35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Adaptive Power Projection Method for Accumulative EEG Classification*

Chun-yue Li, Rong Liu, Yuan-yuan Wang, Yong-xuan Wang, Xiang Li

***Abstract*—For the dynamic classification of motor imagery mind states in the brain-computer interface (BCI), we propose a power projection based feature extraction method to classify the electroencephalogram (EEG) signals by combining information accumulative posterior Bayesian approach. This method improves the classification accuracy by maximizing the average projection energy difference of the two types of signals. The experimental results on two BCI competition datasets show that the classification accuracy is about 90%. The results of the classification accuracy and mutual information demonstrate the effectiveness of this method.**

I. Introduction

A brain-computer interface (BCI) [1] is a system capable of utilizing the brain's electrical signals for direct communication with a computer system, without reliance on the usual neuromuscular pathways. It has been widely used in many fields, such as neurological rehabilitation projects, games, military, brain cognitive, and so on [2].

Many studies have shown that when the subject imagines limb movement, specific frequency components of electroencephalogram (EEG) such as the mu (8-13Hz) and central beta (14-22Hz) rhythms are (de)synchronized over the contralateral (ipsilateral) sensorimotor area [3-5]. This phenomenon is called event-related desynchronization/ synchronization (ERD/ERS). The motor imagery EEG is widely applied in BCI system because it is only generated by movement imagination does not depend on any sensory stimulation. Since 1992, Pfurtscheller's team has designed a motor imagery-based BCI to control the cursor movement [6, 7].

The critical challenge of BCI technology is to classify the brain signals and mental tasks accurately. However, the EEG recorded from the scalp has the characteristics of low strength, low SNR (signal noise ratio), and the EEG difference under different mental tasks is not significant. Therefore, various pattern recognition algorithms were used in BCI system to extract and classify EEG features. Currently, feature extraction for discrimination of left and right hand motor imagery EEG is usually based on EEG band power (BP). For example, autoregression (AR) model [8], discrete Fourier Transformation (DFT) [9] and wavelet transforms (WT) [10] have been used to extract EEG features for classification. In these methods, the characteristic bands are selected by experience. However, the experiments show that the phenomenon of ERD/ERS varies among individuals. Consequently, it's difficult to achieve the best classification results by adopting fixed bands.

In this paper, we propose a power projection (PP) based feature extraction method to classify the EEG signals by combining information accumulative posterior Bayesian approach. This method improves the classification accuracy by maximizing the average projection energy difference of the two different signal classes (right hand motor imagery EEG and left hand motor imagery EEG). We test our method on BCI Competition 2003 Dataset III and BCI Competition 2005 Dataset IIIb. Comparing with the traditional feature extraction methods such as DFT and WT, it has higher classification accuracies for all the subjects.

The rest of this article is organized as follows: Section II describes the PP method; Section III describes Bayesian posterior accumulative classification approach to classify feature vectors. Section IV presents our experimental results, and Section V gives concluding remarks.

II. PP Feature Extraction

In this section, we introduce the PP method for feature extraction. Let $X_c \in R^{M\times N_c}$ be the training dataset from a channel C_3 or C_4 [11], where M denotes the sampling points, N_c denotes the number of trials and $c=\{L,R\}$ denotes the left or right hand motor imagery tasks.

Let $\boldsymbol{u}\in R^M$ be the projection basis which has $\|\boldsymbol{u}\|=1$. The projection power of signal $x_{cj}, j=1,2,\cdots,N_c$ on $\boldsymbol{u}$ is

$$e_{cj}=(x_{cj}{}^T\cdot \boldsymbol{u})^2. \tag{1}$$

The mean projection power $\overline{e_c}$ of two training sets on $\boldsymbol{u}$ can be calculated as follow

$$\overline{e_c}=\boldsymbol{u}^T\cdot \boldsymbol{X}_c\cdot \boldsymbol{X}_c{}^T\cdot \boldsymbol{u}/N_c=\boldsymbol{u}^{\mathrm{T}}\cdot \boldsymbol{R}_c\cdot \boldsymbol{u}\quad, \tag{2}$$

---

*This work was supported in part by the grants from the Natural Science Foundation of China (NSFC) (61005088), the State Key Laboratory of Rob otics and System (HIT) (SKLRS-2010-ZD-07) and the Fundamental Resear ch Funds for the Central Universities (DUT10JS03), National key technolog y support program (2012BAJ18B06).

