Selection of EEG Channels based on Spatial Filter Weights

Naveen Masood
Department of Electrical Engineering
Bahria University
Karachi, Pakistan
navinza@hotmail.com

Humera Farooq
Department of Computer Science
Bahria University
Karachi, Pakistan
humera.farooq@bimcs.edu.pk

Irfan Mustafa
Department of Computer Science
Bahria University
Karachi, Pakistan
irfanmustafas@yahoo.com

Abstract— Most of the Electroencephalography (EEG) based brain computer interfaces (BCIs) use large number of channels to capture the signals from subject's brain. This is one of the major issues in commercial and out of the lab usage of such systems. Common Spatial Pattern (CSP) algorithms are widely used for feature extraction in BCI systems for motor imagery. As the EEG signals have noise and overfitting issues, various regularized CSP algorithms are introduced to overcome these factors. In this work, a method is introduced to find the variant of CSP that achieves maximum classification accuracy with least number of EEG channels. The approach is based on firstly identify the spatial filter weights using complete set of channels. In the next step, the channels are selected based on the maximal filter weights. Channels with highest values are included to find the accuracy in next run and ultimately the optimum combination of number of channels, CSP variant and classification accuracy are reported. Multichannel data comprised of 60 electrodes from BCI Competition III dataset IIIa from three subjects who performed left hand, right hand, foot and tongue MI is considered for this purpose. For the data analyzed in this study, it is found that instead of using data from 60 electrodes only six can be used without significantly compromising the classification accuracy.

Keywords —Brain Computer Interfacing, common spatial pattern, Motor Imagery, Electroencephalography, classification

I. INTRODUCTION

Any system that captures the signals from brain and able to convert them into usable commands is termed as Brain Computer Interface (BCI) [1]. In order to record the brain signals, various modalities are being used. Out of the available modalities, electroencephalography (EEG) is the most common one [1]. Most of the BCI studies make use of large number of EEG sensors to record brain activity. Motor imagery can be defined as any brain activity during which the subject is involved in imagination of doing any movement like imagination of right hand movement, foot movement etc.

The problem that will be addressed in this paper related to the selection of EEG channels responsible for providing most relevant features to achieve improved classification accuracies for specified brain activities. For EEG based BCI devices, using large number of sensors involves a laborious process of electrode placement as well as higher computational time. As specific regions of brain are responsible for specific brain activity, signals from all areas of brain may introduce the issue of overfitting and noise. It is also reported that persons respond differently for same mental task therefore subject-specific optimal placement of EEG channel may vary [1].

During the past many years, researchers have focused on finding the optimal set of EEG sensors for Motor imagery based BCI systems. He et atl. [2] used statistical method based on Bhattacharyya bound of common spatial patterns (CSP) to select the relevant EEG channels for motor imagery data from 59 sensors. Bayes algorithm is used for classification purpose. Out of 59, the authors reported 33 average number of channels having classification accuracy up to 95%. In another study He et al [3] used Genetic Algorithm based on Rayleigh Coefficient maximation for electrode selection on two different EEG datasets comprising of 118 and 32 respectively. For the two datasets high classification accuracies are achieved using reduced set of channels obtained from proposed methodology for channel selection. Wang et al [4] presented the channel reduction approach based on utilizing the maximum values of spatial pattern vectors obtained from CSP filter matrices. It was reported that using only four channels classification accuracies up to 93.45 and 91.88% were achieved.

In this work we aim to validate one of the phases of proposed framework in our previous work to assess the neural biomarkers for Neuroception [5] that specifically focuses towards the problem of finding the neural correlates for fear-induced user brain activity. We proposed that CSP algorithm will be used to perform feature extraction. In this work a method is introduced to find the variant of CSP that achieves maximum classification accuracy with least number of EEG channels. Instead of using any other technique, CSP is itself used to calculate optimum number of reduced channels for given data set. Multichannel data comprised of 60 electrodes from BCI Competition III dataset IIIa from three subjects who performed left hand, right hand, foot and tongue Motor Imagery (MI) is considered for this purpose. Earlier, Wang et al [4]has used CSP spatial filter weights for channel selection. In this work, other than conventional CSP its four variants are being analysed.

