2021 2nd International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) | 978-1-6654-0764-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICHCI54629.2021.00071

Feature Processing of Multi-classification Motor Imagery EEG based on improved ICA and SVM

Xiaoxiao Gong
Fuyang Normal University
School of Physics and Electronic Engineering
Fuyang, China
bettygxx@126.com

Shuguang Chen
Fuyang Normal University
School of Physics and Electronic Engineering
Fuyang, China

Yongbin Ban
Fuyang Unicom
Department of Innovation
Fuyang, China

Miaomiao Wang
Fuyang Normal University
School of Physics and Electronic Engineering
Fuyang, China

***Abstract*—In order to increase the recognition accuracy of multiple motor imagery EEG (MIEEG) signals in Brain Computer Interface (BCI), this paper presents a novel idea for motor imagery recognition based on improved Independent Component Analysis (ICA) and a multi-class Support Vector Machine (SVM) model. After the original EEG signal is preprocessed, an effective filter is designed by improved ICA algorithm to automatically extract the independent components related to motor imagery, and then the independent components are optimized to construct the EEG feature set, which is input into the SVM classifier finally. Our work is tested on BCI III contest data set and self-collected data set, the experimental results show good recognition accuracy and strong robustness, which proves our method is appropriate for EEG processing.**

Keywords-MIEEG; ICA; SVM

I. INTRODUCTION

The technology of BCI is one of the hottest discussed issues in scientific research today. It refers to the ability to obtain EEG signals from our brain, and then convert them to the control of external devices, thereby establishing a direct man-machine interaction. MIEEG are endogenous, spontaneously-generated signals with the characteristics of low signal-to-noise ratio (SNR), low spatial resolution and non-stationary. Human scalp EEG studies have shown that alpha (8-13Hz) and beta (15-25Hz) bands can be used to detect motor imagery activation as they are most associated with sensorimotor processes. Specifically, when brains conduct sensorimotor activities of left or right hands, related neurons are activated, and the rhythmic energy in certain frequency bands increases on the same side of the brain, decreases on the contralateral side of the brain simultaneously, which is called the phenomenon of event-related synchronization (ERS) /desynchronization (ERD)[1]. When sensorimotor activities of feet are conducted, rhythmic energy changes mainly occur in the precentral motor cortex.

The commonly used feature extraction algorithm of EEG signals include Wavelet Transform (WT)[2], Autoregressive model(AR)[3], Common Spatial Pattern(CSP)[4], Principal Component Analysis(PCA) and Independent Component Analysis(ICA) etc. According to the features of EEG signals, the spatial filtering algorithm ICA combines information of time and space to extract the most effective filters, but the order of output independent components is uncertain. To solve the problem, professor Wu proposed an improved ICA algorithm, but it was not combined with the classification algorithm[9]. On this basis, we further optimize the features extracted by the improved ICA and use the pattern recognition algorithm of SVM to achieve better classification recognition results.

II. TECHNICAL PRINCIPLE OF INDEPENDENT COMPONENT ANALYSIS (ICA) AND SUPPORT VECTOR MACHINE (SVM)

A. Principle of improved ICA

In 1994, P.Com systematically analyzed the problem of blind source separation on the instantaneous mixed signals and put forward the concept of ICA after doing a further study on PCA[5]. In 1995, A.J.Bell and T.J. Senowski proposed the maximum-entropy method based on Linsker's information maximization criterion(Infomax)[6]. In 1999, Hyvarinen and Oja proposed FastICA algorithm, which has a fast convergence speed and is appropriate for batch processing[7]. In this paper, we use the extended-Infomax improved by T.W.Lee et al[8], compared with the traditional Infomax algorithm, it can adjust the nonlinear function on the mixing matrix according to the kurtosis change of output results in the process of blind source separation, so as to realize the automatic separation of super-Gaussian and sub-Gaussian sources in EEG signals.