Chun-yue Li and Yuan-yuan Wang are with the Biomedical Engineering Department, Dalian University of Technology, Dalian, Liaoning 116024 P.R.C. (e-mail: Lichunyue@mail.dlut.edu.cn; wangyuan.y@mail.dlut.edu.cn; wyx8904@mail.dlut.edu.cn)

Rong Liu is with the Biomedical Engineering Department, Dalian University of Technology, Dalian, Liaoning 116024 P.R.C. and the State Key Laboratory of Robotics and System (corresponding author: phone: 86-411-84706002-3015; e-mail: rliu@dlut.edu.cn).

Yong-xuan Wang is with the Affiliated Zhongshan Hospital of Dalian University, Dalian, Liaoning 116001 (e-mail: wyx8904@mail.dlut.edu.cn)

Xiang Li is with the department of Radiology, the Second Hospital of Dalian Medical University, Dalian, Liaoning 116023 P.R.C (email: lixiang_5007@163.com)

where $\boldsymbol{R}_c = \boldsymbol{X}_c \cdot \boldsymbol{X}_c^{T} / N_c$ is the autocorrelation matrix and it is usually positive definite. Then the ratio of mean projection power $F(\boldsymbol{u})$ can be obtained

$$F(\boldsymbol{u}) = \frac{\overline{e_L}}{\overline{e_R}} = \frac{\boldsymbol{u}^{\mathrm{T}} \cdot \boldsymbol{R}_L \cdot \boldsymbol{u}}{\boldsymbol{u}^{\mathrm{T}} \cdot \boldsymbol{R}_R \cdot \boldsymbol{u}}. \tag{3}$$

By maximizing and minimizing $F(\boldsymbol{u})$ to be $F_{\max}$ and $F_{\min}$, the corresponding eigenvectors $\boldsymbol{u}_{\max}$ and $\boldsymbol{u}_{\min}$ are the required bases pair. The optimization of (3) could be solved by taking a generalized eigenvalue decomposition method,

$$\boldsymbol{R}_L \cdot \boldsymbol{U} = \boldsymbol{R}_R \cdot \boldsymbol{U} \cdot \boldsymbol{\Lambda}. \tag{4}$$

After contract diagonalization by $\boldsymbol{U}$, the $\boldsymbol{R}_L$ and $\boldsymbol{R}_R$ in (4) turn to,

$$\begin{aligned} &\boldsymbol{U}^{\mathrm{T}} \cdot \boldsymbol{R}_L \cdot \boldsymbol{U} = \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_{\mathrm{M}}), \lambda_{\mathrm{i}} \ge \lambda_{\mathrm{j}} > 0 (\forall i < j), \\ &\boldsymbol{U}^{\mathrm{T}} \cdot \boldsymbol{R}_R \cdot \boldsymbol{U} = \boldsymbol{I}. \end{aligned} \tag{5}$$

Then we have

$$\begin{aligned} F(\boldsymbol{u}) &= \frac{\boldsymbol{u}^{\mathrm{T}} \cdot \boldsymbol{R}_L \cdot \boldsymbol{u}}{\boldsymbol{u}^{\mathrm{T}} \cdot \boldsymbol{R}_R \cdot \boldsymbol{u}} \\ &= \frac{\boldsymbol{u}^{\mathrm{T}} \cdot \boldsymbol{U}^{-\mathrm{T}} \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{U}^{-1} \cdot \boldsymbol{u}}{\boldsymbol{u}^{\mathrm{T}} \cdot \boldsymbol{U}^{-\mathrm{T}} \cdot \boldsymbol{U}^{-1} \cdot \boldsymbol{u}} \\ &= \frac{\boldsymbol{v}^{\mathrm{T}} \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{v}}{\boldsymbol{v}^{\mathrm{T}} \cdot \boldsymbol{v}} \\ &= \frac{\lambda_1 \cdot \boldsymbol{v}_1^{2} + \cdots + \lambda_M \cdot \boldsymbol{v}_M^{2}}{\boldsymbol{v}_1^{2} + \cdots + \boldsymbol{v}_M^{2}}, \end{aligned} \tag{6}$$

