A Preprocessing Method of EEG Based on EMD-ICA in BCI

Banghua Yang, Liangfei He, Qian Wang, Chunting Song, and Yunyuan Zhang

Department of Automation, School of Mechatronics Engineering and Automation, Shanghai Key Laboratory of Power Station Automation Technology, Shanghai University, Shanghai 200072

yangbanghua@shu.edu.cn, heliangfei126@126.com, lvmaogui123@163.com

Abstract. In order to remove artifacts automatically and effectively from the Electroencephalography (EEG) in Brain Computer Interfaces (BCIs), a new preprocessing algorithm called EMD-ICA (Empirical Mode Decomposition, Independent Component Analysis) is explored. The EMD-ICA method includes the following steps: Firstly, EEG signals from single or multiple channels are decomposed into a series of intrinsic mode functions (IMFs) using EMD. Each IMF can be approximately used as an input channel of the ICA, and these IMFs constitute the input matrix of the ICA. Then, the input matrix is separated into a set of statistics independent components (ICs) by ICA. Furthermore, each of statistics ICs is analyzed by using the method of sample entropy to automatically determine whether it is artifact signal. Finally, the ICs determined as artifacts are eliminated and the remaining ICs are reconstructed. The reconstructed EEG is used in the following feature extraction and classification. To evaluate the effect of the proposed method, common spatial patterns (CSP) and support vector machine (SVM) algorithm are used to extract and classify the EEG data from two datasets. The experimental results show that the proposed method can remove various kinds of artifacts effectively, and improve the recognition accuracy greatly.

Keywords: Brain Computer Interface (BCI), Empirical Mode Decomposition and Independent Component Analysis (EMD-ICA), Electroencephalography (EEG).

1 Introduction

Brain computer interfaces (BCIs) are systems that provide an alternative pathway for their users to transmit information to external world, which has become an assistive tool for neuromuscular disordered people's communication and control [1]. Electroencephalography (EEG) might be the most widely used brain imaging modality for noninvasive BCI, because EEG can capture a fast dynamics of brain information processing at a high temporal resolution. However, it is known that EEG has low spatial resolution and high noise level, which make it challenging to extract

S. Ma et al. (Eds.): LSMS/ICSEE 2014, Part I, CCIS 461, pp. 1–12, 2014.

useful information from EEG signals for BCI applications [2]. EEG pre-processing is to eliminate noise to improve the quality of the EEG. It plays an important role on the BCI research. However, many of the preprocessing techniques require minimal number of channels (>6) [3, 4] to work well. In general, increasing the number of channels may improve the processing effect. But the increase of number of channels used in preprocessing techniques not only extends the processing time but also brings noise form non-related channels. Therefore, it is meaningful to propose a preprocessing method that can remove artifacts effectively using less than six channels or even a single channel.

At present, the commonly used methods for pre-processing contains: (1) Wavelet Transform (WT), (2) Principle Component Analysis (PCA), (3).Independent Component Analysis (ICA). ICA is a blind source separation (BSS) technique that can extract the relevant information buried within noisy signals and allow the separation of measured signals into their fundamental underlying independent component(IC) [5]. ICA has already been quite broadly applied to the analysis of biomedical signals, such as analysis of EEG [6]. However, ICA requires minimal number of channels to work well. Furthermore, ICA needs visual inspection to select components manually for correction [7]. Empirical mode decomposition (EMD) is a kind of self-adapting signal processing method and it is very suitable for dealing with nonlinear and non-stationary signals [8]. EMD can broke complex signals down into a set of IMF (Intrinsic Mode Functions), each IMF can be seen as a sub-channel for the input of ICA. The sub-channels are combined into the input matrix of ICA. EMD decomposes a channel into several sub-channels, which results in more input channels for ICA under the certain number of raw EEG. That is also to say that EMD can improve the performance of ICA in case of reducing the number of raw EEG channels required by ICA.

