspects, confirmed by simulations. Finally, we use the proposed algorithm to extract the components of visual event-related potentials evoked by three geometrical figure stimuli, and the classification accuracy based on the extracted components achieves 93.2%.

1 Introduction

It is known that blind source extraction (BSE) algorithms are suitable for extracting a few of temporally correlated source signals from large numbers of sensor signals, say recordings of 128 EEG sensors [1]. In practice they require certain additional *a priori* information of the desired source signals. Thus they generally are implemented in a semi-blind way [2,3,5,6,7].

Among the extraction algorithms there are two famous algorithms, i.e. the cICA algorithm [5] and the FICAR algorithm [6], both of which need to design a so-called reference signal that is closely related to the desired underlying source signal. That is to say, the phase and the morphology of the reference must be matched to that of the desired signal to great extent, or the occurrence time of each impulse of the reference signal is consistent with that of the desired signal [8]. However, in some applications it is difficult to design such a reference, especially when the morphology and the phase of the desired source signals are not expected [3].

Based on our previous primary work [2,3], in this paper we propose a Temporally Correlated signal Extraction algorithm (TCExt algorithm), which does not need the reference, unlike the cICA algorithm and the FICAR algorithm.

I. King et al. (Eds.): ICONIP 2006, Part II, LNCS 4233, pp. 523–532, 2006.

Computer simulations on artificially generated data and experiments on the extraction of event-related potentials show its many advantages.

2 Problem Statement

Suppose that the unknown source signals $\mathbf{s}(k) = [s_1(k), \cdots, s_n(k)]^T$ are mutually statistically independent with zero mean and unit variance, holding the basic simultaneous mixture ICA model [1]. Without lose of generality, we further assume s_1 is the desired temporally correlated source signal, satisfying the following relationship:

$$\begin{cases} E\{s_1(k)s_1(k-\tau^*)\} > 0 \\ E\{s_j(k)s_j(k-\tau^*)\} = 0 \quad \forall j \neq 1 \end{cases} \tag{1}$$

where s_j are other source signals, and τ^* is the optimal lag defined below:

Definition 1. *The non-zero τ^* is called the optimal lag, if the delayed autocorrelation at τ^* of the desired source signal s_1 is non-zero, while the delayed autocorrelation at τ^* of other source signals is zero. Here all of the source signals are supposed to be mutually independent.*

In addition, we give the definition of the optimal weight vector as follows:

Definition 2. *The column vector $\boldsymbol{w}^*$ is called the optimal weight vector of the desired source signal $\boldsymbol{s}_1$, if the following relationship holds:*

$$(\boldsymbol{w}^*)^T \boldsymbol{V}\boldsymbol{A}\boldsymbol{s} = c\boldsymbol{s}_1, \tag{2}$$

where c is a non-zero constant, $\boldsymbol{V}$ is a whitening matrix, and $\boldsymbol{A}$ is a mixing matrix.

3 Framework of the Proposed Algorithm

Based on the assumptions in the previous section, we have proposed a two-stage extraction algorithm framework [3], shown in Fig.1. The first stage is called the capture stage. In this stage, the algorithm coarsely extracts the desired source signal by using correlation information. At the end of the stage, we obtain the weight vector $\hat{\mathbf{w}}$. But it can be shown that due to some practical issues [2,3] $\hat{\mathbf{w}}$ is only close to the optimal weight vector $\mathbf{w}^*$. Therefore the captured source signal $\hat{y} = \hat{\mathbf{w}}^T\mathbf{x}$ is still mixed by the "cross-talk noise".

Next, in the second stage, we exploit the independence assumption and use the output of the first stage, i.e. $\hat{\mathbf{w}}$. At the end of this stage, we obtain a sub-optimal solution $\bar{\mathbf{w}}$ [1], which is much closer to $\mathbf{w}^*$ than $\hat{\mathbf{w}}$ is. Thus we finally obtain the desired source signal $\bar{y} = \bar{\mathbf{w}}^T\mathbf{x}$, which is almost not mixed by the "cross-talk noise".

In the framework we will propose an improved extraction algorithm with higher performance, even if the desired source signals have the same period or are near Gaussian.