where $\boldsymbol{v} = \boldsymbol{U}^{-1} \cdot \boldsymbol{u}$. Obviously, the following expressions can be obtained by

$$\begin{aligned} F_{\max} &= F(\boldsymbol{U} \cdot [1, 0, \cdots, 0]) \triangleq F(\boldsymbol{u}_{\max}) = \lambda_1, \\ F_{\min} &= F(\boldsymbol{U} \cdot [0, \cdots, 0, 1]) \triangleq F(\boldsymbol{u}_{\min}) = \lambda_M. \end{aligned} \tag{7}$$

After obtaining two projection base pairs from electrodes C3 and C4, the projection power for the two electrodes are then stacked together into the 2-dimensional feature vector according to (1)

$$\boldsymbol{z} = [e_{c3}, e_{c4}]^{\mathrm{T}}. \tag{8}$$

In order to use the Bayesian posterior classification, the distribution of feature vectors should be obtained. Based on the Gaussian distribution assumption, the mean vector $\boldsymbol{\mu}_c$ and covariance matrix $\boldsymbol{S}_c$ are expressed as

$$\begin{aligned} \boldsymbol{\mu}_c &= E(\boldsymbol{z}_c) \\ &= \sum_{j=1}^{N_c} \boldsymbol{z}_{cj} / N_c, \\ \boldsymbol{S}_c &= E\left((\boldsymbol{z}_c - \boldsymbol{\mu}_c) \cdot (\boldsymbol{z}_c - \boldsymbol{\mu}_c)^{\mathrm{T}}\right) \\ &= \sum_{j=1}^{N_c} (\boldsymbol{z}_{cj} - \boldsymbol{\mu}_c) \cdot (\boldsymbol{z}_c - \boldsymbol{\mu}_c)^{\mathrm{T}} / N_c. \end{aligned} \tag{9}$$

The probability density function (PDF) is

$$\begin{aligned} f_c(\boldsymbol{z}) = f(\boldsymbol{z} \mid \boldsymbol{\mu}_c, \boldsymbol{S}_c) &= (2\pi)^{-2} \\ \cdot \left|\boldsymbol{S}_c^{-1}\right|^{\frac{1}{2}} \cdot \exp&\left(-\frac{1}{2}(\boldsymbol{z} - \boldsymbol{\mu}_c)^{\mathrm{T}} \boldsymbol{S}_c^{-1} (\boldsymbol{z} - \boldsymbol{\mu}_c)\right). \end{aligned} \tag{10}$$

III. Bayesian Posterior Accumulative Classification Approach

Bayesian posterior accumulative classification approach is used to classify two classes of motor imagery EEG signals. To realize sequential prediction, each trial is divided into certain number of segments. The feature vectors for each segment are computed with the PP method. The distribution of eigenvectors is expressed by $f_{ci}(\boldsymbol{z}), i = 1, 2, \cdots, D$, where D is the total number of segments. The Bayesian posterior probability $p_i(c \mid \boldsymbol{z}_i)$ for a single segment is

$$p_i(c \mid \boldsymbol{z}_i) = \frac{f_{ci}(\boldsymbol{z}_i)}{f_{Li}(\boldsymbol{z}_i) + f_{Ri}(\boldsymbol{z}_i)}, c \in \{L, R\}. \tag{11}$$

In order to derive the online classification at the d_0^{th} segment, we incorporate knowledge from all preceding time segment $d < d_0 \le D$, leading to an evidence accumulation over time about the binary decision process. The temporal combination is then realized by taking the expectation of the class probabilities from (10) with respect to the discriminative power k_i at each segment,

$$p_{d_0}(c \mid \boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_{d_0}) = \frac{\sum_{i=1}^{d_0} k_i \cdot p_i(c \mid \boldsymbol{z}_i)}{\sum_{i=1}^{d_0} k_i}. \tag{12}$$