In this paper, a new EMD-ICA method for removing artifacts in EEG data is proposed, which combines EMD and ICA. The advantage of EMD, compared to WT, is that the EMD is a data driven algorithm. This means that it decomposes a signal in a natural way without prior knowledge about the signal of interest embedded in the data series. The ICA algorithm used in this study is FastICA. FastICA has the advantage of fast convergence and good separation. The EMD is employed to decompose the EEG signal into a series of physically meaningful IMFs. These IMFs are used as input of ICA to extract ICs. Then, sample entropy is used to automatically find out the artifact components among these ICs. Finally, EEG is obtained by reconstructing non-artifactual components. The EMD-ICA method can effectively remove artifacts just using a small number of raw EEG channels.

2 EMD-ICA Methodology

2.1 Introduction of EMD

The EMD is a signal-dependent decomposition. It can decompose a time series into waveforms which are modulated in amplitude and frequency [9]. The iterative extraction of these components is based on the local representation of the signal.

With the EMD, the signal is decomposed into a set of IMFs, which are a kind of complete, adaptive and almost orthogonal representation for the analyzed signals. More precisely, $X_l(t)$ is used to denote the EEG signal from channel l ($l = 1, 2, ...L$), L is the total number of channels). If $X_l(t)$ is input signal, then EMD decomposes the $X_l(t)$ into intrinsic mode functions denoted by $\{d_i(t)\}_{i=1}^{N}$ such that:

$$X_l(t) = \sum_{i=1}^{N} d_i(t) + r(t) \tag{1}$$

Where $r(t)$ denotes the residual monotonic function which reflects the average trend with in the original one signal. In order to get meaningful estimation of instantaneous frequency, the IMF should be designed as close symmetric around the local mean and their number of extrema and zero-crossing must be equal or differ at most by one . The IMFs can be obtained by shifting process, described as:

1 Find all extrema (minimum & maximum) of $X_l(t)$

2 Interpolate (using cubic spline interpolation) between maximums (minimums) to obtain signal upper envelope $e_u(t)$ （a lower envelop $e_l(t)$)

3 Calculate the local mean $m(t) = (e_u(t) + e_l(t)) / 2$

4 Subtract $m(t)$ from $X_l(t)$ to construct oscillating signal $h(t) = X_l(t) - m(t)$

5 If $h(t)$ satisfies all stopping conditions, $d(t) = h(t)$ becomes an IMF; otherwise repeat step1 by setting $X_l(t) = h(t)$

After the shifting, $X_l(t)$ is decomposed into a set of IMFs denoted by $\{d_1(t), d_2(t), \ldots\ldots d_n(t)\}$. n is the number of IMFs. Let $D(t) = [d_1(t), d_2(t), \ldots\ldots d_n(t)]$ and $D(t)$ is used as an input matrix of following ICA.

2.2 Introduction of ICA

ICA is a BSS technique for separating multivariate observed random data into mutually statistically non-Gaussian ICs. ICA can work well without any information about the mixing matrix. The time varying observed signals (mixed signals) are denoted by $D(t) = (d_1(t), d_2(t), \ldots\ldots d_n(t))^T$ and the source signals consisting of ICs by $S(t) = (s_1(t), s_2(t), \ldots\ldots s_m(t))^T$ and therefore

$$D(t) = AS(t) \tag{2}$$

Then there exists a de-mixing matrix W such that

$$S(\mathrm{t}) = WD(t) \tag{3}$$

The object of ICA is to find W. ICA is actually an optimization problem, depending on objective function and optimization algorithm. In this paper, according to negentropy maximum criterion [10], the objective function is defined by

$$C(s) = \sum_{i=1}^{m} J(s_i) \tag{4}$$

Where $s_i = w_i D(t)$, $J(s_i) \approx \rho(E\{G_i(s_i)\} - E\{G_i(v)\})^2$, ρ is a positive constant, $\{G_i(\bullet)\}$ is a non-quadratic function, $E\{\bullet\}$ is a mean function and v is a Gaussian variable having zero mean and unit variance.

