A Review on Spatial Pattern Methods for Brain-Computer Interface

[1][2] Department of Electronics and Communication Engineering
NSS College of Engineering, Palakkad, Kerala, India
[1] bijithabalakrishnan@gmail.com [2] nanakumarpp@hotmail.com

Abstract: Brain-computer interface technology is the one in which the user’s mental intentions are captured and used as control signals for external devices. BCI researches are becoming more popular as it can be used as a means of communication for people with severe disability. The efficiency of a BCI system depends on its signal processing stages. The signal processing stages include the signal acquisition, feature extraction and classification. In this paper, a brief review of one of the feature extraction methods called spatial pattern algorithms for BCI is provided.

Index Terms- brain computer interfaces, common spatial pattern, feature extraction.

I. INTRODUCTION

Brain computer interfaces is also called as Brain machine interfaces (BMI). Researches focus on this technology in the hope that it will be helpful for people with motor disability. Wolpaw et al. [1] have defined BCI as “A BCI is a communication system in which messages or commands that an individual sends to the external world do not pass through the brain’s normal output pathways of peripheral nerves and muscles”. Wolpaw [2] describes that the proper objective of BCI development is to find the signals that the users can control, maximize that control, and translate it into action reliably and accurately. BCI technology has various applications other than assisting disabled people, like environment control [3], entertainment [4], etc. A BCI can be seen as a pattern recognition system. BCI system consists of various stages as signal acquisition, preprocessing, feature extraction and classification. The BCI technology can be classified into different types: according to nature of input signals used it can be classified into exogenous (external stimuli are used) and endogenous (based on the self regulation of brain signals without external stimuli); according to the data processing technique it can be classified into synchronous (brain signals are analyzed in a pre-defined window) and asynchronous (brain signals are analyzed continuously) [5]. Wolpaw et al. [1] have classified BCI into dependant (does not use brain’s normal output pathways to carry the message but the activity in these pathways are used to generate brain activity that carries it) and independent (does not depend on any way on the normal output pathways).

II. BCI TECHNOLOGY

BCI has various stages mainly a signal acquisition part, preprocessing, feature extraction and the classification stage [6]. The brain activity can be measured using various recording methods like EEG, fMRI, EC0G, MEG, PET [7]. These methods are of two types namely, invasive (electrodes are implanted inside the brain) and noninvasive (include haptic controllers and EEG scanners) [5]. In BCI applications the various brain signal types that are used includes the event-related potentials, mu and beta rhythms, event related synchronization/desynchronization, visual evoked potentials and slow cortical potentials [8, 9]. Features need to be extracted from the brain signals and they need to be classified for controlling the external devices [10]. There are various methods for feature extraction as CSP, wavelet transform, autoregressive model [11], principal component analysis(PCA), and independent component analysis(ICA) [12]. Feature translation is the next important signal processing stage in BCI. Wolpaw et al. [13] have specified that the translation algorithm can be based on either on a classifier or a regression function. Regression analysis is a statistical tool for determining relationships between different variables. There are mainly five types of classifier algorithms: linear, nonlinear Bayesian, nearest neighbor, neural networks and a combination of classifiers [14, 15].
III. COMMON SPATIAL PATTERN

Brain computer interfaces can be seen as a pattern recognition system in which the brain activities are converted into control commands for external devices like computer [10]. In order to achieve this, certain features like band powers or power spectral density values need to be extracted from the acquired brain signal [10]. The amplitude modulations of sensorimotor rhythms can be obtained from time/frequency analysis of the EEG signals. The frequency based analysis is mostly done due to its simplicity and efficiency [16]. However frequency features does not provide any time domain information. Also because of the non-stationarity nature of EEG signal [17], time-frequency representations are mostly used [16]. Besides the temporal (time domain) and spectral (frequency domain) features, features related to spatial locations are also important for discriminant feature extraction [16, 18].