The discriminative power is estimated by using the Chernoff bound on the Bayes classification error,

$$k_i = \frac{1}{2} \cdot \left[1 - \min_{0 \le \beta_i \le 1} \int f_{Li}^{\beta_i}(\boldsymbol{z}) \cdot f_{Ri}^{1-\beta_i}(\boldsymbol{z}) \mathrm{d}\boldsymbol{z}\right]. \tag{13}$$

Since the proposed method is based on the Gaussian assumption, the Chernoff bound can be easily estimated at each segment during a trial. Finally, the discriminative power k_i between the two distributions can be approximated by

$$k_i = \frac{1}{2}\cdot\{1-\min_{0\le\beta_i\le 1}\exp[-\frac{1}{2}\cdot(\ln\frac{|\beta_i \boldsymbol{S}_{Ri}+(1-\beta_i)\boldsymbol{S}_{Li}|}{|\boldsymbol{S}_{Ri}|^{\beta_i}|\boldsymbol{S}_{Li}|^{1-\beta_i}} + \beta_i(1-\beta_i)(\boldsymbol{\mu}_{Li}-\boldsymbol{\mu}_{Ri})^{\mathrm{T}}(\beta_i\boldsymbol{S}_{Ri} + (1-\beta_i)\boldsymbol{S}_{Li})^{-1}(\boldsymbol{\mu}_{Li}-\boldsymbol{\mu}_{Ri}))]\}. \tag{14}$$

Let $p_{diff} = p_{d_0}(L\,|\,z_1,z_2,\cdots,z_{d_0}) - p_{d_0}(R\,|\,z_1,z_2,\cdots,z_{d_0})$, the final decision of this segment is

$$y_{d_0} = \begin{cases} L, & if\ p_{diff} > 0.5, \\ R, & if\ p_{diff} < -0.5, \end{cases} \tag{15}$$

where 0.5 reflects the confidence degree.

IV. Experimental Data and Results

To evaluate the performance of our method, we tested it on the four subjects EEG data from BCI Competition data. The task performed was based on left and right hand motor imagination.

1) Dataset III from BCI competition II [12]: contains EEG data from one subject (S2003). The data were recorded from three channels (C3, Cz and C4) and sampled at 128Hz. The data consist of 140 labeled and 140 unlabeled trials with an equal number of left and right hand trials. Each trial has a duration of 9s, where a visual cue (arrow) is presented pointing to the left or the right after 3s preparation period followed by a 6s motor imagery (MI) task.

2) Dataset IIIb from BCI competition III [13]: contains EEG data from three subjects. The data were recorded from two bipolar channels (C3, C4) and sampled at 125Hz. A training and testing set were available for each subject. Except for the subject O3 has only just 320 trials for each set, the subject S4 and X11 contain 540 labeled and 540 unlabeled trials. Every trial has duration of 7s: after a 3s preparation period, a visual cue is presented for 1s, indicating the requested motor intention. This is followed by another 3s for the imagination task.

We compared the proposed PP feature extraction method with two popular algorithms, DFT and WT, on these two datasets. In this study, the window length of the projection base is set to be 1s. The time-domain waveforms of the optimal projection base pairs of the two electrodes for the subject S2003 are shown in Fig. 1(a) and Fig. 1(b). The solid line and the dashed line represent a pair of complementary projection bases with $\pi/2$ phase difference. The corresponding frequency spectrums are shown in Fig. 1(c) and Fig. 1(d). From this figure, we can see the characteristics of the projection bases in time domain are very similar to sine/cosine signals, and the spectral features are similar to wavelet base. The spectrograms show that the band-pass characteristics of the projection bases dominate in the μ rhythm.

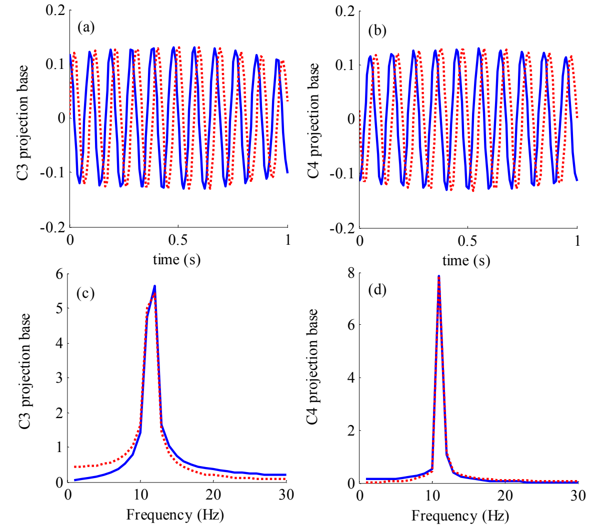

Figure 1. The adaptive projection base of subject S2003. (a) waveform of C3; (b) waveform of C4; (c) spectrum of C3; (d) spectrum of C4.