FastICA is one of the more popular and referenced ICA techniques which is based on its own unique fast fixed-point iterative algorithm. Using newton iteration method and choosing an initial weight vector W, the basic form of FastICA iteration algorithm is as follows:

$$W = E\{DG(W^T D)\} - E\{G^{'}(W^T D)\} W \tag{5}$$

$$W = W / \|W\| \tag{6}$$

The algorithm calculates until convergence. The ICs denoted by $S(t) = (s_1(t), s_2(t), \ldots\ldots s_m(t))^T$ separated by FastICA will be analyzed by the following method of sample entropy.

2.3 Introduction of Sample Entropy

To assess EEG complexity, sample entropy has been introduced. It can be used for short, noisy time series. It measures the irregularity of a time series and does not involve the construction of the attractor. Sample entropy eliminates self matches and has the advantage of being less dependent on time series length and more consistent when comparisons are made over a broad range of conditions [11]. Thus sample entropy is used to measure complexity of EEG. Most of structures of EEG artifacts are relatively simple, such as power frequency interference, electrooculography (EOG), and electromyography (EMG). The sample entropy of those artifacts is small. However, the sample entropy of real EEG is large. Thus the sample entropy can be used to automatically identify common EEG artifact ingredients. Below is a brief description of the sample entropy.

Consider the time series $s_i(t)$ decomposed by ICA, i =1, 2, 3…m. The two input parameters p and r are choosed, where p is the vector length and r is the criterion of similarity. Let the p samples beginning at sample $s_i(t)$ be denoted by the vector $v_{i,p}(i) = [s_i(t), s_i(t+1), \ldots , s_i(t+p-1)]$ and consider the set of all vectors of length p within $s(n)$, that is $[v_{i,p}(1), v_{i,p}(2), \ldots\ldots v_{i,p}(N-p)]$. Let us define

$$B_{i,p} = \frac{n'_{i,p}(r)}{N\text{-p-1}} \tag{7}$$

Where $n'_{i,p}(r)$ is the number of vectors that are similar to $v_{i,p}(i)$, given the similarity criterion r, excluding self -matching. Similar calculations are carried out for each i, with i=1, 2, 3… $N-p$ .The function $B_p(r)$ is then defined as the average of the function $B_{i,p}(r)$

$$B_p(r) = \frac{\sum_{i=1}^{N-p} B_{i,p}(r)}{N-p} \tag{8}$$

Similarly let $B_{i,(p+1)}(r)$ defined as

$$B_{i,(p+1)}(r) = \frac{n'_{i,(p+1)}(r)}{N\text{-p-1}} \tag{9}$$

Where $n'_{i,(p+1)}(r)$ is the number of vectors in the sequence $[v_{i,p+1}(1), v_{i,p+1}(2), \ldots\ldots v_{i,p+1}(N-p)]$ that are similar to $v_{i,p+1}(i) = [s(i), s(i+1), \ldots , s(i+p)]$, given the similarity criterion r, excluding self -matching. Similar calculations are carried out for each i, with i=1, 2, 3… $N-p$. The function $B_{(p+1)}(r)$ is then defined as the average of the function $B_{i,(p+1)}(r)$.

The statistic $SampEn(p,r,N)$ is then defined by

$$SampEn(p,r,N) = -\ln(B_{p+1}(r) / B_p(r)) \tag{10}$$

Fig.1 shows the *SampEn* of different signals. The *SampEn* of the power frequency interference and EOG are 0.2572, 0.6241 respectively, which are much smaller than the SampEn of real EEG with the value of 1.6135. The SampEn of the above ICs are calculated respectively. The SampEn of these ICs can be denoted by $SampEn = \{SampEn_1, SampEn_2, \ldots , SampEn_m\}$. Then, *SampEn* is used to differentiate artifacts components among ICs.

