Motor Imagery BCI Research Based on Sample Entropy and SVM

Lei Wang, Guizhi Xu, Jiang Wang, Shuo Yang, Miaomiao Guo, Weili Yan
The Province-Ministry Joint Key Laboratory of Electromagnetic Field and Electrical Apparatus Reliability, Hebei University of Technology
Tianjin, China

Jiang Wang
TangShan Vocational Technical College
TangShan, China

***Abstract*—Brain Computer Interface (BCI) is a system that provides an artificial communication between the human brain and the external world. It may give disabled people direct control over a neuro-prosthesis by their intentions that are reflected in their brain signals. In this paper, brain electric field data (EEG) was recorded though 28 electrodes placed on the scalp. According to the fact that EEG is non-stationary and non-linear; a non-linear dynamic method called Sample Entropy (SampEn) was applied to extract the features of EEG. A Support Vector Machine (SVM) classifier was structured for pattern classification. The final results show that SampEn is an effective method to extract the feature of different brain states.**

Keywords-component; Brain Computer Interface; Sample Entropy; motor imagery; Support Vector Machine

I. INTRODUCTION

Brain computer interface systems create a novel communication channel from the brain to an output device do not though conventional motor output pathways of nerves and muscles [1]. It is suitable for patients with severest neurological or muscular diseases, such as amyotrophic lateral sclerosis, high-level spinal cord injure or brain stem stroke, and so on. All such "locked-in" patients are completely paralyzed physically and unable to speak, but cognitively intact and alert. BCI can help the patients release their mind and reconnect with the world [2].

ALS

神经生理学观察

According to the neurophysiological observation, imagination of limbs (hands, feet) movement can attenuate or even block the mu and beta rhythm activity, which is termed ERD. On the contrary, the activity enhancement of the mu rhythm at some area of the brain is called ERS. Both ERD and ERS are tightly related with motor imagery [3]. The subjects can imagine different limb movement in order to enhance or block specific brain wave. In this paper, ERD and ERS at some key area were checked, features of EEG were extracted by SampEn, and then subject's intend was determined by SVM classifier.

削弱 阻塞

When SampEn was used to extract feature of different brain states from EEG data, different parameters setting will affect the value of SampEn and then the final classification accuracy. In the last section of this paper, above-mentioned question will be discussed.

II. EXPERIMENT AND DATA ACQUISITION

A. Subjects

Three healthy volunteers (all males, all right-handed, 20~32 years old, mean age was 27.6 years.) participated in this study. All the subjects had normal vision or corrected vision. All of these subjects had accomplished a few hours BCI training.

B. Process of experiments

The subjects sat in an armchair and stared at the computer monitor placed approximately 70 centimeters away from the subject at eye level. Meanwhile, they were asked to keep their arms and hands relax and comfortable and avoid eye movements during the recordings.

The subjects were told to imagine left hand, right hand or foot movement following the cue emerging on the screen. Each trail started with a blank screen lasting for 2 seconds. At second 2, a fixation cross appeared at the center of the monitor. At second 3, the cross disappeared and at the same time an arrow emerged at the center of the monitor, which was pointing left, right, or upward. The subjects were instructed to imagine a movement of the left hand, right hand or the foot, depending on the direction of the arrow. The arrow vanished after 4 seconds, when it disappear the subject should stop imagine movement. The subject can relax for a random period about 2~3 seconds, then coming the next trial. The whole experiment paradigm is shown in Fig.1.

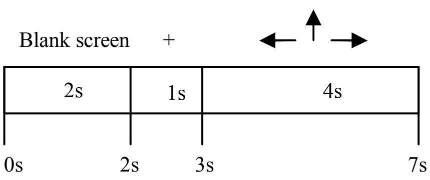


Fig 1 Timing scheme of one trail

One session contains 10 times of each movements. Every subject performed 8~20 sessions which depend on their state and bearing capability.

This work was supported by the Natural Science Foundation of China under Grant No. 31100711, the Natural Science Foundation of Hebei Province, CHINA under Grant No. E2009000062.

C. *Data Recording and Preprocessing*

EEG was recorded using Ag/AgCl electrodes in a 64 channel modified quick cap (NeuroScan, Inc). The electrodes are evenly spaced and symmetrically covered the scalp from nasion to inion and from left to right ear, according to the 10-20 system. The reference electrodes were positioned on left and right mastoids. Horizontal EOG and vertical EOG were also recorded. All signals, including EEG and EOG were sampled at 250 Hz.

The original EEG signals were preprocessed using Neuro Scan 4 software system. During this process, two things were done. One is to remove the disturbances caused by EOG; the other is bandpass filter between 0.1Hz and 40Hz.