A. Importance of Spatial Filtering

The EEG signals recorded from electrodes have not only the neural voltage fluctuations but also have the voltages corresponding to the activities of distant current sources through volume conduction [18]. If EEG signal is represented as $x(t)$ which are generated by brain sources $s(t)$, then the EEG signal can be represented as [19],

$$x(t) = As(t) + n(t),$$

where $A$ is a matrix mapping the activity of each source to the electrode space. The sensory motor rhythms will be attenuated by the movement imagination in the corresponding cortical areas (e.g., left hand motor imagery affects the SMRs over the right motor cortex area) [18, 20]. Hence it becomes necessary to find the sources of the SMR modulation to discriminate between different motor imagery tasks and this can be achieved with spatial filtering as given below

$$s(t) = W^T x(t)$$

where $W = [w_1, w_2, ..., w_d]$ projects the EEG signal to a lower dimension subspace[18] and $w_i$ is a spatial filter which weighs the electrode to get information about the sources.

B. Common Spatial Pattern

CSP is an algorithm to develop the spatial filters for the motor imagery experiments [18]. Haixian et al. [21] has explained CSP in both geometrical and mathematical ways. The method can be explained in geometric basis that it is used to find directions in which the projected scatters between two EEGs are maximized. Mathematically this can be obtained by simultaneously diagonalizing two covariance matrices associated with two EEG populations (or classes) in order to maximize the difference between the two projected populations [21]. The spatial filtering is done by the CSP by linearly transforming the EEG measurements using the following equation,

$$Z = W^T E$$

where $E$ is the EEG measurement, $Z$ is the EEG measurement after spatial filtering and $W$ is the CSP projection matrix [22]. The main function of the CSP algorithm is to find the projection matrix $W$, so that features that can discriminate two classes can be obtained [22]. The projection matrix is obtained by solving the eigenvalue decomposition problem

$$\Sigma_1 w_i = \lambda_i \Sigma_2 w_i$$

where $\Sigma_1$ and $\Sigma_2$ are the covariance matrices of EEG measurements corresponding to two classes and generalized eigenvalues $\lambda_i$ measure the variance ratio between class 1 and class 2 [18]. In a nutshell, CSP finds spatial filters that maximize the variance for one class or population and simultaneously minimizing for the other [23]. CSP can also formulated as a Rayleigh quotient problem [24].

C. Different CSP Variants

Besides the CSP algorithm provides better discrimination it has limitations too. Some of the limitations are that the covariance estimation can be effected by the artifacts in EEG, overfitting problem with the small training sets, the nonstationarity problem, etc. [18, 21]. In order to overcome various limitations of the standard CSP various methods have been developed by researchers.

The projection matrix $W$ remains unchanged within or across sessions, on the other hand the EEG signal changes within time [25]. Hence the direction obtained from projection matrix may not be always optimal. Chen et al. [25] proposed an adaptive method to overcome this problem. Here the covariance matrices are adapted to the varying EEG signals and then the eigenvalue decomposition problem is solved with this adapted covariance matrices. The results they obtained for the adaptive CSP were slightly better than that for traditional CSP.

The standard CSP method is a supervised learning method (requires labeled EEG trials). It is difficult to get large number of labeled information at all the times and if small amount of trials is used then overfitting problem arises which leads to less efficient results [24]. On the other hand it is easy to get unlabeled information. Thus Wang et al. [24] proposed the comprehensive learning scheme of CSP (cCSP) that combines both the labeled and the unlabeled trials. The information from unlabeled trials was incorporated with the help of $\ell_1$ graph. In the objective
function of the comprehensive learning scheme of CSP a regularization term was also included to encode the knowledge of intrinsically temporal structure of the unlabeled EEG trials [24]. Wang et al. [24] obtained better performance of the proposed method for single-trial EEG classification.