2.4 EMD-ICA Method

The flow chart for EMD-ICA method is illustrated below in Fig.2.

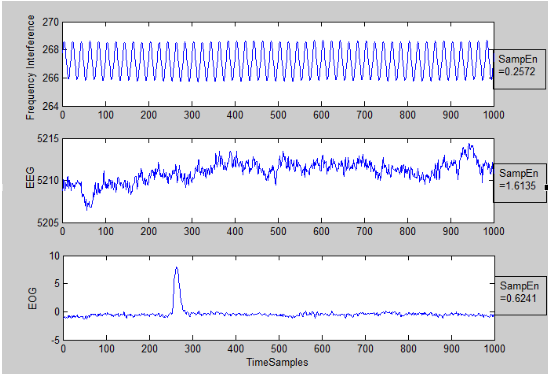


Fig. 1. SampEn of different signals

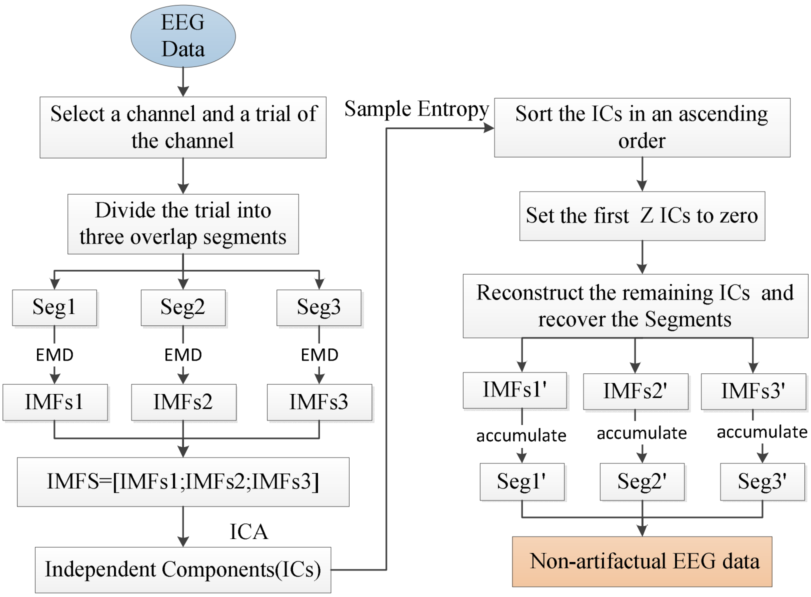


Fig. 2. Flow chart for EMD-ICA

The procedure of EMD-ICA for extracting and removing the artifact just using a single channel is summarized as follows:

1 Select EEG channel related closely with motor imagery. In this paper, CZ is selected.
2 Select one trial from a set of training trials.
3 Divide the trial into three overlap segments (Seg1, Seg2, and Seg3). The Fig.3 below shows how the three segments are divided.
4 Decompose each Seg_i, $i = 1, 2, 3$ into a set of IMF using EMD. Each set of

IMFs constitutes a matrix. The matrix can be denoted by $D_i(t)=[d_{i1}(t),d_{i2}(t),\ldots\ldots d_{in}(t)]$ $i=1,2,3$. Where n is the number of IMFs. We define $d_{i5}(t)=\sum_{j=5}^{n} d_{ij}(t)$. Then $D_i(t)$ can be expressed as $D_i(t)=[d_{i1}(t),d_{i2}(t),\ldots\ldots d_{i5}(t)]$.

5 Then FastICA is employed to separate $D(t)=[D_1(t),D_2(t),D_3(t)]$ into a set of statistics (ICs), denoted by $S(\mathrm{t})=[s_1(t),s_2(t),\ldots\ldots s_m(t)]$ where m is the number of ICs.