Twenty-eight EEG channels which distributed in the primary motor area were chosen to investigate in this study. The electrodes were F1, Fz, F2, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6, P1, Pz, P2 and PO2. EEG data from 3 to 7 second were used in following investigation.

III. METHOD

A. *Sample Entropy*

Sample Entropy is a statistic method to quantify the unpredictability of fluctuations in both deterministic and stochastic signals, which first proposed by Richman. Sample Entropy can be applied to short time series and it is adequate for stochastic, noisy deterministic and composite processes. Sample Entropy is used extensively in earthquake prediction, speech recognition, hydrologic detection and physiological time series analyzation [4]. In this paper, SampEn was used to extract the particular features of EEG recorded during different motor imagery.

For calculating the SampEn of a time series $x(1)$, $x(2)$, … ,$x(N)$, where N is the length of the time series, two fixed parameters, m and r, must be determined firstly. The parameter m is the series of vectors length, and r is the similar tolerance. The SampEn algorithm is as blow:

Step 1. Form a series of vector $X(1) \sim X(N\text{-}m\text{+}1)$, whose length is m, defined by:

$$X(i)=\left[x(i),x(i+1),\cdots,x(i+m-1)\right] \quad i=1\sim N\text{-}m\text{+}1 \quad (1)$$

Step 2. The distance $d[X(i), X(j)]$ between two vectors $X(i)$ and $X(j)$ is defined as the maximum absolute difference in the scalar components of $X(i)$ and $X(j)$.

$$d[X(i),X(j)]=\max_{k=0\sim m-1}[|x(i+k)-x(j+k)|] \quad (2)$$

Step 3. Define $N^m(i)$, the number of vector $X(j)$ $(j=1\sim N-m+1, j\neq i)$ such that the distance between the vectors $X(j)$ and the generic vector $X(i)$ is lower than r.

Define $B_i^m(r)$ as:

$$B_i^m(r)=\frac{N^m(i)}{N-m+1} \qquad i=1\sim N\text{-}m\text{+}1 \quad (3)$$

Step 4. The actual logarithmic average over all the vectors of the $B_i^m(r)$ probability is computed as:

$$B^m(r)=\frac{1}{N-m+1}\sum_{i=1}^{N-m+1}B_i^m(r) \quad (4)$$

Step 5. Increase the length of vector to m+1. Repeat steps 1 to 4 and find $B_i^{m+1}(r)$

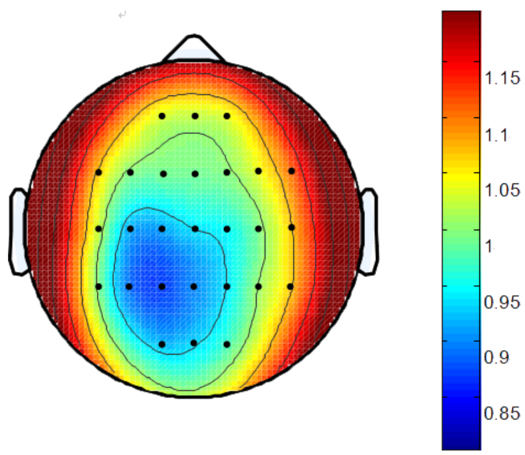
Step 6. Theoretically, the sample entropy is defined as:

$$SampEn(m,\ r)=\lim_{N\to\infty}\{-\ln[B^{m+1}(r)/B^m(r)]\} \quad (5)$$

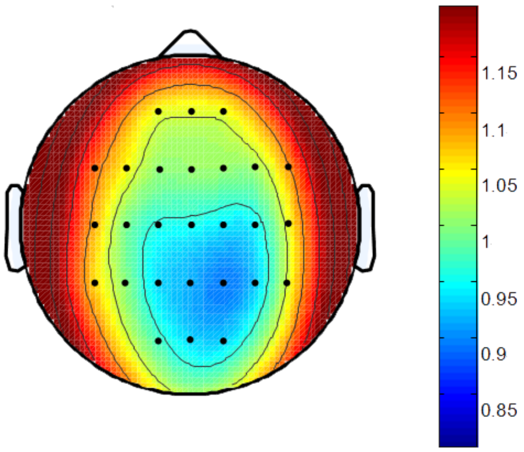
In practice, the length of time series N is finite, therefore, we can only get the estimate result instead of precise result. When the length is N, estimation of SampEn denoted as:

