Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Classifying ECoG/EEG-Based Motor Imagery Tasks

Bin An, Yan Ning, Zhaohui Jiang, Huanqing Feng, Heqin Zhou

Abstract—The multichannel electrocorticogram (ECoG)/ electroencephalogram (EEG) signals are commonly used to classify two kinds of motor imagery (MI) tasks. In this paper, the ECoG and EEG data sets are composed of training and test data, which are recorded during different time/days. Power spectral density (PSD) is selected as features; fisher discriminant analysis (FDA) and common spatial patterns (CSP) are used to filter redundancy; K-Nearest-Neighbor (KNN) classifier is applied to classify MI tasks; and a new function R (k) is presented to estimate the value of k. Using these methods, we obtain the predictive accuracy of MI tasks based on ECoG data (which is 92%) and EEG data (which is 81%). The results show that we can effectively classify two kinds of MI tasks based on EEG as well as ECoG.

Keywords—MI, ECoG, EEG, KNN

I. INTRODUCTION

THE MI is defined as a mental representation of movement without any body movement [1]. Classifying MI tasks is to distinguish the different mental movement, which is helpful to human beings to communicate with each other, especially to the handicapped. Compared with other signals used to distinguish mental movements, the EEG and ECoG signals possess lots of advantage, such as high temporal resolution and easy recording for EEG signals, high spatial resolution and low noise in ECoG signals. With the development of microelectrode technology [2], the ECoG signals can be recorded more conveniently than before. Extracting features, filtering redundancy and designing classifier are three important aspects which will affect the results of classification of MI tasks. The temporal, frequency and statistic features are usually used in the realm of classifying MI tasks. Among these features, PSD is most commonly used by researchers. Filtering redundancy is a necessary progressing step, and several methods are proposed, such as common spatial subspace decomposition [3], [4] and FDA [5]. There are various classifiers that can be selected for EEG and ECoG classification, such as support vector machines, clustering, via cluster mean, linear discriminant analysis [6], [7]. In this paper, we propose a new method to classify two kinds of MI tasks not only based on EEG signals but also on ECoG signals. We extract the PSD as feature, and use the methods of FDA and CSP to filter redundancy of the signals. The KNN classifier [8] is selected to distinguish two kinds of MI tasks. Finally, we will discuss the relation between EEG and ECoG in classifying MI tasks.

This work was sponsored by China Nature and Science Foundation (project No.60422201) & Creativity Fund of University of Science & Technology of China (USTC) (200509).

Bin An, Department of Electronic Science & Technology, USTC, POBox4 Hefei Anhui, 230027 China. (E-mail: ban@mail.ustc.edu.cn).

Yan Ning, Department of Electronic Science & Technology, USTC, POBox4 Hefei Anhui, 230027 China. (E-mail: ningyan@mail.ustc.edu.cn).

Zhaohui Jiang, Department of Electronic Science & Technology, USTC, POBox4 Hefei Anhui, 230027 China. (E-mail: jiangzhh@ustc.edu.cn).

Huanqing Feng, Department of Electronic Science & Technology, USTC, POBox4 Hefei Anhui, 230027 China. (E-mail: hqfeng@ustc.edu.cn).

Heqin Zhou, Department of Automation, USTC, POBox4 Hefei Anhui, 230027 China. (phone: 0086-551-3601800; fax: 0086-551-3601522; E-mail: hqzhou@ustc.edu.cn).

II. METHODS

A. Data

There are two groups of data used in this paper [7], which are named as Data Set I and II. Data Set I was recorded when one subject executed a task of motor imagery (left pinky, tongue). It is composed of training (278 trials) and test data (100 trials), both of which are the same subject's ECoG recordings from two different sessions with about one week in between. The data was collected by using an 8x8 ECoG platinum electrode grid that was placed on the right motor cortex. In Data Set II, another subject executed the MI tasks, in which he pressed computer keyboard with his left hand and right hand before body movement. This data are also composed of training and test data, both of which are the same subject's EEG recordings from two different sessions with some minutes break in between.