Following are the major objectives of this work:

978-1-5090-4448-1/17/$31.00 ©2017 IEEE

i. To apply CSP methodology both for feature extraction and feature reduction purpose.
ii. The second objective of this work is to find out which variant of CSP achieves maximum classification accuracy with least number of electrodes.
iii. If there are any common channels selected for different subjects by using above-mentioned approach

The next section II will brief the proposed methodology and Section III reports the results. Section IV discusses the results and in the last section V conclusion and future work is presented.

II. METHODOLOGY

CSP algorithm has been widely used for feature extraction in EEG based BCI systems for motor imagery [6, 7]. As the EEG signals have noise and over fitting issue, various regularized CSP algorithms are introduced to cater these issues [8, 9]. In this work, a framework is introduced to find the variant of CSP that achieves maximum classification accuracy with least number of EEG channels. The approach is based on firstly identify the filter weights using complete set of channels. In the next step, the channels are selected based on the highest value of CSP coefficient or vector weight. Channels with largest values are included to find the accuracy in next run and ultimately the optimum combination of number of channels, CSP variant and classification accuracy is figured out.

The structure of our proposed method is elaborated in Fig 1. It comprised of five main phases, signal acquisition, filtering, feature extraction using conventional CSP and its variants, classification and selection of six EEG channels based on highest coefficient or weights. For feature extraction, following CSP algorithms are considered:

i. Conventional CSP
ii. Composite CSP with weighted sum of covariance matrices of other subjects (CSP-Comp)
iii. CSP with weighted sum of covariance matrices using Kullback – Leibler divergence (CSP-KLD)
iv. CSP with weighted Tikhonov Regularization (CSP-WTR)
v. CSP with diagonal loading using cross validation (CSP-DL)

