Grenze International Journal of Engineering and Technology, July Issue

A Novel Channel Selection Method for Motor Imagery BCI System

Sruthy Parvathy K[1] and Anish Babu K K[2]
[1-2]Dept. of ECE, Government Engineering College, Thrissur, Kerala, India,
Email: sruthyparvathy395@gmail.com, anish.babu@gectcr.ac.in

***Abstract*—Brain Computer Interfacing (BCI) is a powerful communication tool that exists between a man and a machine. Electroencephalography (EEG) based BCI is popular in the area of rehabilitation of motor disabled people. Motor Imagery (MI) EEG is the recording of brain potentials that occurs while imagining a body movement. Identification of relevant channels in a multi-channel MI EEG is necessary for a reliable MI BCI system. Reducing the large number of channels will greatly aid in reducing system complexity and improving performance. A channel selection method based on Inverse Coefficient of Variation is proposed in this work. Common Spatial Pattern is used for feature extraction utilizing the principle of sub-band CSP. Various classifier models are used for classification to identify the best performing classifier for the particular MI task. The reduced 10 channel subset yields better performance than the original 64 channel set. Enhancement in performance accuracy when the number of channels gets reduced is evident for all classifiers. The accuracy is 52.38% when the whole set of electrodes is employed, whereas accuracy increases to a maximum of 92.85% when only chosen 10 channels are used.**

***Index Terms*— Brain Computer Interfacing, Motor Imagery, Electroencephalography, Inverse Coefficient of Variation and Sub-band CSP.**

I. INTRODUCTION

Brain computer interface technology is a growing field of research which is intended in improving the quality of life of differently-abled persons. EEG is the most popular input tool in BCI because of its non-invasive nature. Advantages of the EEG based systems also include its low cost and high temporal resolution features. EEG signals measure the brain potential acquired through the electrodes placed in the scalp. Electrodes record the brain activity and fed the input electrical signal to the system. EEG is the popular method in BCI systems because of its non-invasiveness, low cost and high temporal resolution characteristics. Non-invasive EEG technique utilizes both dry and wet electrodes for signal acquisition based on the application. Basically, the International 10-20 system is a standard electrode positioning scheme for EEG acquisition. EEG is dominant in five major frequency bands, namely, the delta band (0.5-4Hz), theta band (4-8Hz), alpha band (8-13Hz), beta band (13-30Hz) [1].
MI EEG occurs when the person thinks of doing an action without actually performing the action. This avoids the unnecessary muscle movements involved while doing a particular action. MI EEG finds immense application in the neuro-prosthetics such as active rehabilitation of motor disabled persons. When a person thinks of doing an action, several regions in the brain get stimulated and generated potentials can be

Grenze ID: 01.GIJET.6.2.10_1

recorded. The motor conditions are well obtained from the sensorimotor cortex region of the brain. Sensorimotor rhythms based MI BCI are called as mu (8-12Hz) and beta (12-30Hz) rhythms [2]. There occurs an increase or decrease in the brain potential known as Event Related Desynchronization (ERD) and Event Related Synchronization (ERS). The MI EEG signals are pre-processed using spectral or spatial filters and then features are extracted and represent the signal in a compact form. Then the features can be classified and translated into commands to feed a BCI application. Multichannel EEG carries a large amount of redundant information which leads to degradation in the performance accuracy when compared to an optimal set of channels. Channel selection can identify the dominant brain areas for a particular application in an effective manner. A number of research works are being done in this field to develop sophisticated algorithms for channel selection. Basic channel selection process involves generating channel subsets from the original set of channels and evaluating the sets using proper algorithms. Channel selection techniques in MI EEG are widely classified into five categories [3]; filtering, wrapper, hybrid, embedded and human-based techniques. The channel subsets are chosen on the basis of certain specified criteria, taking into account the channels' position and redundancy factors. The channel selection approaches utilize several search algorithms [4], like the complete search, sequential search, and heuristic/random search. The complete search methods ensure in finding the best channel subset in accordance with the evaluation measure. In sequential search method, the channels can be added and removed in a sequential manner, and these methods are usually simple to implement and fast. Sequential Floating Forward Selection (SFFS) method in [5] is a commonly used sequential method. The random search method brings in randomness into the former two search algorithms and creates the next subset in a random fashion. Genetic Algorithm is an example for random search. These methods turn to be useful for larger data and also when there is a need for low computation time. Selection of relevant minimum channels can reduce the computational complexity, cost and minimize the hardware design. The main objective of this work is to identify the relevant channels and scalp locations for a two class MI application and analyze the classifier performance.

The remaining portion of the paper is organized as follows. Section II describes the methodology used in the project in detail. Section III discusses the results obtained. Section IV concludes the paper, followed by references.