[1] Note that in practice we almost cannot obtain the optimal solution $\mathbf{w}^*$.

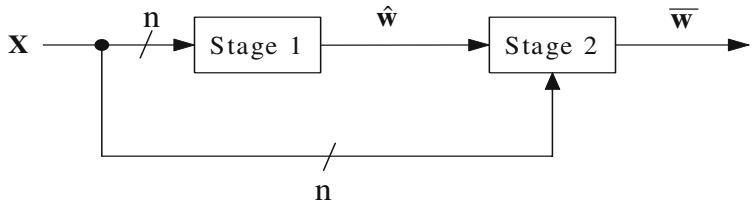


Fig. 1. The framework of the proposed algorithm

3.1 Finding Lags

In practice we cannot find the optimal lag, and we can only find lags that satisfy $E\{s_1(k)s_1(k-\tau_i)\} > E\{s_j(k)s_j(k-\tau_i)\}$, $\forall j \neq 1$, $i = 1, \cdots, P$. Thus, due to performance consideration [2] we suggest to select several suitable lags that correspond to the time structure of the desire source signal, instead of selecting only one lag. For example, for a periodic signal we select the lags corresponding to its fundamental period and multiple periods. The use of several lags, instead of only one lag, can improve the extraction performance, as shown in [3].

There are many methods for finding these lags or the temporal structures [7]. For example, the cepstrally transformed discrete cosine transform [11] can be used to detect the periods of source signals, even if the strengths of signals differ by about 60 dB. In addition, in some applications, such as biomedical signal processing, the lags are often readily available [7,8].

3.2 The First Stage: Coarse Recovery

After choosing suitable lags $\tau_i (i = 1, \cdots, P)$ and whitening the original observations, the first stage employs our previously proposed algorithm [2] to obtain the weight vector $\hat{\mathbf{w}}$:

$$\hat{\mathbf{w}} = EIG\Big(\sum_{i=1}^{P}\Big(\mathbf{R}_{\mathbf{Z}}(\tau_i) + \mathbf{R}_{\mathbf{Z}}(\tau_i)^T\Big)\Big), \tag{3}$$

where $\mathbf{R}_{\mathbf{Z}}(\tau_i) = E\{\mathbf{z}(k)\mathbf{z}(k-\tau_i)^T\}$, $\mathbf{z}(k)$ are the whitened observations, and $EIG(\mathbf{Q})$ is the operator that calculates the normalized eigenvector corresponding to the maximal eigenvalue of the matrix $\mathbf{Q}$.

If the desired signal is periodic, then the algorithm (3) can be rewritten as

$$\hat{\mathbf{w}} = EIG\Big(\sum_{i=1}^{P}\Big(\mathbf{R}_{\mathbf{Z}}(i\tau) + \mathbf{R}_{\mathbf{Z}}(i\tau)^T\Big)\Big), \tag{4}$$

where τ is the fundamental period of the desired source signal. If several desired source signals have the same period, they still can be extracted under some weak conditions, confirmed by the following theorem.

Theorem 1. *Suppose there are q source signals $(s_1, \cdots, s_q)$ that are mutually uncorrelated and have the same period N, and also suppose their autocorrelations satisfy $E\{s_i(k)s_i(k-N)\} \neq E\{s_j(k)s_j(k-N)\}$, $\forall i \neq j$ and $1 \leq i, j \leq q$. Without lose of generality, further suppose $r_1 > \cdots > r_q$, where $r_i = E\{s_i(k)s_i(k-N)\}$. Then the i-th source signal can be perfectly extracted by the weight vector $\boldsymbol{w}_i$ that is the normalized eigenvector corresponding to the i-th largest eigenvalue of $E\{\boldsymbol{z}(k)\boldsymbol{z}(k-N)^T\}$.*