The time-varied average projection power for subject S2003 during imagined movement of the right (red line) and left hand (blue line) for the C3 and C4 electrodes are displayed in Fig. 2. Starting with cue presentation (t=3s), subject S2003 displayed a transient contralateral desynchronization and an ipsilateral synchronization.

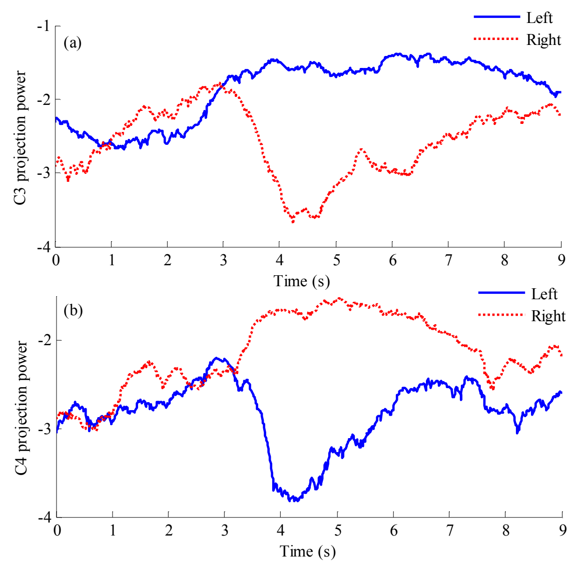

Figure 2. The time-varied projection power of subject S2003during imagined movement of the right (red line) and left hand (blue line) for the C3 and C4 electrodes.

The time courses of discriminative power and accumulative process of classification information for S2003 are shown in Fig. 3. Fig. 3(a) shows that the discriminative power increases significantly from 4s and reach the peak at 5s then fall gradually. Therefore, the discriminative power enhances the effect of the information at the mid trial and decreases the impact of information at the beginning and end of the trial. The accumulative process of classification information also illustrates this phenomenon. Fig. 3(b) shows that the Bayesian posterior accumulative classification approach gains information due to the integration process from 4s. The cumulative Bayesian posterior probabilities will reach to extrema at around 5s, indicating a peak decision confidence at this time. However, the accumulative information will fall down at the end of trial. The result shows that the effective control takes place during the middle of a trial.

The classification accuracy (ACC) and mutual information (MI) of these methods are listed in Table I, where Avg. denotes the averaged indexes over all four subjects. The results of WT are derived from Lemm's method which won the BCI competition 2003 and 2005 for motor imagery dataset. From Table I, we can see that the proposed method outperforms all other methods on every subject consistently. Compared with Lemm's method, the average MI of our method increased 11%, from 0.577 to 0.642 bit.

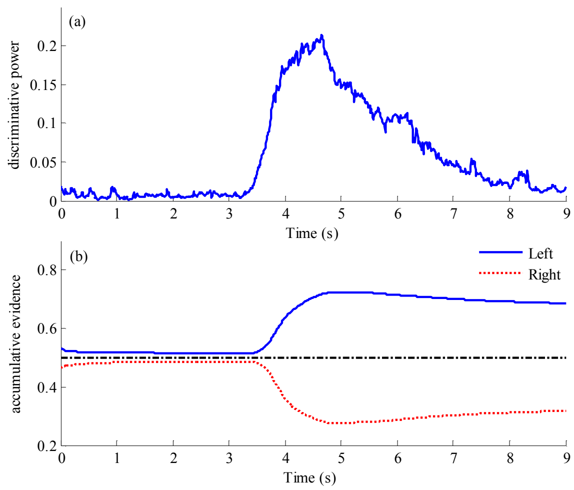

Figure 3. The time courses of average discriminative power and the average accumulative process of classification information for subject S2003. (a) discriminative power; (b) the average accumulative process of classification information.