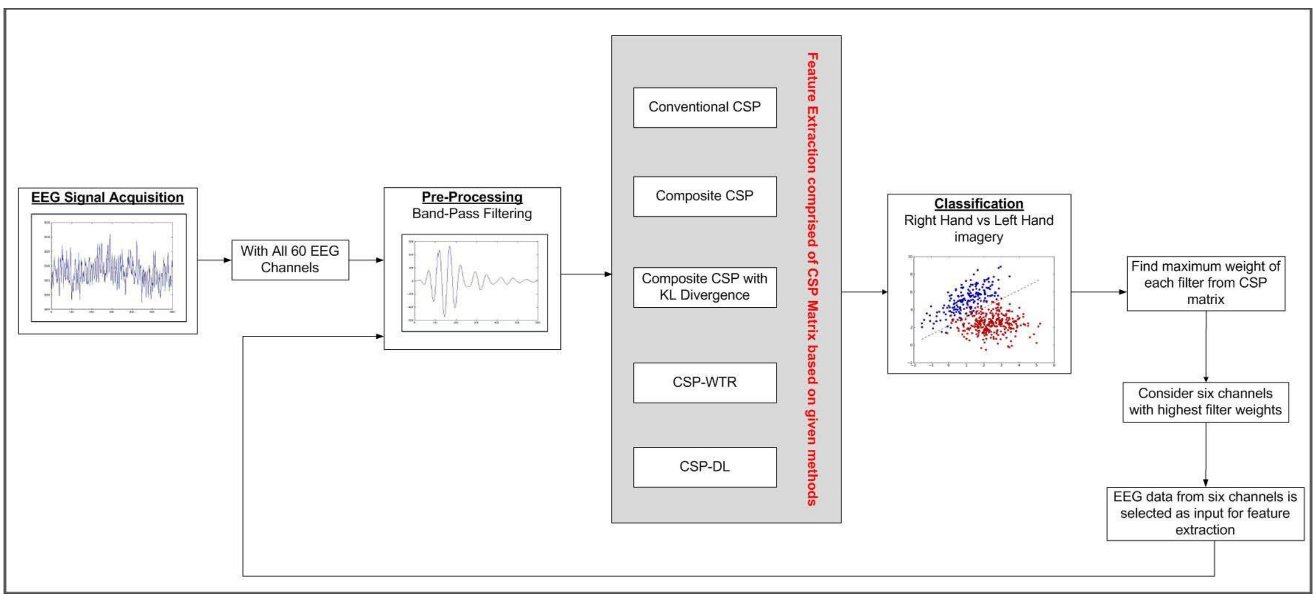

Fig. 1 Framework to reduce EEG channels based on CSP Filter weights

Multichannel data comprised of 60 electrodes from BCI Competition III dataset IIIa from three subjects who performed left hand, right hand, foot and tongue MI is considered for this purpose. Each trial was band-pass filtered in 8-30 Hz, as in [3], using a 5th order Butterworth filter. With each (R) CSP algorithm, we used 3 pairs of filters for feature extraction ($N_f = 3$), as recommended in [4]. Once the spatial filter matrix is obtained from above-mentioned algorithms, the six channels that have highest filter weight are selected. Here the RCSP Toolbox [10] is used that provides the open source code for execution of these algorithms.

In the following section, the algorithms for feature extraction considered in this paper are discussed:

A. Conventional CSP

The Common Spatial Patterns (CSP) algorithm is a feature extraction method which can learn spatial filters maximizing the discriminability of two classes [6, 7]. CSP has been proven to be one of the most popular and efficient algorithms for BCI design, notably during BCI competitions. CSP aims at learning spatial filters which maximize the variance of band-pass filtered EEG signals from one class while minimizing their variance from the other class. As the variance of EEG signals filtered in a given frequency band corresponds to the signal power in this band, CSP aims at achieving optimal discrimination for BCI based on band power features. Formally, CSP uses the spatial filters w which extremize the following function:

$$J(W) = \frac{\omega^T X_1^T X_1 \omega}{\omega^T X_2^T X_2 \omega} = \frac{\omega^T C_1 \omega}{\omega^T C_2 \omega}$$

where T denotes transpose, Xi is the data matrix for class i (with the training samples as rows and the channels as columns) and Ci is the spatial covariance matrix from class i, assuming a zero mean for EEG signals.

B. Composite CSP

Kang et al. [11] proposed Composite CSP algorithm based on regularization of the covariance matrices from other subjects' data. It considers a generic matrix that represents a specific prior on how the covariance matrix for given mental activity considered should be. In composite CSP, this generic matrix is constructed as the weighted sum of the covariance matrices of other subjects corresponding to the same brain activity. Complete details of this technique are available in [11].

C. Composite CSP with Kullback-Leibler divergence

This method is also proposed by Kang et al [11]. Same criterion is used for construction of generic matrix mentioned above but here the weights are defined as per Kullback-Leibler (KL) divergence amongst the subjects.

D. CSP with Weighted Tikhonov Regularization

Fabien et al [10] proposed CSP with weighted Tikhonov Regularization based on the CSP objective function regularization using quadratic penalties. Using this approach, for each electrode high weights are penalized equally. Due to different contribution of channels in spatial filters, penalty levels are defined based on information available in literature for which region of brain and ultimately sensors placement is more relevant to the specific mental activity.

E. CSP with Diagonal loading using cross validation

One of the techniques for covariance matrix regularization is Diagonal Loading that shrinks the matrix towards identity matrix. Fabien et al. [10] have used cross validation of regularization parameter to check the efficiency of this approach.

F. Channel Selection

As a result of above-mentioned CSP based algorithms, spatial filters are obtained in the form of square matrix of n x n where n is the number of channels called spatial patterns. These patterns can be used as good indicators for EEG source distribution as reported by Wang et al. [4]. Channels having maximal values of spatial pattern coefficients are assumed to be the channels that are most correlated with the mental task in consideration. As mentioned earlier in this work three filter pairs are used so minimum six channels are to be considered for optimal selection based on maximal coefficient values.