The principle of the algorithm is as follows: For the observed signal $X(t)=[x_1(t),\dots x_N(T)]^T$, A is the unknown mixing matrix. According to the transmission criterion of maximum information, the extended-Infomax continuously adjusts the separation matrix W through the post-learning of network, aimed at the maximum of mutual information $I(x,y)$ between the nonlinear output Y and input X, which is equivalent to the maximum of output entropy $H(y)$, as is shown in Equation (1):

$$H(y)=H(y_1)+H(y_2)+\dots+H(y_n)-I(y_1,y_2\dots y_n) \qquad (1)$$

978-1-6654-0764-9/21/$31.00 ©2021 IEEE
DOI 10.1109/ICHCI54629.2021.00071

Only if the nonlinear outputs y_i are independent of each other, the equations can be established: $I(y_1, y_2, \cdots y_n) = 0$ and $H(y) = H(y_1) + H(y_2) + \ldots + H(y_n)$. Based on the information theory of mutual information maxima, it can be transformed to an optimization problem of separation matrix listed in equation (2): Where $g_i(\,)$ is a reversible monotone nonlinear function.

$$\begin{cases} \max : O(W) \\ O(W) = H(y) = H(g(Wx)) \end{cases} \tag{2}$$

Mixing matrix A can be obtained from the separation matrix W. Spatial characteristics exist in the column vectors of A. The weight of column vectors in A represents the projection intensity of signal sources on each electrode.

B. *Principle of SVM*

Vapnik et al proposed the Support Vector Machine (SVM) approach in 1995[10]. For non-linearly separable data samples, SVM maps input variables to a high-dimensional space, so that data can be classified in the high-dimensional space and the optimal classification hyperplane can be obtained. SVM specializes in samples of small size, nonlinear, high dimension. In our experiments, SVM has achieved good classification performance.

The problem of optimal hyperplane in SVM can be transformed into the problem of following minimization model:

$$\begin{cases} \min\limits_{\omega, b} \ \dfrac{1}{2} \| \omega \|^2 \\ s.t. \ \ y_i(\omega \cdot x_i + b) \geq 1, i = 1, 2, \ldots, N \end{cases} \tag{3}$$

After the Lagrange operator is introduced, it can be converted into a problem of optimal quadratic function with inequality constraints. Finally, we can get an optimal classification function:

$$g(x) = \mathrm{sgn}\{(\omega \cdot x) + b\} = \mathrm{sgn}\{\sum_{i=1}^{N} a_i^* y_i (x_i \cdot x) + b^*\} \tag{4}$$

Where a_i^* is the corresponding Lagrange multiplier and b^* is the classification threshold.

Equation (5) is the classification and discrimination formula in the high-dimensional space, where the decision function can be replaced by the kernel function:

$$f(x) = \mathrm{sgn}[\sum_{i=1}^{N} y_i a_i K(x_i, x) + b] \tag{5}$$

Choosing different kernel functions will affect the results of SVM algorithm. Under the constraints of Mercer's theorem, the kernel function can be selected freely. After the kernel function is selected, the relevant parameters can be adjusted, then training data set can be input to construct a SVM classifier model. Finally we use the model to predict the labels of test data set.

III. Feature Processing Algorithms Based on Improved ICA and SVM

A. *Feature Construction*

Flexible means can be used to construct the spatial filters by ICA, ICA has lower requirements on data quality compared with CSP. Therefore, we can extract the filters with EEG data of one or some consecutive trials, or data segment from one trial. In this paper, we choose ten consecutive trials as the raw EEG data.

As it is known, the order and amplitude of independent components separated by ICA algorithm are uncertain, and each column of the matrix A represents the projection coefficients of an independent component. Fig.1 shows an example of the brain topographic mapping to visualize the projection. Due to the attenuation effects in the propagation process of EEG signals[8], for an independent component, the distance to an electrode location is inversely proportional to its projection intensity on the electrode. Because the positions of electrode C3, C4 and Cz are mostly associated with motor imagery activities of left hand, right hand or foot movements. MIICs have maximum projection coefficient in C3, C4 or Cz channel.