B. Feature

In this study, neither can we mark the starting or ending moments of each mental task, nor can we compare the temporal waveform of each channel of signal to another [10]. So we analyze ECoG signals in the frequency domain. Considering the randomicity of EEG signals, autoregressive (AR) model is used to estimate the signal spectrum and PSD is chosen [3], [11] as features. Then we locate the frequency range corresponding to the MI tasks.

The process of features extraction is illustrated by using the ECoG signals recorded by the electrode 39 and 1 (Fig. 1 Panel A and B). As shown in Panel B(electrode39), PSD of two kinds of tasks are obviously distinguished, especially in the range of 8~12Hz. However, in Panel A (electrode1) two curves are almost overlapped. Therefore, we extract features from the PSD between 8Hz to 12Hz.

C. Wipe off the Redundancy

The research on the magnetic resonance imaging infers that

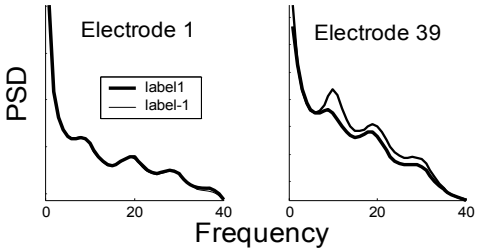

Fig. 1. PSD Comparing (One task is labeled with thick line while the other is labeled with thin line).

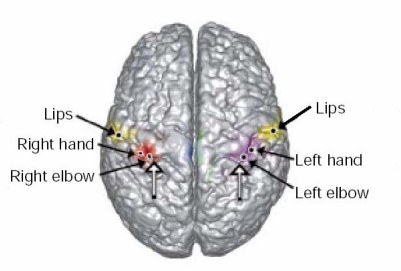

Fig. 2. The function of motor cortex [9]

the activity of motor cortex is close related to organic activity. Some areas in Fig.2 show the brain areas which are related to several kinds of MI tasks, such as imaged lips movement, hands movement and elbows movement. In recording, the platinum electrode grid was assumed to cover the right motor cortex completely, but it also partly covered surrounding areas outside motor cortex due to its size (Approx. 8x8cm). It is possible that the data recorded inside the motor cortex offer more distinct features than those of outside motor cortex. In other words, the data recorded by some electrodes affect the predictive accuracy obviously, but the data from other channels may be redundant and will lower the classification precision. So we wipe off the redundant channels with less discrimination risk. Based on the above requirements, we use FDA to sort out the channels on which the features are obviously distinguished.

According to Fisher discriminator theory that maximizing the discriminant leads to the best separation between two projected data sets, we seek directions that are efficient for the discrimination of the samples.

D. CSP

The EEG/ECoG signals and the noises contained in them possess obvious features of space distribution. One common effective method to improve signal noise ratio (SNR) is CSP filtering [12]. By using CSP filter designed based on the training data, several channels are selected from multichannel signals to carry most discrimination information. We extract the spatial features from signals of two kinds of motor imagery to maximize the discrimination of two kinds of tasks.

E. Classification

We use KNN algorithm to perform the classification. Based on the KNN theory, the error ratio of the algorithm decreases while k increases, and finally equals to the minimum error ratio of Bayes decision. To make sure the k_th nearest neighbor is near enough to the sample, k<<N must be tenable. So how to estimate k is very important and is also one of the difficulties in this study. By investigating the m-fold cross-validation method, we construct a function R (k). The k is obtained by maximizing R. The algorithm to select k is as followed:

Define X as a sample set and Y a test set. N is the number of samples, and M is the number of test samples.

Separate X to m subsets:

$$X_{\mathrm{i}}, \quad \mathrm{i}=1,2,\cdots \mathrm{m} \tag{1}$$

The new test set:

$$NewY = X_j, \quad j = 1,2,\cdots m \tag{2}$$

New sample set:

$$NewX = \{X_{\mathrm{i}}, \quad \mathrm{i} \neq \mathrm{j}, \mathrm{i}=1,2,\cdots \mathrm{m}\} \tag{3}$$

Then we use kNN algorithm to predict NewY and obtain the test precision:

$$A_j(k), \quad j = 1,2,\cdots m \tag{4}$$

Define a function:

