EEG SIGNAL CLASSIFICATION USING NONLINEAR INDEPENDENT COMPONENT ANALYSIS

Farid Oveisi

MSR Research Institute, Tehran, IRAN

ABSTRACT

One of the preprocessors can be used to improve the performance of brain–computer interface (BCI) systems is independent component analysis (ICA). ICA is a signal processing technique in which observed random data are transformed into components that are statistically independent from each other. This suggests the possibility of using ICA to separate different independent brain activities during motor imagery into separate components. However, there is no guarantee for linear combination of brain sources in EEG signals. Thus the identification of non-linear dynamic of EEG signals should be taken into consideration. In this paper, a new method is proposed for EEG signal classification in BCI systems by using non-linear ICA algorithm. The effectiveness of the proposed method is evaluated by using the classification of EEG signals. The tasks to be discriminated are the imaginative hand movement and the resting state. The results demonstrate that the proposed method performed well in several experiments on different subjects and can improve the classification accuracy in the BCI systems.

Index Terms— EEG, brain-computer interface, nonlinear independent component analysis, classification, genetic algorithm

1. INTRODUCTION

Recently, many efforts have been done to use the electro-encephalogram (EEG) as a new communication channel between human brain and computer. This new communication channel is called EEG-based brain–computer interface (BCI). Most of these efforts have been dedicated to the improvement of the accuracy and capacity of this EEG-based communication channel. One of the most important factors about the performance of BCI systems is classification system. A classification system typically consists of both a preprocessor and a classifier. Preprocessors are used to improve the performance of classifier systems. One of the preprocessors can be used to improve the performance of brain–computer interface (BCI) systems is independent component analysis (ICA) [1-2]. ICA is a signal processing technique in which observed random data are transformed into components that are statistically independent from each other [3]. ICA is a useful technique for blind separation of independent sources from their mixtures. Sources are usually original, uncorrupted signals or noise sources. Linear ICA was used to separate neural activity from muscle and blink artifacts in spontaneous EEG data [4]. It was verified that the ICA can separate artifactual, stimulus locked, response-locked, and non-event related background EEG activities into separate components [5]. Furthermore, ICA would appear to be able to separate task-related potentials from other neural and artifactual EEG sources during hand movement imagination in form of independent components. In [1] has been showed that the power spectra of the linear ICA transformations provided feature subsets with higher classification accuracy than the power spectra of the original EEG signals. However, there is no guarantee for linear combination of brain sources in EEG signals. Thus the identification of non-linear dynamic of EEG signals should be taken into consideration. For non-linear mixing model, linear ICA algorithms fail to extract original signals and become inapplicable because the assumption of linear mixtures is violated and the linear algorithm cannot compensate for the information distorted by the non-linearity. Therefore, in this paper, a nonlinear ICA has been used to separate task-related potentials from other neural and artifactual EEG sources. The proposed method has been tested on several different subjects. Moreover, the results of proposed method were compared to the results obtained using linear ICA, and original EEG signals.

2. METHODS

2.1. Mutual Information

Mutual information is a non-parametric measure of relevance between two variables. Shannon's information theory provides a suitable formalism for quantifying these concepts. In accordance with Shannon's information theory, the mutual information can be expressed as

$$I(X,C)=\sum_{c\in C}\int_{\mathbf{x}} p(\mathbf{x},c)\log\frac{p(\mathbf{x},c)}{p(c)p(\mathbf{x})}d\mathbf{x}\,. \tag{1}$$

where $p(c)$ and $p(\mathbf{x})$ represent the probability of the discrete random variable C and $\mathbf{x}$, respectively.

If the mutual information between two random variables is large, it means two variables are closely related. Indeed, MI is zero if and only if the two random variables are strictly independent.

2.2. Nonlinear Independent Component Analysis

Conventional linear ICA approaches assume that the mixture is linear by virtue of its simplicity. However, this assumption is often violated and may not characterize real-life signals accurately. A realistic mixture needs to be non-linear and concurrently capable of treating the linear mixture as a special case. Generally, a non-linear ICA problem can be defined as follows: given a set of observations, $x(t)=[x_1(t),x_2(t),\ldots,x_n(t)]^T$ which are random variables and generated as a mixture of independent components $s(t)=[s_1(t),s_2(t),\ldots,s_n(t)]^T$ according to

$$x(t)=f[s(t)] \tag{2}$$

where f is an unknown nonlinear mixing transform (NMT). The block diagram of the nonlinear ICA is shown in Figure 1.