$$SampEn(m,\ r,N)=-\ln[B^{m+1}(r)/B^m(r)] \quad (6)$$

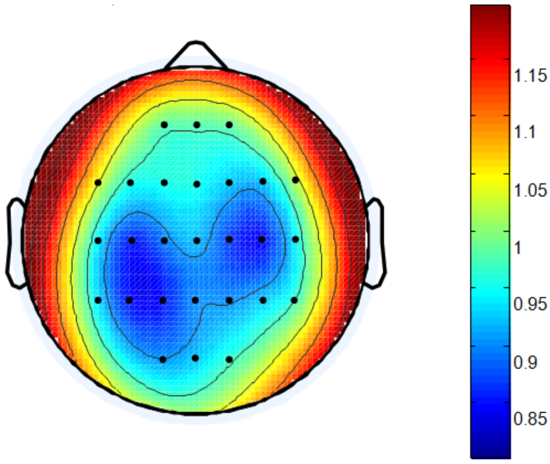
In the literature, m can be set as 2. Similar tolerance r can be chosen between 0.1~2 SD, where SD is the standard deviation of the time series $x(i)$. In our research, the parameter m was taken as 2. The value of r affected the final accuracy greatly. In different subject's EEG data, the parameter setting is different, details will be discussed in the last section. The values of SampEn at different EEG channels were computed as the features. Totally 28 features were got, which represent each movement imagery state. Fig 2 is the average SampEn distribution of subject 1 during three motor imagery tasks. The parameter r was set as 0.1 SD.



(a)Imagine left hand movement



(b)Imagine right hand movement



(c)Imagine foot movement

Fig 2 Distribution of Sample Entropy during three imagery tasks

From Fig 2, we can find that different imagery states show different distribute patterns. SampEn as the EEG features can be used in the following classification.

B. SVM classifier

There are a number of standard classification techniques in literature, such as simple rule based and nearest neighbor classifiers, Bayesian classifiers, artificial neural networks, fisher classifiers, support vector machine (SVM), etc. Among these techniques, 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. In this paper, a SVM classifier with RBF kernel function was structured in order to classify the features extracted by SampEn. A RBF SVM two parameters C and γ must be set before the classifier established. Four combinations of C and γ were researched. Two of them, C and γ were random selected as C=100,γ=1 and C=500,γ=2. The third was decided by the method grid search. In the last situation, optimize method GA was used to find appropriate values of C and γ [5]. The accuracies of classifications were shown in the following section.

IV. Results

The data about each subject consisted of 480 trails, including 160 repetitions of each type of metal tasks (left or right hand imagination or foot imagination). Four fold cross validation was applied to test the performance of proposed procedure. The three-category average accuracy of each subject was shown in Table I. The results of optimized SVM classifier using grid search and GA method were also shown in Table I.

TABLE I. The Classification Results using SVM (%)

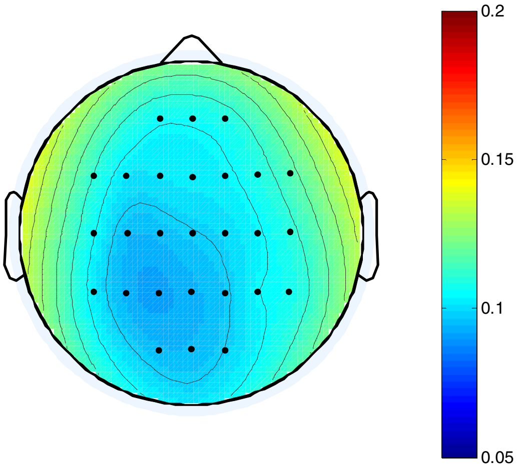
Subject	**SVM** (C=100,γ=1)	**SVM** (C=500,γ=2)	**SVM Optimized by Grid Search**	**SVM Optimized by GA**
Subject 1	75.62 ± 3.7	69.17 ± 1.8	76.88 ± 1.6	77.08 ± 1.6
Subject 2	66.25 ± 3.2	61.87 ± 4.2	68.75 ± 2.2	70.42 ± 1.4
Subject 3	56.46 ± 1.7	60.63 ± 1.2	61.25 ± 2.2	62.29 ± 1.4

The results show that SampEn is an effective method to extract the feature of different brain states. The parameters C and γ can impact the accuracy of classification. For the purpose of gain higher accuracy, optimization of SVM classifier is necessary.