The imagery classifications are done based on the changes in the mu and the beta rhythms and corresponding spatial distributions but these rhythmic changes will vary from subject to subject [26]. Hence fine tuning process will be required. Novi et al. [26] have proposed a method called sub-band common spatial pattern (SBCSP) to solve this problem. In this case they have decomposed the EEG signals into various sub-bands using a filter bank and CSP is performed on each sub-band [26]. Novi et al. [26] fused the scores from each sub-band and decision was made. They compared this method with the CSSP (common spatio-spectral pattern [27]) and CSSSP (common sparse spatio-spectral pattern [28]) method and found that the proposed method outperformed the two. The performance of CSP is dependent on the frequency band, hence setting up a broad frequency range or manual selection of subject specific frequency range [29]. The CSSP [27] and CSSSP [28] were proposed to solve the problem of manual selection of the frequency range. Ang et al. [29] proposed filter bank common spatial pattern, in which there are four stages as band pass filtering, feature extraction, feature selection and classification. FBCSP was compared with the SBCSP and the result outperformed that of SBCSP. In FBCSP fixed filter bank was used by all the subjects. Thomas et al. [30] proposed subject specific discriminative FB (DFBCSP), in which the parent FB filters the EEG from one channel (C3 or C4) and then fisher ratio of the filtered EEG signal was used to determine the subject specific discriminative frequency bands and they obtained successful results. In FBCSP there involves multiple spatial filtering which requires multiple estimations of the covariance matrices and thus this increases the sensitivity of the FBCSP to noise, artifacts and outliers. Ang et al. [31] proposed a method called composite FBCSP that employs a single spatial filter (computed from a weighted sum of covariance matrices) instead of multiple spatial filters. The results obtained had a better kappa values as compared to FBCSP [31]. Aghaei et al. [32] proposed a novel method called separable common spatio-spectral patterns (SCSSP) where a heteroscedastic matrix-variate Gaussian model and was found to be computationally efficient and outperformed the FBCSP method.

Hyohyeong et al. [33] provides modified CSP for subject to subject transfer (transfer useful information of subjects involving the same task to the subject with lower training samples). They exploited the composite covariance matrices determined by a linear combination of covariance matrices for all subjects in consideration. Two different methods were proposed to determine appropriate weights in evaluating composite covariance matrices: Method 1 de-emphasized covariance matrices involving fewer samples, while Method 2 emphasized covariance matrices for subjects with similar characteristics to the subject in consideration [33]. The method proposed by Hyohyeong et al. [33] worked well for subjects with a small number of samples, while the traditional methods worked better for subjects with a sufficient number of training samples.

One of the important objectives of the BCI research is to reduce the number of training trials needed and at the same time conventional CSP algorithm is based on the sample-based covariance matrix estimation and the accuracy of estimation is badly affected if only small training set is available [34]. In order to tackle this problem Haiping et al. [34] proposed regularized CSP, to regularize the covariance estimation in CSP; two parameters have been employed, one to lower the estimation bias and the other to lower the estimation variance. Later Fabien et al. [35] reviewed several regularization techniques for CSP and also proposed four RCSP algorithms. They got the best results from the Tikhonov regularization method, which is a classical form of regularization generally used with regression problems [35].

In EEG based BCI technology, to get better performance signals from multiple sites of the scalp are required and at the same time large number of EEG channels may increase noise, redundant signals and long preparation time [36]. Therefore, there is a need to select least number of channels that yields best accuracy. Arvaneh et al. [36] proposed a novel sparse CSP (SCSP) for optimal EEG channel selection. The method outperformed the CSP in the case where the number of selected channels was small. Wang et al. [37] proposed L1 based CSP to make it less sensitive to outliers. In standard CSP the formulation is based on variance using L2 norm, and in the proposed method in [37] the formulation of variance is based on L1 norm. The spatial filters in this case are obtained through an iterative algorithm and Wang et al. [37] obtained better results compared to other CSP variants. Kam et al. [38] proposed time dependent CSP inorder to overcome the limitation of standard CSP that it does not consider the temporal information of EEG signals. The standard CSP is usually used for two-class problems, however its extension to multiclass problems are also studied [39]. Wentrup et al. [39] have mentioned various multiclass extensions of CSP and have proposed an information theoretic feature extraction method for multiclass.
IV. CONCLUSION

The common spatial pattern method had been widely used in brain-computer interface technology for feature extraction. Though it provides discrimination of EEG data, its performance are degraded by the artifacts in EEG data, the nonstationarity problem, overfitting problem that arises due to availability of small training data sets. Hence inorder to overcome these limitations of the basic CSP algorithm different variants have been developed by various researchers.

REFERENCES


[18]. Wojciech Samek, Motoaki Kawanabe, and Klaus-Robert Muller, “Divergence-Based Framework for