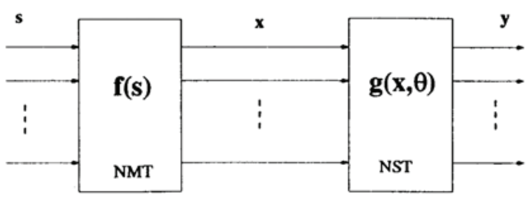

Fig. 1. Nonlinear mixing and separating systems for independent component analysis.

The separating system $g(.,\theta)$ in the right part of Fig. 1, called nonlinear separation transform (NST) is used to recover the original signals $x(t)$ from the nonlinear mixture without the knowledge of the source signals $s(t)$ and the mixing nonlinear function f . However, a fundamental difficulty in nonlinear ICA is that it is highly non-unique without some extra constraints; therefore, finding independent components does not lead us necessarily to the original sources [6].

ICA in the nonlinear case is, in general, impossible. In [7] has been added some extra constraints to the nonlinear mixture so that the nonlinearities are independently applied in each channel after a linear mixture. As figure 2 shows, the proposed algorithm in [7] needs to estimate two different mixtures: a family of nonlinearities g which approximates the inverse of the nonlinear mixtures f and a linear unmixing matrix W which approximates the inverse of the linear mixture A.

For the demixing system, first we need to approximate g_i, which is the inverse of the nonlinear function in each

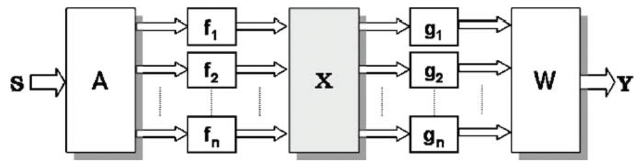

Fig. 2. Post-nonlinear mixing and demixing models for independent component analysis.

channel, and then separate the linear mixing by applying W to the output of the g_i nonlinear function

$$y_i(t)=\sum_{j=1}^{n} w_{ij}\, g_i\left(x_j(t)\right) \tag{3}$$

In order to develop a more general and flexible model of the function g_i, can be used a M th order odd polynomial expression of nonlinear transfer function (g_i):

$$p_j\left(x_j\right)=\sum_{k=1}^{M} p_{jk}\, x_j^{2k-1} \tag{4}$$

where $p_j = \lfloor p_{j1}, p_{j2}, \ldots, p_{jM} \rfloor$ is a parameter vector to be determined. By using relations (3) and (4), we can write the following criterion for the output sources y_i **:**

$$y_i(t)=\sum_{j=1}^{n} w_{ij} \sum_{k=1}^{M} p_{jk}\,.x_j^{2k-1} \tag{5}$$

The parameter vector p_j should be determined so that the inverse of the mutual information of the output sources y_i is maximized. To achieve this objective, can be defined the following criterion [7]:

$$eval\_function(y)=\frac{1}{I(y)} \tag{6}$$

In this paper, a genetic algorithm (GA) [8] was used for mutual information optimization. Unlike many classical optimization techniques, GA does not rely on computing local first- or second-order derivatives to guide the search process; GA is a more general and flexible method that is capable of searching wide solution spaces and avoiding local minima (i.e., it provides more possibilities of finding an optimal or near-optimal solution). To implement the GA, I use genetic algorithm and direct search toolbox for use in Matlab (The Mathworks, R2007b).

The linear demixing stage has been performed by the well-known Infomax algorithm [3]. To be precise, Infomax has been embedded into the GA in order to approximate the linear mixture**.**

In this application, the genetic algorithm is run for 30 generations with population size of 20, crossover probability 0.8, and uniform mutation probability of 0.01. The number of individuals that automatically survive to the next generation (i.e., elite individuals) is selected to be 2. The scattered function is used to create the crossover children by creating a random binary vector and selects the genes where the vector is a 1 from the first parent, and the genes where the vector is a 0 from the second parent.