6 Calculate the sample entropy of each $s_i(t)$, $i=1,2,\ldots m$. Then a *SampEn* matrix denoted by $SampEn=\{SampEn_1,SampEn_2,\ldots\ldots,SampEn_m\}$ can be obtained. According to the *SampEn*, ICs are sorted in an ascending order. $S'(\mathrm{t})=[s_1'(t),s_2'(t),\ldots\ldots s_m'(t)]$.

7 The first Z ICs are determined as artifacts according to our experience.

8 Set the first Z ICs to zero from $S'(t)$ That is, $s_1'(t),s_2'(t),\ldots\ldots s_z'(t)=0$.

9 The remaining ICs $s_{z+1}'(t),s_{z+2}'(t),\ldots\ldots s_m'(t)$ are reconstructed as $D_{i(reconstructed)}(t)$, denoted by $D_{i(reconstructed)}(t)=[d_{i1}'(t),d_{i2}'(t),\ldots\ldots d_{i5}'(t)]$.

10 Recover Seg_i' by accumulating component $d_{ij}'(t)$. $Seg_i'=\sum_{j=1}^{5} d_{ij}'$.

11 Seg_i' （$i=1,2,3$） are constituted into a non-artifact trial.

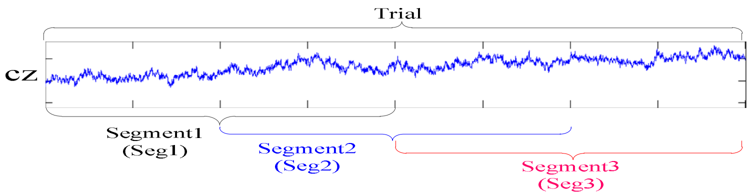


Fig. 3. Selection process of the three Segments

With the method of EMD-ICA, artifacts among the EEG are removed. Following, data from 2008 BCI Competition and our laboratory are used to validate the effect of EMD-ICA.

3 Preparation

3.1 Dataset 1—2008 BCI Competition Data

Dataset 1 used in the study is the publicly available dataset used for BCI Competition IV, which was launched on July 3rd 2008. Dataset 1 was recorded from 7 healthy subjects. For each subject, the two classes of motor imagery were selected from the three classes of left hand, right hand, and foot imagery. EEG signals were recorded from 59 channels, which positioned over sensorimotor areas densely. The signals were band-pass filtered between 0.05 and 200Hz and then sampled at 100Hz. More details are described in [12].

3.2 Dataset 2—Our Laboratory Data

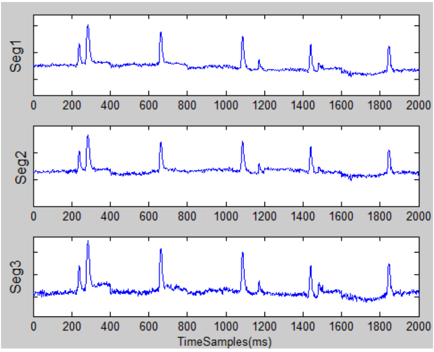
Dataset 2 was from the authors' laboratory experiments. The authors recorded EEG signals using a 16-channel electrode cap. The EEG amplifier was a high-precision biological amplifier developed by Tsinghua University. The EEG signals were transformed by a 24-bit A/D converter and then collected through EEG signal acquisition software. The sampling frequency was 100 Hz. In this experiment, each of seven healthy subjects was asked to complete 60 trials in each session. Each trial included a 4 s left or right hand imagination task. There were eight sessions for each subject and then eight data sets for each subject were obtained.