Proof: Since $\mathbf{w}_i$ is the normalized eigenvector corresponding to the i-th largest eigenvalue of $E\{\mathbf{z}(k)\mathbf{z}(k-N)^T\}$, we have $E\{\mathbf{z}(k)\mathbf{z}(k-N)^T\}\mathbf{w}_i = \lambda_i\mathbf{w}_i$, $i = 1, \cdots, q$, where λ_i is the i-th largest eigenvalue. In other words, $\mathbf{VA}E\{\mathbf{s}(k)\mathbf{s}(k-N)^T\}\mathbf{A}^T\mathbf{V}^T\mathbf{w}_i = \lambda_i\mathbf{w}_i$. Since $\mathbf{VA}$ is an orthogonal matrix, then $E\{\mathbf{s}(k)\mathbf{s}(k-N)^T\}(\mathbf{A}^T\mathbf{V}^T\mathbf{w}_i) = \lambda_i(\mathbf{A}^T\mathbf{V}^T\mathbf{w}_i)$, indicating that $(\mathbf{A}^T\mathbf{V}^T\mathbf{w}_i)$ is the normalized eigenvector corresponding to the eigenvalue λ_i of $E\{\mathbf{s}(k)\mathbf{s}(k-N)^T\}$. Due to the distinction of the eigenvalues of $E\{\mathbf{s}(k)\mathbf{s}(k-N)^T\}$, we can deduce that λ_i is its i-th largest eigenvalue, i.e., $\lambda_i = r_i$. According to the assumptions and the previous development, $E\{\mathbf{s}(k)\mathbf{s}(k-N)^T\}$ is a diagonal matrix, and thus we have $(\mathbf{A}^T\mathbf{V}^T\mathbf{w}_i) = \mathbf{e}_i$, whose the i-th element is one while other elements are zero. On the other hand, we have $y = \mathbf{w}_i^T\mathbf{z} = \mathbf{w}_i^T\mathbf{VAs} = \mathbf{e}_i^T\mathbf{s}$, implying the i-th source signal is perfectly extracted. ■

The algorithm (3) has many advantages (see [2,3] for details). However, although it can achieve good extraction quality, it can be shown that the algorithm is insufficient to perfectly recover the desired source signal, and that the solution $\widehat{\mathbf{w}}$ in this stage is just close to the optimal weight vector $\mathbf{w}^*$ [3]. Thus, to make the solution $\widehat{\mathbf{w}}$ further closer to $\mathbf{w}^*$, in the second stage we derive a higher-order statistics based algorithm.

3.3 The Second Stage: Fine Extraction

Under the constraint $\|\mathbf{w}\| = 1$, the maximum likelihood criteria for extracting one source signal is given by

$$\begin{cases} \min & l(\mathbf{w}) = -E\{\log p(\mathbf{w}^T\mathbf{z}(k))\} \\ s.t. & \|\mathbf{w}\| = 1 \end{cases} \tag{5}$$

where $p(\cdot)$ denotes the probability density function (pdf) of the desired source signal. Note that minimizing (5) only leads to one source signal, but not necessarily the desired source signal s_1. However, if we use the $\widehat{\mathbf{w}}$ from the first stage as the initial value, then we can necessarily obtain the s_1.

By the Newton optimization method, we obtain the following algorithm for extracting the desired source signal s_1:

$$\begin{cases} \mathbf{w}^+ = \mathbf{w} - \mu E\big\{f\big(\mathbf{w}^T\mathbf{z}\big)\mathbf{z}\big\}/E\big\{f'\big(\mathbf{w}^T\mathbf{z}\big)\big\} \\ \mathbf{w} = \mathbf{w}^+/\|\mathbf{w}^+\|, \end{cases} \tag{6}$$

with the initial value $\mathbf{w}(0) = \widehat{\mathbf{w}}$. μ is a step-size that may change with the iteration count. In particular, it is often a good strategy to start with $\mu = 1$. $f(\cdot)$ is a nonlinearity, given by $f(\cdot) = -(\log p(\cdot))' = -p(\cdot)'/p(\cdot)$.

In general, the pdf p is unknown and should be estimated. We present a density model that combines the t-distribution density model, the generalized Gaussian distribution density model and the Pearson system model. Our motivation is that the nonlinearity derived from the t-distribution is more robust to the outliers and avoids the stability problem [9], and that the nonlinearity derived from Pearson system can achieve good performance when the desired source signals are skewed and/or near Gaussian [10].