2.3. Multiple Classifiers

Multiple classifiers are employed for classification of extracted components. The Multiple Classifiers are used if different sensors are available to give information on one object. Each of the classifiers works independently on its own domain. The single classifiers are built and trained for their specific task. The final decision is made on the results of the individual classifiers. In this work, for each component, separate classifier is trained and the final decision is implemented by a simple logical majority vote function. The desired output of each classifier is -1 or $+1$. The output of classifiers is added and the *signum function* is used for computing the actual response of the classifier. The diagonal linear discrimination analysis (DLDA) [9] is here considered as the classifier. The classifier is trained to distinguish between rest state and imaginative hand movement.

3. EXPERIMENTAL SETUP

The EEG data of healthy right-handed volunteer subjects were recorded at a sampling rate of 256 from positions Cz, T5, Pz, F3, F4, Fz, and C3 by Ag/AgCl scalp electrodes placed according to the International 10-20 system. The eye blinks were recorded by placing an electrode on the forehead above the left brow line. The signals were referenced to the right earlobe.

Data were recorded for $5\,\mathrm{s}$ during each trial experiment and low-pass filtered with a cutoff 45 Hz. There were 100 trails acquired from each subject during each experiment day. At $t = 2\,\mathrm{s}$, a cross ("+") was displayed on the monitor of computer as a cue visual stimulus. The subjects were asked to imagine the hand grasping in synchronization with the cue and to not perform a specific mental task before displaying the cue. In the present study, the tasks to be discriminated are the imaginative hand movement and the idle state.

Eye blink artifact was suppressed by using independent component analysis. The artifactual independent components were visually identified and set to zero.

4. RESULTS

The nonlinear ICA algorithm, proposed in [7], was applied to given training 7-channel EEG data sets associated to the hand movement imagination and resting state. Original features are formed from 1second interval of each component, in the time period 2.3–3.3 seconds, during each trial of experiment. The window starting 0.3 seconds after cue presentation is used for classification. The number of local extrema within interval, zero crossing, 5 AR parameters, variance, the mean absolute value (MAV), and 1Hz frequency components between 1 and 35Hz constitute the full set of features with size 44. The classifier is trained to distinguish between rest state and imaginative hand movement. The imaginative hand movement can be hand closing or hand opening. From 200 data sets, 100 sets are randomly selected for training, while the rest is kept aside for validation purposes. Training and validating procedure is repeated 10 times and the results are averaged.

The results have been recorded for four subjects (AE, ME, BM, SN) for different experiment days. Table 1 summarizes the results of classification accuracy of the original EEG signals. The average classification accuracy is 73.84%.

Table 2 summarizes the results of classification accuracy for different subjects by using linear ICA (the Infomax algorithm [3]). The average classification accuracy over all subjects is 74.61% which 1% better than that obtained original EEG signals. An average classification rate of 77.95% is achieved by using nonlinear ICA. As can be observed, components which are obtained by nonlinear ICA improved the EEG classification accuracy compared to the linear ICA and original EEG signals. These results are 4 percent higher than average classification results by using the raw EEG data. Fig. 3 shows the classification accuracy rate obtained by nonlinear ICA (NICA), linear ICA (LICA), and original EEG signals (channel).

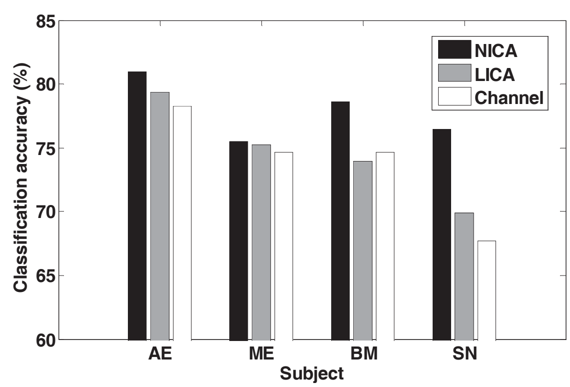

Fig. 3. Mean classification accuracy of EEG patterns for different subjects using nonlinear ICA (NICA), linear ICA (LICA), and original EEG signals (channel).