For 60 channels, we find firstly the maximum value of each channel and named as

$$C_1, C_2, C_3 \ldots C_{60}.$$

Six channels with highest values are considered for feature extraction in the next run as mentioned in Fig. 1. The features extracted from CSP and its variants are given as input to LDA classifier to determine classification accuracies firstly using all sixty channels and in the next run with six selected channels.

TABLE I. CLASSIFICATION ACCURACIES ACHIEVED USING ALL 60 CHANNELS AND 6 CHANNELS FROM CSP, CSP-COMP & CSP-KLD FOR SUBJECT-1

Algorithms	All 60 Channels	6 Channels
CSP	95.56	94.44
CSP-COMP	98.89	90.00
CSP-KLD	95.56	94.44
CSP-WTR	98.89	91.11
CSP-DL	95.56	94.44

TABLE II. CLASSIFICATION ACCURACIES ACHIEVED USING ALL 60 CHANNELS AND 6 CHANNELS FROM CSP, CSP-COMP & CSP-KLD FOR SUBJECT-2

Algorithms	All 60 Channels	6 Channels
CSP	61.67	63.33
CSP-COMP	45.00	50.00
CSP-KLD	61.67	51.67
CSP-WTR	71.67	55.00
CSP-DL	78.33	63.33

TABLE III. CLASSIFICATION ACCURACIES ACHIEVED USING ALL 60 CHANNELS AND 6 CHANNELS FROM CSP, CSP-COMP & CSP-KLD FOR SUBJECT-3

Algorithms	All 60 Channels	% (Increase / Decrease)
CSP	93.33	0
CSP-COMP	93.33	11
CSP-KLD	93.33	11
CSP-WTR	93.33	11
CSP-DL	93.33	5

III. RESULTS

In this paper, we have compared conventional CSP with its regularized techniques, to determine if they could be used for EEG sensor selection. Based on the filter weights, electrodes are selected for feature extraction and classified through LDA. BCI Competition III Dataset is used in this work. Table I-III reports classification accuracies of Subjects 1, 2 and 3 respectively using two combinations of channels. First one using all sensors while second one with reduced number of six channels based on filter weights criterion. In the last column, the increase / decrease of accuracy between the two scenarios is presented.

TABLE IV. CLASSIFICATION ACCURACIES USING 6 CHANNELS FOR EACH SUBJECT

Algorithms	**Sub 1**	**Sub 2**	**Sub 3**
CSP	**94.44**	**63.33**	**93.33**
CSP-COMP	90.00	50.00	83.33
CSP-KLD	**94.44**	51.67	83.33
CSP-WTR	91.11	55.00	83.33
CSP-DL	**94.44**	**63.33**	88.33

IV. DISCUSSION

Various CSP regularized algorithms have been proposed for feature extraction. To the best of our knowledge, no study is observed where they are compared or evaluated for EEG sensor selection. In this work, five regularized CSP methods have been compared based on spatial filter weights. As mentioned in Table I, for subject 1 the classification accuracy has decreased when six electrodes are considered. But with CSP, CSP-KLD and CSP-DL only 1% decrease is reported that answer our question that using filter weights we could reduce EEG sensors without significant degradation in accuracy. Results for subject 2 are quite different from that of Subject 1. Instead of decrease, the accuracy has improved using six electrodes for CSP and CSP-COMP algorithm. While with other algorithms, the accuracy is decreased significantly. The results for subject 3 are also different from other two. For CSP, no change in accuracy is observed while decreased significantly using other algorithms.

In table IV, results from tables I-III are summarized for six channels. Maximum classification accuracies are mentioned in bold for each subject. It is quite interesting that maximum accuracies are achieved from multiple algorithms but for each subject CSP produces highest accuracy. From this, it can be concluded that to cater the issue of overfitting and noise, instead of using regularized CSP algorithms, the conventional one can be used with less number of EEG sensors based on largest filter weights.

In table V, the channels extracted for each subjects are mentioned to find if common channels are selected from CSP algorithms. For subject 1 and 2 channels are selected quite randomly but for subject 3, the results are quite interesting as each algorithm except CSP has produced same set of six electrodes. From the results it is concluded that the true optimal EEG channel constellation varies between the subjects for same mental activity. It is also probable that it differs between sessions of a same subject [12]. In order to find a standard configuration comprised of less number of channels sessions and trials could be increased substantially. Moreover, the number of channels may be increased from six without significant degradation of classification accuracy. Based on selected channels, experiments could be repeated so that the results based on accuracies could be reproduced and standardized.