We use the t-distribution [9] to model the super-Gaussian distribution. The derived nonlinearity is

$$f(y) = -p(y)'/p(y) = \frac{(1+\beta)y}{y^2 + \frac{\beta}{\lambda^2}}. \tag{7}$$

where parameters β and λ^2 can be calculated by $\lambda^2 = \beta\Gamma(\frac{\beta-2}{2})/(2m_2\Gamma(\frac{\beta}{2}))$ and $\kappa_t = \frac{m_4}{m_2^2} - 3 = 3\Gamma(\frac{\beta-4}{2})\Gamma(\frac{\beta}{2})/(\Gamma(\frac{\beta-2}{2})^2) - 3$, where m_2 and m_4 are respectively the second-order moment and the fourth-order moment of the distribution. It is clear to see that the function $f(y)$ approaches to zero when the value of y abruptly increases, implying that it is robust to the undue influence of outliers.

To extract the sub-Gaussian source signal, we use the well-known fixed nonlinearity

$$f(y) = y^3, \tag{8}$$

which belongs to the generalized Gaussian density model.

In some applications the desired source signals are skewed, such as the components of the ECG with absolute skewness ranging from 1 to 10. In addition, in some cases the desired source signals are close to Gaussian. Due to these facts, we use the Pearson system to derive a family nonlinearities that are more suitable to extract the skewed and/or near Gaussian signals than the ones derived from the t-distribution and the generalized Gaussian distribution.

The nonlinearity derived from the Pearson system is given by [10]

$$f(y) = -\frac{p_p'(y)}{p_p(y)} = -\frac{(y-a)}{b_0 + b_1 y + b_2 y^2}, \tag{9}$$

where a, b_0, b_1 and b_2 are the parameters of the distribution, calculated by $a = b_1 = -m_3(m_4 + 3m_2^2)/C, b_0 = -m_2(4m_2m_4 - 3m_3^2)/C, b_2 = -(2m_2m_4 - 3m_3^2 - 6m_2^3/C$, where $C = 10m_4m_2 - 12m_3^2 - 18m_2^3$. Note that this type of nonlinearity is also robust to the outliers, just as the nonlinearity given in (7).

Now we have presented three types of nonlinearities for three types of signals. According to the estimated moments, the algorithm (6) adopts suitable nonlinearities. A procedure for the adaptive nonlinearity selection using the sample moments may be given as follows.

Repeat until convergence:

1. Calculate the second, third and fourth sample moments $\hat{m}_2, \hat{m}_3, \hat{m}_4$ for current data $\mathbf{y}(l) = \mathbf{w}^T(l)\mathbf{z}$, where l represents the iteration number.

2. According to the estimated moments, select the nonlinearity as follows:
   - If $\hat{m}_4 > \hat{m}_3^2 + 4.5$, then calculate the nonlinearity (7);
   - if $\hat{m}_4 < 2.5$, then use the nonlinearity (8);
   - if $2.5 \leq \hat{m}_4 \leq \hat{m}_3^2 + 4.5$, then calculate the nonlinearity (9).
3. Calculate the weight vector $\mathbf{w}(l+1)$ using the algorithm (6).

4 Simulations

In the first simulation, we generated seven zero-mean and unit-variance source signals, shown in Fig.2. Each signal had 2000 samples, and its statistics property is shown in Table 1. These signals were randomly mixed and whitened. Our goal was to extract the temporally correlated source signals s_1, s_2, s_3, s_6 and s_7 one by one. After estimated the suitable lags for extracting each desired signal, we employed our proposed two-stage algorithm (TCExt). To make comparisons, we also employed the akExt algorithm [4], the cICA algorithm [5], the FICAR algorithm [6], the SOS algorithm [7], the CPursuit algorithm [13], the SOBI algorithm [12] and the pBSS algorithm [14] on the whitened signals. Note that, in this simulation both the cICA and the FICAR could not extract the source signals due to the difficulty to design the reference signals, but in order to compare the extraction quality, we designed suitable reference signals in advance according to the waveforms of the source signals. To compare the extraction performance we used the following performance index