3.3 Feature Extraction and Classification Method

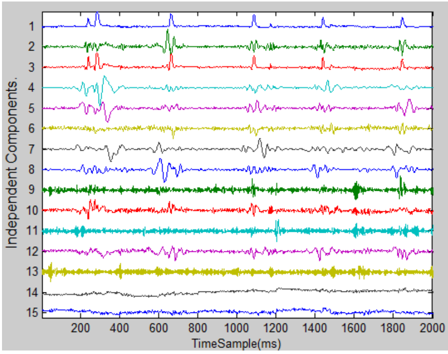
The common spatial patterns (CSP) method is a widely used spatial filtering technique that can extract discriminative features for EEG-based BCI classification tasks [13]. The support vector machine (SVM) is one of the best-known techniques for its good theoretic foundations and high classification accuracy. It has been used extensively in biomedical signal analysis, speech recognition and face recognition [14]. In this paper, the method of EMD-ICA described in 2.4 is applied into the dataset2 described in 3.2 to get a non-artifact EEG dataset. Then, the CSP is used to extract six-dimensional feature of non-artifact EEG dataset. A SVM classifier is used to classify the features extracted by the CSP.

4 Results and Discussions of EMD-ICA

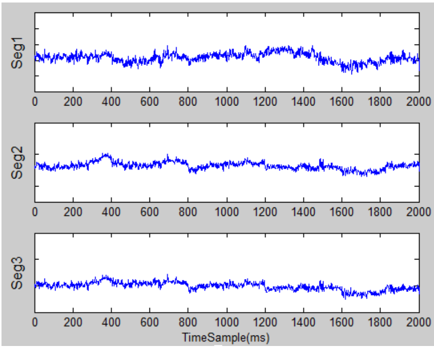
A trial from dataset 1 is randomly selected, and then the trial is used as the input of EMD-ICA described in 2.4. Fig. 4 shows the processing results of the trial. Fig. 4 (a) shows the raw EEG data of CZ channel. It is obvious that the raw data contains many artifacts. Fig.4 (b) shows the ICs of CZ channel. IC 1 and IC 3 can be seen as EOG artifacts, while IC 7 can be seen as ECG. The sample entropies of these ICs (marked by bold red line) are showed in Table 1. It can be seen that the sample entropies of ICs that identified as artifacts are smaller than those of other ICs. Fig.4 (c) shows the reconstructed non-artifactual EEG components of CZ channel. Compared with Fig.4 (a), EOG, and ECG have been significantly removed.