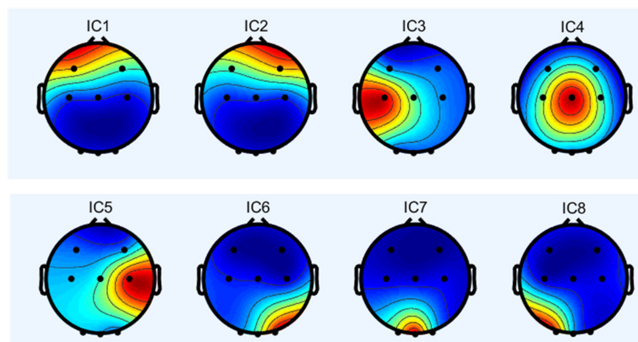

Figure 1. Brain topographic mapping

Fig.2 shows an example, firstly we calculate the absolute value of matrix A and get a set of electrodes according to the row number of the maximum value of each column. Then we identify the electrode C3, C4 and Cz by the row numbers, the corresponding column number are 3, 4 and 5 respectively. Note that if the electrode set does not contain C3, C4 or Cz, the filter design has failed. Finally, we extract the third, fourth, and fifth column in W^T to construct a set of filters $\{\mathrm{w}_1, \mathrm{w}_r, \mathrm{w}_f\}$. The obtained filter set will be utilized to perform spatial filtering on every single trial of test data to extract three types of MRICs.

Besides three classical motor imagery tasks, motor imagery of tongue movements can be added as the fourth task[11]. Tongue is connected to five of the twelve pairs of brain nerves. Therefore, the imagination of tongue movement also causes changes of scalp EEG signals. The nerve that controls tongue movement is the hypoglossal nerve, which is in the occipital brain. So in our experiment, we choose Oz lead as location of the independent component related to tongue movements. The corresponding column from the separation matrix W^T can be selected as the spatial filter, then we have a filter group $\{\mathrm{w}_1, \mathrm{w}_r, \mathrm{w}_f, \mathrm{w}_t\}$ for four motor imagery tasks.

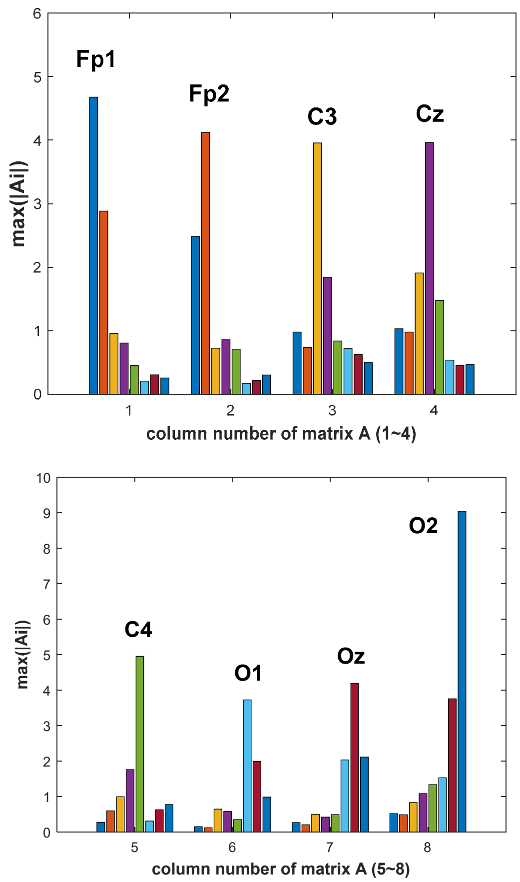

Figure 2. Analyzation of matrix A

B. *Feature optimization and parameters adjustment*

Feature selection and optimization are very important that strongly affect the design and performance of classifier, so we select variances of independent components as the basis to detect the phenomenon of ERD/ERS. The main parameters that we need to adjust include regularization parameter C and kernel function parameter. We aim to maintain the balance between over-fitting and under-fitting of the model and make the model have relatively high accuracy in both training set and test set.