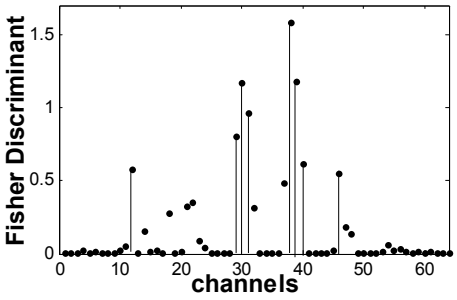

Fig. 3. Fisher Discriminant

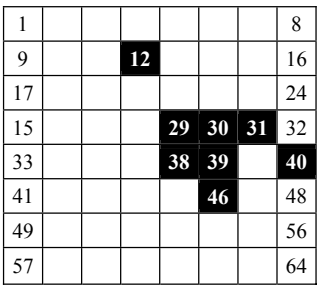

Fig. 4. The eight dark areas show the position of eight channels on the platinum electrode grid.

$$R(k) = \sum_{m=[N/M]+1}^{N} \left(\frac{1}{m}\sum_{j=1}^{m} A_j\right) \quad (5)$$

When m= [N/M] +1, the number of samples of NewY approaches that of Y; When m=N, the number of samples of NewX is closest to X. So the range of m is [N/M] +1~N.

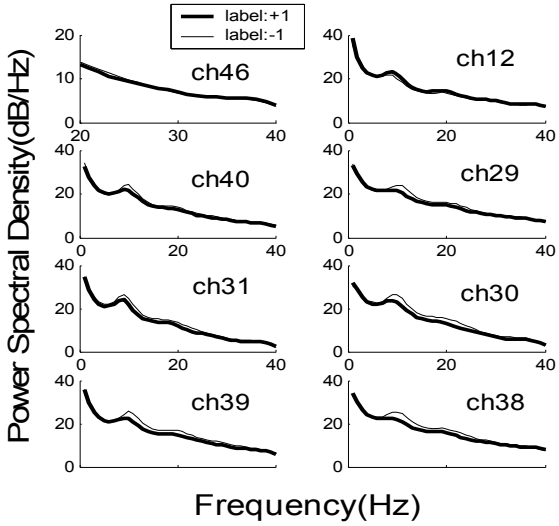

Fig. 5. The PSD Comparing of eight outputs channels of FDA (One task is labeled with thick line while the other is labeled with thin line).

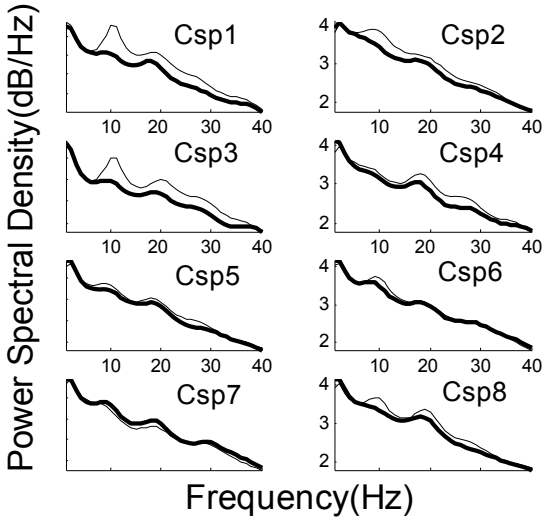

Fig. 6. PSD Comparing of outputs of CSP filter on training data (One task is labeled with thick line while the other is labeled with thin line).

III. RESULTS

The results of Data Set I is shown in Fig.3 as an example; Fig. 3 shows that big differences are existed among the values of Fisher discriminant for each channel. That is, it is easier to classify two kinds of MI tasks in some channels than others. So we choose eight channels, which are the eight most efficient directions for the discrimination of the samples.