(a) Origin data of CZ channel



(b) CZ data after EMD-ICA



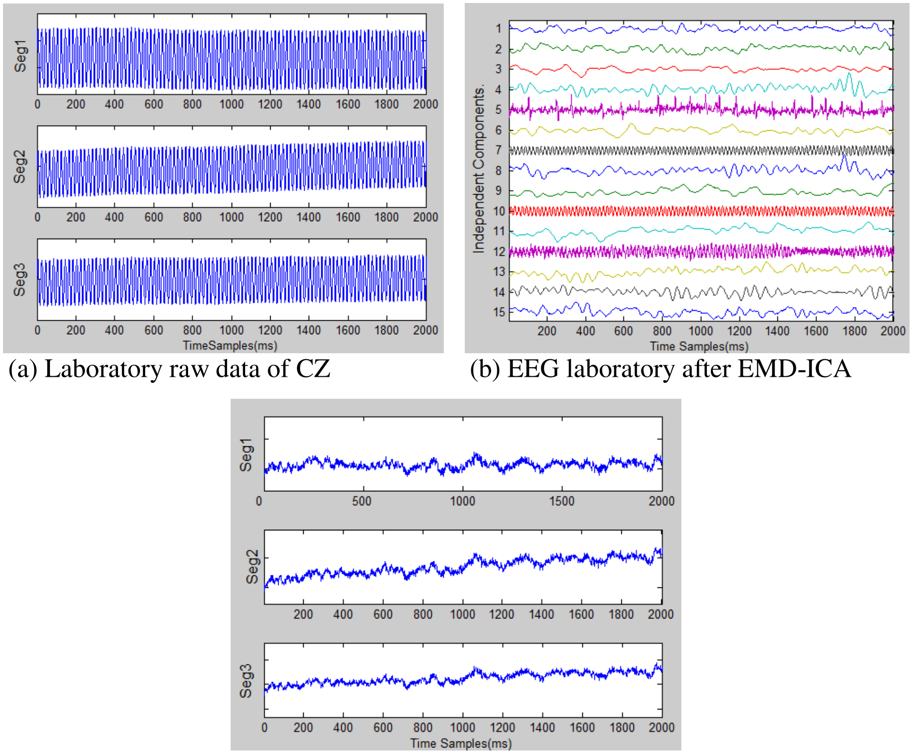
(c) ICs of BCI data after EMD-ICA

Fig. 4. Processing result of dataset 1

Table 1. Sample Entropy of ICs

ICs	**1**	2	**3**	4	5	6	**7**	8
SampEn	**0.3505**	0.5835	**0.3205**	0.7591	0.5518	1.0399	**0.5419**	0.5624
ICs	9	10	11	12	13	14	15	
SampEn	1.9504	1.9317	1.897	1.1469	0.8530	1.2542	0.7570	

Fig.5 shows the results of a random trial from dataset2 after the processing of EMD-ICA described in 2.4. Similar to Fig.4, Fig.5 (a), Fig.5 (b) and Fig.5 (c) describe the raw EEG data, ICs separated by ICA and the reconstructed non-artifactual EEG components respectively. It can be seen from Fig.5 (a) that the raw EEG is covered by power frequency. In Fig.5 (b), IC 7, IC 10 and IC12 can be seen as frequency artifacts, while IC 5 can be seen as EMG artifact. Fig.5 (c) shows that Frequency and EMG artifacts are significantly removed automatically.



(a) Laboratory raw data of CZ

(b) EEG laboratory after EMD-ICA

(c) ICs of Laboratory data after EMD-ICA

Fig. 5. Processing result of dataset 2

The processing results showed in Fig.4 and Fig.5 are from a random trial respectively. The left trials were also processed, and similar results were obtained. In addition to the comparison of results showed above, recognition accuracy is also used to evaluate the effect of the proposed method. Here the CSP is used for feature extraction and the SVM for classification. The details of the CSP and the SVM are described in 3.3. Dataset 2 is processed by EMD-ICA and ICA respectively just using a single channel. In the case of 5-fold cross-validation, the recognition accuracies are shown in Table2. The result shows that the EMD-ICA leads to a higher 14% and 13% recognition accuracy than that of the raw EEG and ICA respectively.

Table 2. Recognition accuracy of different methods with single channel

Method	Recognition Accuracy						
	Subject1	Subject2	Subject3	Subject4	Subject5	Subject6	Subject7
None	0.50	0.52	0.51	0.52	0.53	0.53	0.54
ICA	0.51	0.52	0.52	0.54	0.54	0.53	0.54
EMD-ICA	0.62	0.65	0.71	0.67	0.68	0.67	0.66

Furthermore, the effect of the EMD-ICA is studied in the case of three channels. The result is showed in Table 3. What can be learned from the Table 3 is that the EMD-ICA leads to higher 17% and 4% recognition accuracy than that of the raw EEG and ICA respectively.

Table 3. Recognition accuracy of different methods with three channels

Method	Recognition Accuracy						
	Subject1	Subject2	Subject3	Subject4	Subject5	Subject6	Subject7
None	0.51	0.53	0.54	0.53	0.52	0.53	0.54
ICA	0.64	0.63	0.70	0.66	0.68	0.67	0.65
EMD-ICA	0.66	0.71	0.70	0.68	0.72	0.69	0.74

5 Conclusion

In this paper, a new method for automatically removal artifacts in EEG data based on EMD-ICA is presented. From the experiment results, the proposed method not only can remove the artifacts effectively but also can reduce the number of the raw EEG channels. The EMD-ICA can lead to higher 14% and 13% recognition accuracies than those of the raw EEG and ICA respectively in the case of just using a small number of channels, and higher 17% and 4% by using three channels. Thus, the EMD-ICA provides a foundation for online and practical application.