$$PI = -10E\{lg(s(k) - \tilde{s}(k))^2\}\ (dB) \tag{10}$$

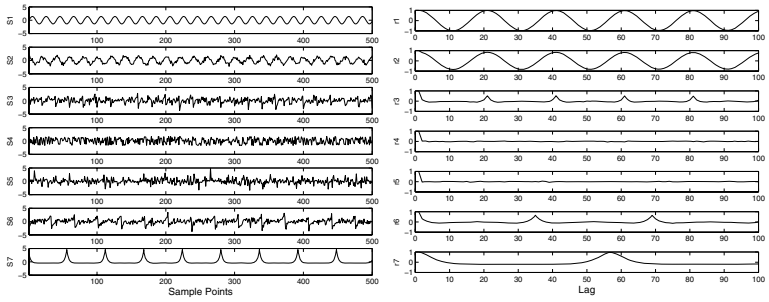
where $s(k)$ is the desired source signal, and $\tilde{s}(k)$ is the extracted signal (both of them are normalized to be zero-mean and unit-variance). The higher PI is, the better the performance is. The averaged performance indexes over 100 independent trials of each algorithm are shown in Table 2, from which we can see that the proposed algorithm generally has better performance than the other algorithms.

Table 1. The properties of the source signals in Fig.2. 'p' denotes the corresponding signal was strictly periodic; 'c' denotes temporally correlated but not strictly periodic; 'n' denotes random noise without any time structure.

source signal	s_1	s_2	s_3	s_4	s_5	s_6	s_7
periodicity	p	c	c	n	n	c	p
kurtosis	-1.5	-1.0	0.7	-1.2	2.8	0.4	7.5

In the next experiment we applied our algorithm to extract potentials evoked by three types of geometrical figures stimuli, and our objective is to classify each type of figures according to the extracted visual evoked potentials (VEPs).

One right-handed subject, aged 21, volunteered to participate in the present study. The subject was healthy both in psychological and neurological, and had



(a) (b)

Fig. 2. Source signals. (a) A segment of the seven source signals. Note that s_1, s_2, s_3 had the same period, but differ in autocorrelations. (b) The corresponding autocorrelation functions of the source signals of (a).

Table 2. The averaged performance index of each algorithm over 100 independent trials. '-' indicates that PI was less than 5 dB or the algorithm could not converge in all the trials. 'akExt(1)' indicates that the value of the parameter τ of the akExt algorithm was equal to the fundamental period of the desired signal; 'akExt(2)' indicates the value of τ was equal to the doubled fundamental period. The same with 'SOS(1)' and 'SOS(2)'.

	TCExt	akExt(1)	akExt(2)	SOS(1)	SOS(2)	cICA	FICAR	SOBI	CPursuit	pBSS
PI of s_1	48.0	17.6	15.9	41.2	37.3	20.6	13.2	8.9	48.3	7.9
PI of s_2	26.9	-	-	-	-	-	-	8.7	27.8	10.2
PI of s_3	12.2	-	-	-	-	-	-	14.6	10.7	6.7
PI of s_6	22.0	34.9	-	-	-	-	20.9	32.2	21.4	-
PI of s_7	57.3	42.0	37.7	45.6	41.4	39.4	34.3	35.7	36.2	20.1

a normal vision. He was seated in a comfort and fixed chair, 0.7m far from the screen of monitor, in a sound and light attenuated RF shielded room.

Three types of geometrical figures(five different-size units for each type) were presented to the subject, i.e. the circle, the square, and the triangle figures. In each trial, a type of geometrical figures, say the circle figure, appeared according to the sequence illustrated in Fig.3. In order to reduce the subject's expectation, each trial showed a type of figures randomly (each type was showed in identical probability). EEG signals were recorded (see Fig.4), sampled at 1000 Hz (thus each trial had 3174 samples) and bandpass filtered between 0.1 Hz and 200 Hz, by a 64-channel EEG system (SynAmps2, Neuroscan, at our Lab for Perception Computing at Shanghai Jiao Tong University, China).