TABLE I
PREDICTION ACCURACY OF DATE SET I

Data Set I (ECoG)					
	With CSP			Without CSP	
K	R(k)	ACCU	K	R(k)	ACCU
1	0.8294	0.85	1	0.8112	0.8
2	0.8197	0.86	2	0.8134	0.77
3	0.8311	0.88	**3**	**0.8387**	**0.83**
4	0.8524	0.88	4	0.8175	0.8
5	0.8484	0.91	5	0.8355	0.83
6	0.8553	0.89	6	0.8377	0.82
7	0.8566	0.91	7	0.834	0.84
8	**0.8599**	**0.92**	8	0.8273	0.84
9	0.8431	0.9	9	0.8326	0.86
10	0.8561	0.92	10	0.8245	0.84
11	0.8462	0.89	11	0.8343	0.85
12	0.8468	0.9	12	0.8187	0.85
13	0.8424	0.93	13	0.8195	0.88
14	0.8497	0.93	14	0.827	0.84
15	0.8439	0.92	15	0.8322	0.86
16	0.8504	0.92	16	0.828	0.85
17	0.8437	0.93	17	0.8244	0.87
18	0.8413	0.92	18	0.8282	0.85
19	0.8431	0.91	19	0.8264	0.86
20	0.8447	0.9	20	0.8251	0.87

ACCU means accuracy and 0.92=92%.

TABLE II
PREDICTION ACCURACY OF DATE SET II

Data Set II (EEG)					
	With CSP			Without CSP	
K	R(k)	ACCU	K	R(k)	ACCU
1	0.755	0.74	1	0.5264	0.58
2	0.7189	0.71	2	0.4962	0.48
3	0.7729	0.77	3	0.5102	0.49
4	0.7487	0.73	4	0.4992	0.45
5	0.7805	0.72	**5**	**0.5331**	**0.45**
6	0.7759	0.74	6	0.5119	0.41
7	0.7827	0.78	7	0.4932	0.49
8	0.7774	0.77	8	0.4786	0.5
9	0.7921	0.79	9	0.4942	0.48
10	0.8044	0.76	10	0.479	0.52
11	0.8124	0.78	11	0.4976	0.46
12	0.8013	0.8	12	0.4731	0.48
13	0.8019	0.8	13	0.4891	0.42
14	0.7949	0.8	14	0.4599	0.46
15	0.8122	0.79	15	0.4822	0.44
16	0.8087	0.79	16	0.4554	0.48
17	0.7999	0.79	17	0.4686	0.49
18	0.7959	0.8	18	0.4586	0.45
19	**0.8133**	**0.81**	19	0.476	0.49
20	0.7997	0.8	20	0.4759	0.46

ACCU means accuracy and 0.92=92%.

They are arranged in decreasing order as [38 39 30 31 29 40 12 46] and their locations on the platinum electrode grid are shown in Fig.4. We find that the dark area in Fig.4 almost matches the brain area in Fig.2. In addition, Fig. 5 shows that

the PSD of two classes of tasks of these eight electrodes are obviously distinguishing

As shown in Fig. 6, the CSP filtering is applied on the signals recorded at these eight electrodes and the resulting PSD is obtained. The feature of two kinds of MI tasks extracted from each of the eight outputs are obviously distinguishing, and their PSD curves are not overlapped like those in Fig. 1 Panel A.

We used KNN to classify two classes of MI tasks. During this process, the most important step is to select an appropriate value of k. Since the number of nearest neighbor should be a subsection of total number of training samples, it is required that k<<278 for Data Set I. Therefore, we choose k not lager than 20. In the condition that the labels of test data are unknown, we firstly calculate R (k) based on training data, and then investigate the results, shown in Table 1, and find that R reaches the maximum when k=8. So we choose 7NN classification algorithm to predict the test data, and finally obtain a very good prediction accuracy of 92%.

For comparison, we perform the prediction without using CSP method and only obtain the prediction accuracy of 83%, which makes R reach maximum when k=3, shown in Table I.

With the same methods, we also investigated the prediction accuracy of Data Set II. The results are shown in Table II.

IV. Conclusions

In this paper, we study the classification algorithm for two classes of MI tasks based on the multi-dimensional EEG and ECoG signals. Based on ECoG data and EEG data, high prediction accuracy are achieved at 92% and 81%, respectively.

Using the extracted features in this paper, we can preferably classify two kinds of MI tasks, especially those tasks based on ECoG. After the FDA is applied to wipe off the redundant channel signals and CSP is applied to include more information, the prediction accuracy is remarkably improved. Therefore, we conclude that CSP method can effectively improve SNR and then improve the prediction accuracy.