IV. Experiments and Result Analysis of Experiments

Self-collected data set. Three-class motor imagery experimental data set were obtained from 16-channel MIEEG signals collected by Neuroscan system. The electrode positions were set as follows: VEOU, VEOL, Fp1, Fp2, FC3, FCZ, FC4, C3, CZ, C4, CP3, CPZ, CP4, O1, OZ, O2. Original EEG was sampled at 250Hz and filtered between 1 and 50z with Notchfilter on. In this paper, three healthy subjects named A, B and C with no history of brain disease were selected. Each data set contained 75 trials, including 45 times left hand, right hand and foot motor imagery respectively. The experimental process of one trial is shown in the figure below:

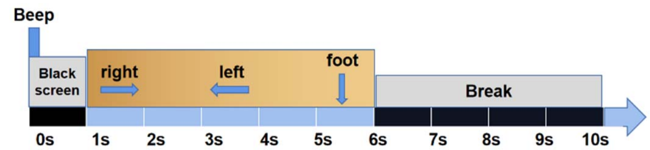

Figure 3. The experimental process

BCI competition data set. We used four-class motor imagery data set IIIa from BCI III Competition. Data from two subjects named k3b and l1b, the experiment was performed in several rounds (at least six rounds), each contained 40 trials. Each trial lasted for 7 seconds. In the first 2s shown a blank screen, at t=2s, a cross appeared on the screen. From 3s, an arrow appeared for 1s, to remind the subject to imagine a corresponding movement of left hand, right hand, foot or tongue until the cross disappeared at 7s. More details about the data set can be got from http://www.bbci.de/competition/iii/.

Raw MIEEG data is digitally bandpass filtered (8 - 30Hz) and averaged, then fold cross-validation is used to test algorithm accuracy. We can see from fig.4, in general, we have achieved good results in both data sets. Self-collected data sets have a higher recognition rate, which is associated with fewer imagination tasks.

As the choices of frequency bands and lead positions determine the recognition rate to a certain extent. In the future, we will optimize the option to achieve better results in multi-classification motor imagery. Due to the existence of individual physiological differences, for some subjects, the recognition with EEG rhythm at 10-14Hz is good, while for others, at 8-12Hz may be better. We know that tongue movement is related to EEG signals in the occipital part of the brain. Oz lead has been selected in the experiment. Next, we will choose other leads, such as CPz or Pz lead, to further explore the influence of lead positions on the results.

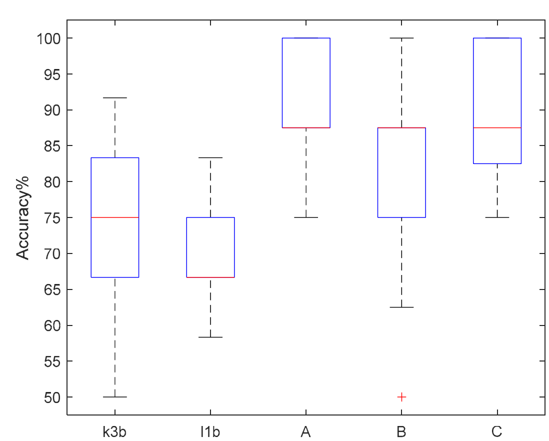

Figure 4. Recognition results of improved ICA+SVM

In order to verify the influence of data quality on our algorithms, we select effective motor imagery data vs all data of subject A to construct filters, and make a comparison of recognition rates on the same data set of self-collected experiments. The experimental result is shown in fig.5, there is little difference in recognition rates. Which proves our method has strong robustness to artifacts and abnormal values.