From the original EEG data, we used the proposed algorithm to extract three VEP components by the following procedure. Suppose we had extracted q ($q < 3$) components of VEPs, which corresponded to the first q largest eigenvalues among

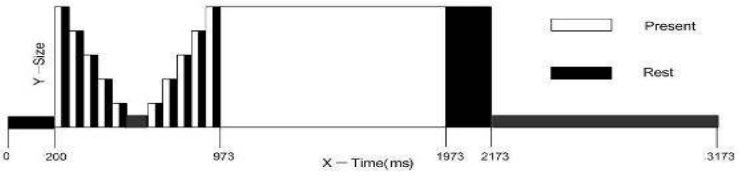
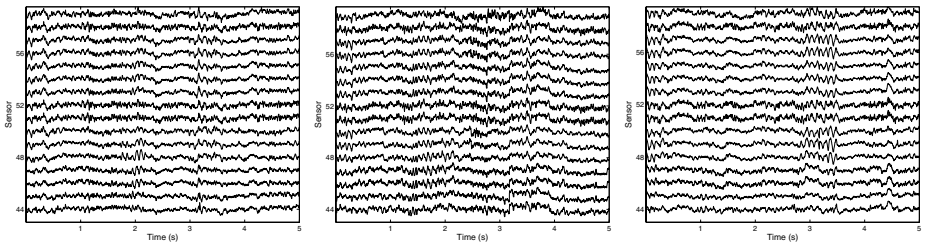


Fig. 3. A stimuli sequence in one trial. The X axis showed the lasting time of the presence or the non-presence of the figure stimuli. The Y axis showed the relative size of the geometrical figure.



(a) (b) (c)

Fig. 4. Five-second segments of the original EEGs recorded by sensors (from Channel 44 to Channel 58). (a) EEGs of the Circle Class after the epoch-finding. (b) EEGs of the Square Class after the epoch-finding. (c) EEGs of the Triangle Class after the epoch-finding.

all of the eigenvalues of $\sum_{i=1}^{P}(\mathbf{R_Z}(i\tau) + \mathbf{R_Z}(i\tau)^T)$, and we extracted the next VEP component:

1. Applied the proposed algorithm to extract the component that corresponded to the $(q+1)$-th largest eigenvalue of $\sum_{i=1}^{P}(\mathbf{R_Z}(i\tau) + \mathbf{R_Z}(i\tau)^T)$;
2. To ensure the extracted component was not the component of artifacts, we calculated its autocorrelation;
3. Since the components of VEPs exhibited time-locked activation to task-related events and those of artifacts did not, the autocorrelations of VEP components had peaks locating at lag 3174, lag 6348, lag 9522, et al., while those of artifacts components did not. By this method, if we found the extracted component was not a component of VEPs, then we discarded it and went back to step 1. It should be noticed that there are many approaches, e.g. [15], that can help us further distinguishing artifacts from evoked potentials.

This loop continued until we extracted three VEP components; each component consisted of 120 trials. Then the epoch-finding was conducted using the Neuroscan toolbox so that the trials corresponding to the same type of figures were gathered into a class (Fig.5 shows the average result of the trials belonging to the same extracted VEP component in each class). Thereby we had three classes, namely the Circle Class, the Square Class and the Triangle Class.

We randomly selected 60 trials of each extracted component as the training set and the remained trials of each extracted component as the test set. For classifi-

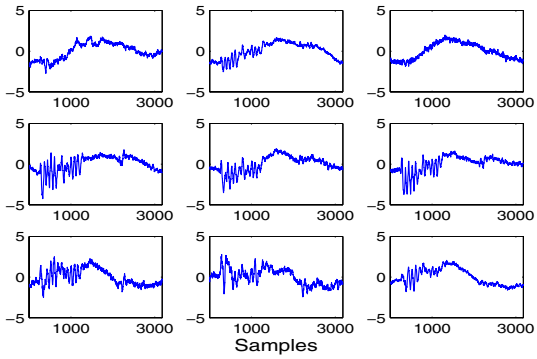


Fig. 5. The averaged trials. The signal in i-th row and j-th column is the average result of the trails belonging to the i-th extracted VEP component of the j-th class $(i, j = 1, 2, 3)$.

cation, the feature vectors of each class were constructed as follows: we selected some features from each trial of the first, the second and the third extracted components, respectively, and these features were concatenated orderly to form a feature vector. Here we selected 30 features from the frequency components of each trial according to the MIFS-U algorithm [16], an effective feature selection method based on mutual information. Finally, we used the multi-category SVM as the classifier, and the classification accuracy reached 93.2%.