TABLE V. SIX CHANNELS SELECTED FROM ALL ALGORITHMS

Algorithms	**Sub I**	**Sub II**	**Sub III**
CSP	23,29,30,34,56,57	[23,27,29,43,44,52]	[28,33,34,47,51,52]
CSP-COMP	[19,23,28,29,30,37]	[22,23,26,28,35,44]	[28,33,34,37,47,52]
CSP-KLD	[23,29,30,34,56,57]	[2,15,27,32,44,47]	[28,33,34,37,47,52]
CSP-WTR	[2,7,23,27,29,56]	[8,20,23,39,44,45]	[28,33,34,37,47,52]
CSP-DL	[23,29,30,34,56,57]	[5,6,14,26,45,58]	[28,33,34,37,47,52]

V. CONCLUSION

Using spatial filter weights the number of electrodes required for an EEG-based BCI system can be reduced without significantly compromising classification accuracy. The application of the proposed method is to identify and validate a reduced electrode set using conventional and four regularized algorithms of CSP. The approach is applied on BCI competition dataset. In future, the approach could be applied to our work mentioned in Secton I for assessment of neural correlates of neuroception.

ACKNOWLEDGMENT

This work is funded by Higher Education Commission (HEC), Pakistan and is being conducted and supervised under the ‘Intelligent Systems and Robotics’ research group at Computer Science (CS) Department, Bahria University, Karachi, Pakistan.

REFERENCES

[1] F. Lotte, L. Bougrain, and M. Clerc, "Electroencephalography (EEG)‐Based Brain–Computer Interfaces," *Wiley Encyclopedia of Electrical and Electronics Engineering,* 2015.

[2] H. Lin, Y. Zhuliang, G. Zhenghui, and L. Yuanqing, "Bhattacharyya bound based channel selection for classification of motor imageries in EEG signals," in *2009 Chinese Control and Decision Conference*, 2009, pp. 2353-2356.

[3] L. He, Y. Hu, Y. Li, and D. Li, "Channel selection by Rayleigh coefficient maximization based genetic algorithm for classifying single-trial motor imagery EEG," *Neurocomputing,* vol. 121, pp. 423-433, 2013.

[4] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selelction in motor imagery based brain-computer interface," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2006, pp. 5392-5395.

[5] N. Irtiza and H. Farooq, "The study of fear-induced power modulations for Cognitive Man-Machine Communication," in *Cognitive Informatics & Cognitive Computing (ICCI* CC), 2015 IEEE 14th International Conference on*, 2015, pp. 346-351.

[6] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface," *Journal of neuroscience methods,* vol. 255, pp. 85-91, 2015.

[7] J. S. Kirar and R. Agrawal, "Optimal Spatio-spectral Variable Size Subbands Filter for Motor Imagery Brain Computer Interface," *Procedia Computer Science,* vol. 84, pp. 14-21, 2016.

[8] W. Samek, C. Vidaurre, K.-R. Müller, and M. Kawanabe, "Stationary common spatial patterns for brain–computer interfacing," *Journal of neural engineering,* vol. 9, p. 026013, 2012.

[9] B. Blankertz, M. Kawanabe, R. Tomioka, F. Hohlefeld, K.-r. Müller, and V. V. Nikulin, "Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing," in *Advances in neural information processing systems*, 2007, pp. 113-120.

[10] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *Biomedical Engineering, IEEE Transactions on,* vol. 58, pp. 355-362, 2011.

[11] H. Kang, Y. Nam, and S. Choi, "Composite common spatial pattern for subject-to-subject transfer," *Signal Processing Letters, IEEE,* vol. 16, pp. 683-686, 2009.

[12] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components—a tutorial," *NeuroImage,* vol. 56, pp. 814-825, 2011.