II. METHODOLOGY

A. Data Description

The MI EEG data used in this study is obtained from the physionet databank [6,7]. The EEG data were acquired using 64-channel EEG using BCI2000 system, according to the International 10-20 system. The subjects were asked to perform the imagination of closing and opening of either right or left fist as per the target cue. The cue appeared on the left or the right side of the screen and the subject imagines till the target disappears and then relaxes. Each experiment run consists of 15 trials. Each subject performed 3 such runs, so in total 45 trials per subject. A uniform time window of 4.1 seconds was taken for each MI trials. The data were provided in the EDF format with 64 EEG channels, each sampled at a frequency rate of 160 Hz. Annotation details is specified by the additional 65th channel.

B. Preprocessing

The original data is to be pre-processed before applying channel selection algorithm, to remove the artifacts present in the signal. The EEG data were pre-processed using a fifth order Butterworth IIR filter. The choice of the filter and the filter order is taken based on the available literature [5,8,9], which shows considerably good results for the fifth order filter. The left and right trials were extracted and then passed through a filter bank. The filter bank was constructed by decomposing the 8-32Hz frequency band into six overlapping bands each of bandwidth 4Hz. The frequency filters were 8-12Hz, 12-16Hz, 16-20Hz, 20-24Hz, 24-28Hz and 28-32Hz.The filtered signal is then passed onto the channel selection stage. The work flow of the process is shown in Figure 1.

C. Channel Selection based on Inv CV

Channel selection algorithm based on the EEG signal statistical measure Inverse CV can be categorized as a filtering model for channel selection. To begin with, initialize the number of channels, the number of samples and the number of trials. The energy of the signal over N trials is computed. Then mean and variance of the obtained energy is calculated. Standard deviation is obtained from the square root of variance. Inv CV metric is computed as the ratio of mean to standard deviation (1). The higher the inverse CV ratio, lower is the

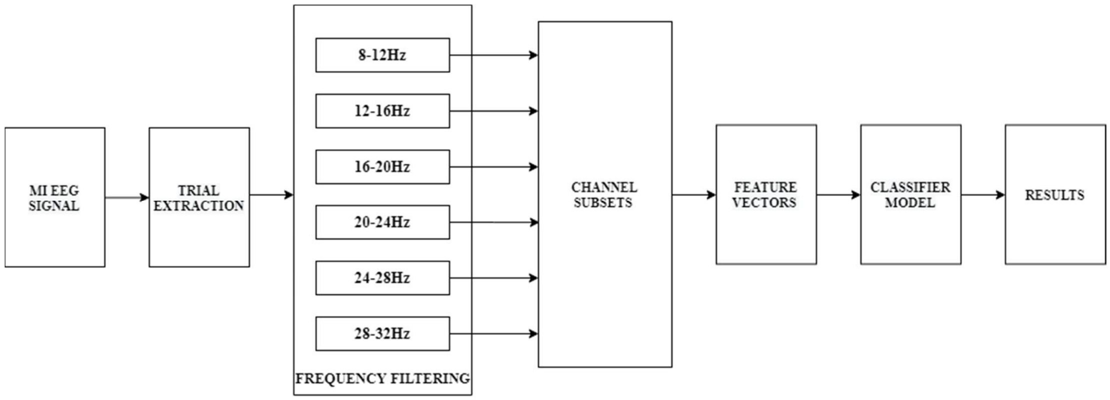

Figure 1. Work flow of MI EEG Channel Selection

deviation of the sample from the mean. Inv CV obtained for each channel is averaged over 31 subjects' data. Channels are sorted in descending order of Inv CV, so that channels with lower variation can be identified. Repeat the procedure for data in all the six frequency bands. The channels with higher Inv CV for both left and right trials are only selected. Those channels are ranked in the order of their frequency of occurrences over all the bands. Using these channels 8 different subsets are formed and tested for accuracy.

$$InvCV = \mu / \sigma \qquad (1)$$

D. Feature Extraction and Classification

CSP is a spatial filtering method in which the filtered signal is passed through a number of spatial filters that computes the features whose variances are optimal to discriminate two MI classes. This method is based on diagonalization of two co-variance matrices simultaneously [10]. The spatially filtered signal Z of single trial EEG E is given as (2):

$$Z = WE\,. \qquad (2)$$

E is an NxS matrix EEG data of a single trial, with N number of channels and S number of measurement samples per channel. Each row of the CSP projection matrix W represents stationary spatial filters. Z maximizes the differences in the variance of the two classes. The m first and last rows of Z, i.e. Zp, $p \in \{1,2,...2m\}$ form the feature vector Xp. The feature vector is computed as in (3):

$$\mathrm{Xp} = \log(\mathrm{Var(Zp)} / \sum_{i=1}^{2m} \mathrm{Var(Zp)}) \qquad \mathrm{p} \in \{1,2,...2\mathrm{m}\}\,. \qquad (3)$$