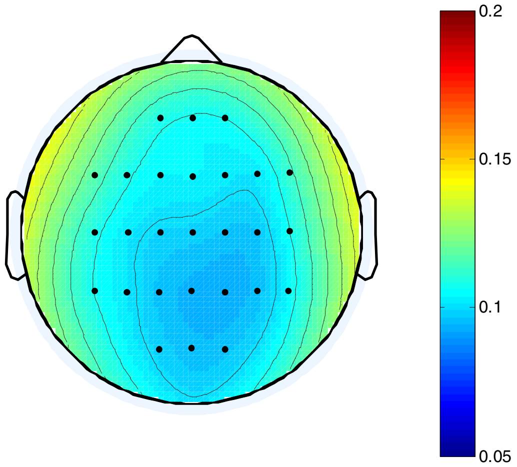
V. Discussion

As mentioned in section III, similar tolerance r will affect the value of SampEn greatly. If r was set as a big value, then the number of vectors which satisfy equation in step 3 will decrease, SampEn represents the signal's tendency rather than the details. If r was set as a small value, $N^m(i)$ will be big, and SampEn represents the signal's details rather than the tendency. The situation of r takes different values can be regarded as SampEn represents the complexities of signal under certain scales. In our research, the value of r will impact the SampEn value of each EEG data recording during different mental tasks. As SampEn is taken as EEG fetures, r will affect the accuracy of classification ultimately. Fig 3 is the averages SampEn distribution of subject 1 same to Fig 2. But in Fig 3 similar tolerance r was set as 1.5 SD. It can be easily seen that the difference of three states is not as obviously as in Fig 1.

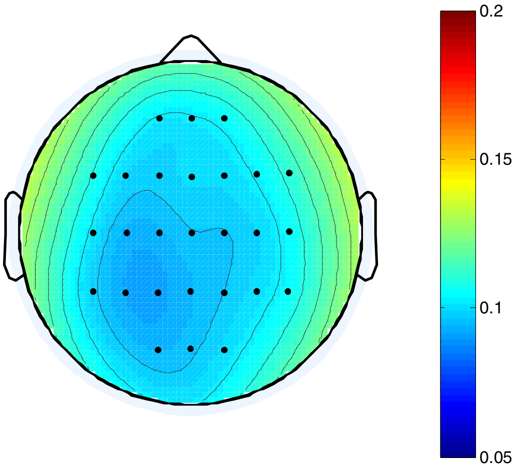
The curves in Fig 4 have shown the relationship between similar tolerance r and the final classification accuracy of all three subjects. From the figure we can see that, every subject's highest accuracy has great difference. The accuracy changes with the value of r. In order to gain the highest accuracy of ecah subject, the value of r must be considered deeply and exhaustive.



(a)Imagine left hand movement



(b)Imagine right hand movemen



(c)Imagine foot movement

Fig 3 Distribution of sample entropy during three imagery tasks as r=1.5SD

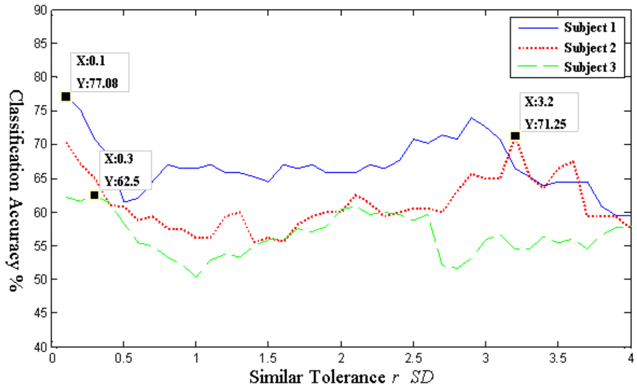


Fig 4 Classification accuracy of three subjects at different r

VI. CONCLUSION

In this paper, the method used in our BCI reseacrh, named SampEn was introduced. Features of EEG data recoded during three different motor imagery tasks were extracted by the method. Based on the extracted feature sets, SVM classifier was structured. Two algorithms were applied to optimize the performance of the SVM classifier. The final accuracy showed that SampEn as the method to extract the EEG features has certain advantage. In practical applications, in order to gain the highest accuracy the effect of r must be considered.

Further study will focus on enhancing the classification accuracy. An on-line BCI system design based on SampEn will be performed also.

REFERENCES

[1] J.R.Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. “Brian-computer interfaces for communication and control,” Clinical Neurophysiology, vol. 113, 2002, pp. 767-791.

[2] I.H. Iversen, N. Ghanayim , N. Neumann , et al. “A brain−computer interface tool to assess cognitive functions in completely paralyzed patients with amyotrophic lateral sclerosis,”Clinical Neurophysiology , vol 119, ,2008, pp. 2214-2223.

[3] G. Pfurtscheller, C. Brunner, A. schlogl and F.H. Lopes da Silva, “Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks,” NeuroImage, vol. 31, 2006, pp. 153-159.

[4] Richman Joshua S, I. Randall Moorman. “Physiological time series analysis using approximate entropy and sample entropy.” American journal of pahysiology. Heart and circulatory physiology, vol. 6, 2000, pp. 2039-2049.

[5] Olivier Devos, Cyril Ruckebusch, Alexandra Durand, et al. “Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation,” Chemometrics and Intelligent Laboratory Systems, vol 96, 2009, pp. 27-33