5. CONCLUSION

Preprocessing plays an important role in the performance of BCI systems. One of the preprocessors can be used to improve the performance of BCI systems is independent component analysis (ICA). ICA would appear to be able to separate task-related potentials from other neural and artifactual EEG sources during hand movement imagination in form of independent components. However, there is no guarantee for linear combination of brain sources in EEG signals. Therefore, in this paper a novel method was proposed for EEG signal classification in BCI systems by using non-linear ICA algorithm. The results of applying this method on four subjects have demonstrated that the proposed method in this paper has improved the mean classification accuracies in relation to raw EEG data and linear ICA. The analysis of variance (ANOVA) shows that the mean classification accuracies achieved by using non-linear ICA are significantly different $(p < 0.01)$.

11. REFERENCES

[1] D. A. Peterson, J. N. Knight, M. J. Kirby, C. Anderson, and M. Thaut, ″ Feature selection and blind source separation in an EEG-based brain-computer interface,″ *EURASIP Journal on Applied Signal Processing*, vol. 19, pp. 3128–3140, 2005.

[2] H. Serby, E. Y. Tov, and F. Inbar, ″ An improved P300-based brain-computer interface,″ *IEEE Trans. Neural Systems and Rehabilitation Engineering* , vol. 13, no. 1, pp. 89-98, March 2005.

[3] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis,* John Wiley & Sons, 2001.

[4] T. Jung, S. makeig, C. humphiers,T. Lee, M. Mckeown, and T. Sejnowski, ″ Removing electroencephalographic arifacts by blind source separation, ″ *Psychophysiology*, pp. 163- 178, 2000.

[5] T. Jung,S. makeig, M. Westerfield, J. Townsend, E. Courchesne, T. Sejnowski, ″Analysis and visualizion of single-trial event-related potentials, ″ *Human Brain Mapping,* vol. 14, pp. 166-185, 2001.

[6] S. Achard, and C. Jutten, ″ Identifiability of post-nonlinear mixtures, ″ *IEEE Trans. Signal Processing Letters*, vol. 12, no. 5, pp. 423-426, May 2005.

[7] F. Rojas, C. Puntonet, M. Alvarez, I. Rojas, and R. Clemente, ″ Blind source separation in post-nonlinear mixtures using competitive learning, simulated annealing, and a genetic algorithm,″ *IEEE Trans. Systems, Man, and Cybernetics.*, vol. 34, no. 4, pp. 407-416, nov. 2004.

[8] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989.

[9] W. J. Krzanowski, *Principles of multivariate analysis: a user's perspective*, Oxford University Press, Oxford, 2000.

TABLE 1
CLASSIFICATION ACCURACY RATE OF ORIGINAL EEG SIGNALS DURING HAND MOVEMENT IMAGINATION.

Subject	Day1	Day2	Day3	Day4	Day5	mean
AE	77.3	76.4	75.5	83.9	-	78.27
ME	65.3	84.9	74.6	73.8	-	74.65
BM	67.2	90.6	66	75.2	-	74.75
SN	77.4	66.1	61.6	69.4	64.1	67.7
mean	71.8	79.5	69.42	75.57	64.1	73.84

TABLE 2
CLASSIFICATION ACCURACY RATE OF EXTRACTED COMPONENTS DURING HAND MOVEMENT IMAGINATION USING LINEAR ICA.

Subject	Day1	Day2	Day3	Day4	Day5	mean
AE	76.3	81.9	77.9	81.4	-	79.37
ME	68.7	84.1	77.2	71.1	-	75.27
BM	67.1	93.3	63	72.5	-	73.97
SN	78.9	71.1	64.1	67.6	67.6	69.86
mean	72.75	82.6	70.55	73.15	67.6	74.61

TABLE 3
ACCURACY RATE OF EXTRACTED COMPONENTS DURING HAND MOVEMENT IMAGINATION USING NONLINEAR ICA.

Subject	Day1	Day2	Day3	Day4	Day5	mean
AE	77.6	81	80.1	85.3	-	81
ME	72.8	80.5	76.6	72	-	75.47
BM	76.2	93	69	76.2	-	78.6
SN	78.5	79	81.7	72.5	72	76.74
mean	76.28	83.38	76.85	76.5	72	77.95