Acknowledgments. This project is supported by National Natural Science Foundation of China (60975079, 31100709), Innovation project of Shanghai Education Commission(11YZ19), Shanghai University, "11th Five-Year Plan" 211 Construction Project.

References

1. Wei, H., Pengfei, W., Liping, W.: A Novel EMD-Based Common Spatial Pattern for Motor Imagery Brain-Computer Interface. In: Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI 2012), Hong Kong and Shenzhen, China, January 2-7, pp. 216–219 (2012)

2. Hassan, A., Xiaomu, S.: A Study of Kernel CSP-based Motor Imagery Brain Computer Interface Classification. In: 2012 IEEE on Signal Processing in Medicine and Biology Symposium (SPMB), pp. 1–4. IEEE Press, New York (2012)
3. Azzerboni, B., Foresta, F.L., Mammone, N.: A new approach based on Wavelet-ICA algorithms for fetal electrocardiogram extraction. In: 13th European Symposium on Artificial Neural Networks, pp. 193–198. D-side Publication, Bruges (2005)
4. Jui, C., WeiYeh, L., Ju, H.: An Online Recursive ICA Based Real-time Multi-channel EEG System on Chip Design with Automatic Eye Blink Artifact Rejection. In: 2013 International Symposium on VLSI Design, Automation, and Test, Taiwan (2013)
5. Xiaojing, S., Yantao, T., Yang, L.: Feature Extraction and Classification of sEMG Based on ICA and EMD Decomposition of AR Model. In: 2011 International Conference on Electronics, Communications and Control (ICECC), pp. 1464–1467. IEEE Press, Ningbo (2011)
6. Soomro, M.H., Badruddin, N.: A Method for Automatic Removal of Eye Blink Artifacts from EEG Based on EMD-ICA. In: 9th IEEE International Colloquium on Signal Processing and its Applications, pp. 129–134. IEEE Press, Malaysia (2013)
7. Mijović, B., Vos, M.D., Gligorijević, I.: Combining EMD with ICA for Extracting Independent Sources from Single Channel and Two-Channel Data. In: 32nd Annual International Conference of the IEEE EMBS, pp. 5387–5390. IEEE Press, Argentina (2010)
8. Simon, R.H.D., Christopher, J.J.: Novel use of Empirical Mode Decomposition in single-trial classification of Motor Imagery for use in Brain-Computer Interfaces. In: 35th Annual International Conference of the IEEE EMBS, pp. 5610–5613. IEEE Press, Osaka (2013)
9. Abdollah, A., Amin, J.M., Hassan, M.: A Survey on EMD Sensitivity to SNR for EEG Feature Extraction in BCI Application. In: 2010 5th Cairo International Biomedical Engineering Conference, Cairo, pp. 176–179 (2010)
10. Huang, L., Wang, H.: Reducing the Computation Time for BCI Using Improved ICA Algorithms. In: Guo, C., Hou, Z.-G., Zeng, Z. (eds.) ISNN 2013, Part II. LNCS, vol. 7952, pp. 299–304. Springer, Heidelberg (2013)
11. Lei, W., Guizhi, X., Jiang, W.: Motor Imagery BCI Research Based on Sample Entropy and SVM. In: 2012 Sixth International Conference on Electromagnetic Field Problems and Applications, Dalian, pp. 1–4 (2012)
12. Benjamin, B., Guido, D., Matthias, K.: The non-invasive Berlin Brain-Computer Interface Fast acquisition of effective performance in untrained subjects. J. NeuroImage 37, 539–550 (2007)
13. Yan, L., Yasuharu, K.: A real-time BCI with a small number of channels based on CSP. J. Neural Computing and Applications. 20, 1187–1192 (2012)
14. He, X., Wei, S., Zhiping, H., Cheng, C.: A Speedup SVM Decision Method for Online EEG processing in Motor Imagery BCI. In: 2010 10th International Conference on Intelligent Systems Design and Applications, ISDA 2010, Cairo, pp. 149–153 (2010)