The prediction accuracy based on EEG (81%) is lower than that based on ECoG and CSP affected EEG data more than ECoG data. Comparing to ECoG signals, EEG signals may contain less information that can be used to classify MI tasks than ECoG signals do. We attribute the case to the influence of scalp and skull, which make the ECoG signals of nerve cells chaotic. As a result, it is more difficult and inaccurate to classify MI tasks using EEG than ECoG signals.

The algorithm used to find k makes classification of two kinds of MI tasks more accurate. For instance, in Data Set II, when S (19) =0.8133 (maximum) the prediction accuracy is maximum accordingly; in Data Set I, The maximum of S (k) is 0.8599 when k equals to 8 and the prediction accuracy is 92%. This accuracy approaches to the maximum (93%, when S (14) =0.8497). These results show that the algorithm is effective. Meanwhile, we find a problem that the accuracy is not the maximum in Table II. In the future work, we will try to solve the problem and investigate how to improve the accuracy further.

The algorithm in this paper is robust, because the high prediction accuracy is achieved by using the classifier trained on the first day to classify data recorded on a following day. During this interval, the subject might change into a different state concerning motivation, fatigue etc. so that his brain will show different electrical activities. In addition, the recording system might have undergone slight changes concerning electrode positions and impedances.

Furthermore, we will study other MI tasks and apply these methods to classify three and more kinds of MI tasks.

Acknowledgment

We thank University of Tübingen, Germany, Dept. of Computer Engineering (Prof. Rosenstiel) and Institute of Medical Psychology and Behavioral Neurobiology (Niels Birbaumer), and Max-Planck-Institute for Biological Cybernetics, Tübingen, Germany (Bernhard Schölkopf), and Universität Bonn, Germany, Dept. of Epileptology (Prof. Elger) to provide Data Set I.

We thank Fraunhofer-FIRST, Intelligent Data Analysis Group (Klaus-Robert Müller), and Freie Universität Berlin, Department of Neurology, Neurophysics Group (Gabriel Curio) to provide Data Set II.

Bin An thank Dr. Lei Feng (SUNY-Stony Brook) for helpful discussions and critical reading of this manuscript.

References

[1] M. Jeannerod, The Representing Brain: Neural Correlates Of Motor Intention And Imagery, Behav. Brain Sci. 17 (1994) 187– 245.

[2] John P. Donoghue, Connecting cortex to machines: recent advances in brain interfaces http://www.nature.com/natureneuroscience 2002 Nature Publishing Group

[3] M Cheng, WY Jia, et al. Mu Rhythm-Based Cursor Control: An Offline Analysis. Clinical Neurophysiology, 115(4): 745–751, 2004.

[4] Wang Y, Berg P, Scherg M. Common Spatial Subspace Decomposition Applied To Analysis Of Brain Responses Under Multiple Task Conditions: asimulation study. Clin Neurophysiol 1999; 110: 604–14.

[5] Sebastian Mikat, Gunnar fitscht, et al Fisher Discriminant Analysis With Kernels 0-7803-5673-X/99/$10.00 01 999 IEEE

[6] TN Lal, T Hinterberger, G Widman, et al. Methods Towards Invasive Human Brain Computer Interfaces. http://books.nips.cc/papers/files/nips17/2004_0443.pdf

[7] IEEE BCI Competition III, http://ida.first.fraunhofer.de/projects/bci/competition_iii

[8] COVERTM, HARTPE Nearest neighbor pattern classification [J] In Trans IEEE Inform Theory, 1967, IT-13: 21-27.

[9] S Shoham, E Halgren, EM Maynard, et al. Motor-cortical activity in tetraplegics. Nature, 413: 793, 2001.

[10] SP Levine, JE Huggins, et al. A Direct Brain Interface Based on Event-Related Potentials. IEEE Trans Rehab Eng, 8(2): 180-185, 2000.

[11] SC O'Connor, PA Robinson. Wavenumber power spectrum of the EED, ECoG, and ERP. Neurocomputing, 58: 1181-1186, 2004.

[12] H Ramoser, J Müller-Gerking, G Pfurtscheller. Optimal Spatial Filtering of Single Trial EEG during Imagined Hand Movement. IEEE Trans Rehab Eng, 8(4): 441-446, 2000.