TABLE I. A COMPARISON OF ACC AND MI FOR THREE METHODS

Methods	subjects	ACC (%)	MI(B)
DFT	S2003	92.1	0.612
	O3	91.1	0.568
	S4	74.1	0.175
	X11	81.3	0.308
	Avg.	84.7	0.416
WT	S2003	89.3	0.610
	O3	89.3	0.602
	S4	88.5	0.608
	X11	83.3	0.486
	Avg.	87.6	0.577
PP	S2003	93.2	0.643
	O3	95.2	0.724
	S4	89.7	0.621
	X11	91.5	0.579
	Avg.	92.4	0.642

V. CONCLUSION

In this paper, we propose a power projection based feature extraction method to classify the EEG signals by combining information accumulative posterior Bayesian approach. This method improves the classification accuracy by maximizing the average projection energy difference of the two types of signals. The results show that the method could effectively improve the performance of BCI system and have good practicability.

REFERENCES

[1] C. Brunner, M. Billinger, C Vidaurre, C. Neuper, "A comparison of univariate, vector, bilinear autoregressive, and band power features for brain-computer interfaces (Periodical style)," *Med Biol Eng Computer*, vol. 49, pp. 1337-1346, 2011.

[2] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, "Brain-computer interfaces for communication and control (Periodical style)," Clin. Neurophysiol. vol. 113, pp. 767-791, 2002.

[3] G. Pfurtscheller, C. Neuper, D. Flotzinger, M. Pregenzer, "EEG-based discriminination between imagination of right and left hand movement (Periodical style)," *Electroencephalogr Clin Neurophysiol*, vol. 103, pp. 642-651, 1997.

[4] G. Pfurtscheller, C. Neuper, "Motor imagery and direct brain-computer communication (Periodical style)," *Proc IEEE,* vol. 89, pp. 1123-1134, 2001.

[5] G. Pfurtscheller, C. Brunner, A. Schlogl, F.H.L. daSilva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks (Periodical style)," Neuroimage, vol. 31, pp. 153-159, 2006.

[6] J. R. Wolpaw, D. J. McFarland, G. W. Neat and C. A. Forneris, "An EEG based brain-computer interface for cursor control (Periodical style)," *Electroencephalogr Clin Neurophysiol*, vol. 78, pp. 252-259, 1991.

[7] M. A. Lebedev, and A. L. Nicolelis, "Brain-machine interfaces: past, present and future (Periodical style)," Trends in Neurosciences, vol. 20, pp. 536-549, 2006.

[8] Liu R, Newman G, and Thakor NV. "Improved BCI performance with sequential hypothesis testing (Presented Conference Paper style)," in *Proc. of the 33rd IEEE Eng Med Biol Soc.*, USA: IEEE, 2011, pp, 4215-4218.

[9] M. Zhong, F. Lotte, M. Girolami, and A. Lecuyer, "Classifying EEG for brain computer interfaces using gaussian processes (Periodical style)," *Pattern Recognition Letters*, vol. 29, pp. 354-359, 2008.

[10] S. Lemm, C. Shafer, and G. Curio, "Aggregating classification accuracy accross time: application to single trial EEG (Periodical style)," *Advances in Neural Information Processing Systems*, vol. 19, pp. 825-832, 2007.

[11] D. McFarland, L. Miner, T. Vaughan, and J. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements,"*Brain Topogr.,* vol. 12, no. 3, pp. 177-186, 2000.

[12] B. Blankertz, K. R. Müller, G. Curio, T. M. Vaughan, et al., "The BCI competition 2003 set III: probabilistic modeling of sensorimotor μ rhythms for classification of imaginary hand movements (Periodical style)," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1044-1051, 2004.

[13] B. Blankertz, K. R. Müller, D. J. Krusienski, et al., "The BCI competition III: Validating alternative approaches to actual BCI problems (Periodical style)," *IEEE Trans. Neural Syst. Rehabil*, vol. 14, pp. 153-159, 2006.