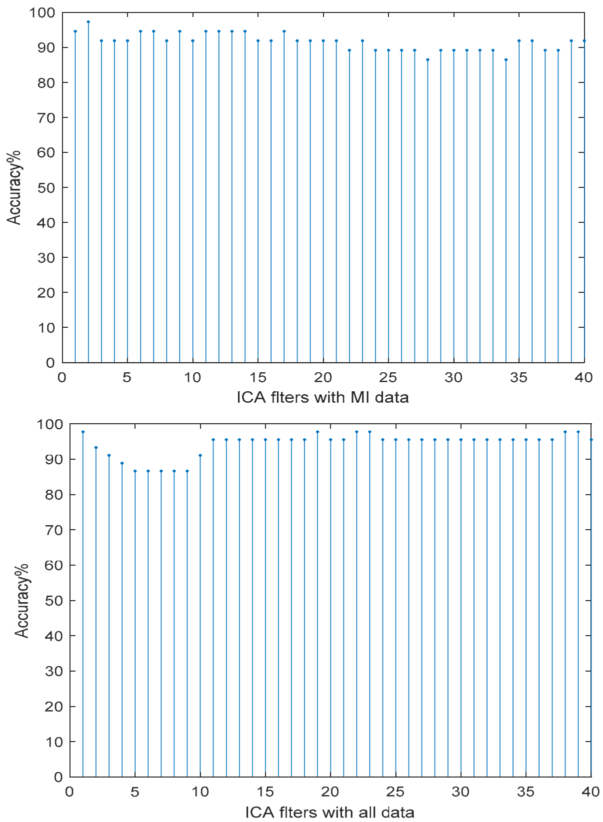

Figure 5. A comparison

V. Conclusion

In our study, a new strategy based on improved ICA algorithm and SVM model to process multi-classification MIEEG is proposed. ICA is commonly used for EEG data processing, but in the past, we needed background knowledge when selecting the filter set related to motor imagery. In this paper, we use the improved ICA to automatically extract spatial filter set of MIICs, and feature optimization is introduced in the pattern recognition, we choose the variance of MIICs as the feature to detect the ERD/ERS phenomenon, and it is no need to analyze how the EEG rhythm changes when conducting motor imagery activities, we only need to adjust the parameters of SVM classifier to train a model and predict test labels, experimental results show that the algorithm has good nonlinear approximation capability and strong robustness.

Acknowledgement

This research work is supported by project of Department of Science and Technology of Anhui Province (S202004a05020071).

Reference

[1] Rebreikina AB, Larionova EV, Varlamov AA. Evoked Changes in the Rhythmic Activity of the Brain During Processing of Visually Presented Target, Nontarget, and Unknown Words[J]. Neuroscience and Behavioral Physiology, 2016, 46(5).

[2] Yang T, Ma Y, Meng M, et al. Motor Imagery Eeg Feature Extraction Based on Fuzzy Entropy with Wavelet Transform[C]//Proceedings of the 10th International Conference on Computer Engineering and Networks (cenet2020), 2020: 1680-1690.

[3] Tang Q,Zhou J,Xin J, et al. Autoregressive Model-based Structural Damage Identification and Localization Using Convolutional Neural Networks[J]. Ksce Journal of Civil Engineering, 2020(prepublish).

[4] Xiaoou L. [eeg Feature Extraction Based on Ica and Csp Algorithms].[J]. Sheng Wu Yi Xue Gong Cheng Xue Za Zhi = Journal of Biomedical Engineering = Shengwu Yixue Gongchengxue Zazhi, 2010, 27(6).

[5] James C. Introduction and Overview of the Bss/ica Problem - Specifically When Applied to Biomedicine[J]. Iee Seminar on Blind Source Separation in Biomedicine, 2004.

[6] Mental Eeg Analysis Based on Infomax Algorithm[J]. Chinese Journal of Biomedical Engineering(english Edition), 2004(1): 7-15.

[7] Junliang Y,Haipeng R,Qing L. Fixed-point Ica Algorithm for Blind Separation of Complex Mixtures Containing Both Circular and Noncircular Sources[J]. The Journal of China Universities of Posts and Telecommunications, 2016, 23(2): 15-23.

[8] Zhou weidong JLAYKOISSUJSOCSSUJ. Blind Separation of Biosignals Using Extended Infomax Algorithm[C]//Icbme'2002 of the 6th International Conference on Biomedical Engineering and Rehabilitation Engineering, 2002: 295-296.

[9] Jing R, Xiaopei W, Bangyan Z, et al. An Automatic Channel Selection Approach for Ica-based Motor Imagery Brain Computer Interface.[J]. Journal of Medical Systems, 2018, 42(12).g Z. An Overview of Ica (independent Component Analysis) Applications in Remote Sensed Data[J]. Spie Remote Sensing, 2005.

[10] Chen CH, Wang Z. An Overview of Ica (independent Component Analysis) Applications in Remote Sensed Data[J]. Spie Remote Sensing, 2005.

[11] H SS, V NC, R OC, et al. Motor Imagery Practice and Increased Tongue Strength: a Case Series Feasibility Report.[J]. Journal of Speech, Language, and Hearing Research: Jslhr, 2019, 62(6).