5 Conclusions

We propose a two-stage algorithm for extracting source signals that satisfy some given temporal structure. The algorithm is suitable to extract the periodic or quasi-periodic source signals, even if the desired source signals have the same period (but they should have different autocorrelation structure). Compared with many widely-used extraction algorithms, the algorithm has better performance, verified by simulations and experiments.

Acknowledgements

The work was supported by the National Basic Research Program of China (Grant No. 2005CB724301) and National Natural Science Foundation of China (Grant No.60375015).

References

1. Cichocki, A., Amari, S.: Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. John Wiley & Sons, New York (2002)
2. Zhang, Z.-L., Yi, Z.: Robust Extraction of Specific Signals with Temporal Structure. Neurocomputing **69 (7-9)** (2006) 888-893

3. Zhang, Z.-L., Zhang, L.: A Two-Stage Based Approach for Extracting Periodic Signals. Proc. of the 6th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2006). LNCS **3889** (2006) 303-310
4. Zhang, Z.-L., Yi, Z.: Extraction of Temporally Correlated Sources with Its Application to Non-invasive Fetal Electrocardiogram Extraction. Neurocomputing **69 (7-9)** (2006) 900-904
5. Lu, W., Rajapakse, J.C.: Approach and Applications of Constrained ICA. IEEE Trans. Neural Networks **16 (1)** (2005) 203-212
6. Barros, A.K., Vigário, R., Jousmäki, V., Ohnishi, N.: Extraction of Event-related Signals from Multichannel Bioelectrical Measurements. IEEE Trans. Biomedical Engineering **47 (5)** (2000) 583-588
7. Barros, A.K., Cichocki, A.: Extraction of Specific Signals with Temporal Structure. Neural Computation **13 (9)** (2001) 1995-2003
8. James, C.J., Gibson, O.J.: Temporally Constrained ICA: An Application to Artifact Rejection in Electromagnetic Brain Signal Analysis. IEEE Trans. Biomedical Engineering **50 (9)** (2003) 1108-1116
9. Cao, J., Murata, N., Amari, S.-I., Cichocki, A., Takeda, T.: A Robust Approach to Independent Component Analysis of Signals with High-Level Noise Measurements. IEEE Trans. Neural Networks **14 (3)** (2003) 631-645
10. Karvanen, J., Koivunen, V.: Blind Separation Methods Based on Pearson System and Its Extensions. Signal Processing **82** (2002) 663-673
11. Paul, J.S., Reddy, M.R., Kumar, V.J.: A Cepstral Transformation Technique for Dissociation of Wide QRS-Type ECG Signals Using DCT. Signal Processing **75** (1999) 29-39
12. Belouchrani, A., Abed-Meraim, K., Cardoso, J.F., Moulines, E.: A Blind Source Separation Technique Using Second-Order Statistics. IEEE Trans. Signal Processing **45 (2)** (1997) 434-444
13. Hyvärinen, A.: Complexity Pursuit: Combining Nongaussianity and Autocorrelations for Signal Separation. Proc. of the 2th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2000), pp. 175-180
14. Jafari, M.G., Wang, W., Chambers, J.A., et al.: Sequential Blind Source Separation Based Exclusively on Second Order Statistics Developed for a Class of Periodic Signals. IEEE Trans. Signal Processing **54 (3)** (2006) 1028-1040
15. Shoker, L., Sanei, L., Chambers, J.: Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm. IEEE Signal Processing Letters **12 (10)** (2005) 721-724
16. Kwak N., Choi C.-H.: Input Feature Selection for Classification Problems. IEEE Trans. Neural Networks **13 (1)** (2002) 143-159