Classification is the last stage in EEG processing and the accuracy depends on the classifier and also on the data. The various machine learning algorithms used for classification are Support Vector Machine (SVM), Linear
Discriminant Analyzer (LDA), Quadratic Discriminant Analyzer (QDA) and k-Nearest Neighbor (KNN). The classification was done for five subjects each having 21 trials in both classes. Four subjects' data i.e. 84 trials were used to train the model and then 1 subject's data i.e. 21 trials were used for testing. Five different classifier models are tested in this work. Classifier rate or Accuracy (4) defined as the ratio of the total number of correct predictions obtained by a model to the total predictions.

$$\mathrm{Accuracy} = (\mathrm{TP} + \mathrm{TN}) / (\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})\,. \qquad (4)$$

F1 Score (5) combines precision and recall, where precision is obtained from the ratio of total correct predictions to the sum of true predictions and the false predictions of that class and recall is the ratio of correct predictions to the total number of positive observations. A good F1 score suggests a low false positives and low false negatives. A perfect F1 score will be 1. F1 Score metric can be computed as follows:

$$F1Score = 2 * (precision * recall) / (precision + recall) \qquad (5)$$

III. RESULTS AND DISCUSSIONS

The EEG data was obtained in the European Data Format (EDF) format, so the signals were converted into MATLAB readable format. The imagery motion of left and right fists was extracted from the signals. After the trial extraction, each subjects' signal was filtered out to remove the artifacts. Channel selection algorithm is applied on the filtered signals obtained over six frequency bands. The number of channels obtained after applying the channel selection algorithm reduced to nearly half of the total number of channels. The Table I shows the 34 selected channels which occurred maximum number of times while evaluating the signal, with higher Inv CV ratios.

TABLE I. REDUCED SET OF CHANNELS

4 : FCz	3 : FC1	9 : C3	5 : FC2	7 : FC6	31:F5	32:F3
12 : C2	17 : CP1	16 : CP3	22 : FP1	26 : AF3	49:P3	56:P8
33 : F1	34 : FZ	36 : F4	41 : T7	42 : T8	2:FC3	61:O1
50 : P1	52 : P2	57 : PO3	58 : PO7	62 : Oz	59:P04	
21 : CP6	29 : AF8	37 : F6	47 : P7	54 : P6	1:FC5	

When the subset with 10 channels is passed through the classifiers, the accuracy increases tremendously compared to that with 64 channels. To obtain the prediction accuracy, the trained classifier models were tested on unseen data, where training and test data were taken in the ratio of 80:20. The results showed a high increase in the accuracy showing almost 80% increase from 64 channel set to 10 channel subset for Discriminant analyzers. SVM with Gaussian kernel and KNN models also showed a similar trend in the performance. Those best performing channels included in the 10 channel subset is given in Table II. While linear SVM showed a higher accuracy for 15 channel subset, with an accuracy of 85.71%. Table III shows the accuracy rates of all the classifiers used in this work. Classifier models were trained using default MATLAB functions for 7-fold cross validation.

TABLE II. BEST PERFORMING CHANNEL SUBSET

Sl. No.	Channel Label	Channel Number
1	FCz	4
2	FC1	3
3	C3	9
4	FC2	5
5	FC6	7
6	C5	8
7	C2	12
8	CP1	17
9	CP3	16
10	FP1	22

The confusion matrix in Figure 2 shows the total number of actual trials and predicted trials for both 64 channel set and 10 channel set for QDA classifier evaluating data of 5 subjects. Out of 21 test trials in each class, 21 right trials and 18 left trials are correctly predicted for 10 channel subset yielding an accuracy of 92.85%. In case of 64 channel set, only right trials are correctly predicted and all the left trials are misclassified as the right class giving an accuracy of 50%. The right MI class is denoted as class 0 and class 1

TABLE III. PERFORMANCE ACCURACY OF CLASSIFIERS

Number of channels	LDA	QDA	Linear SVM	Gaussian SVM	KNN
64	50	50	50	50	52.38
34	50	50	50	50	50
30	50	50	50	50	50
25	50	50	52.38	50	50
20	50	66.66	50	50	50
15	50	50	85.71	50	50
10	90.47	92.85	69.04	80.95	73.80
5	54.76	76.19	64.28	61.9	54.76
2	50	38.09	38.09	50	54.76

corresponds to left MI. Thus there is an increase of about 42% by reducing the channels For the best performing classifier QDA, the F1-score for 64 channel set is obtained to be 0.67 whereas it is 0.93 for 10 channel subset. It is observed that after a particular number of channels, there is no improvement in the performance accuracy, which indicate that the new channels added does not give any additional signal information. For every classifier except for linear SVM, there is a tremendous increase of accuracy for 10 channel subset. The accuracy rate decreases thereafter, to a minimum value for 2 channel subset. All the channels in the 10 channel subset are found to fall in the sensorimotor area of the brain where the MI EEG is dominant, validating the significance of channels selected. Each of the classifier models displays different trends in the accuracy as the number of channels is varied.

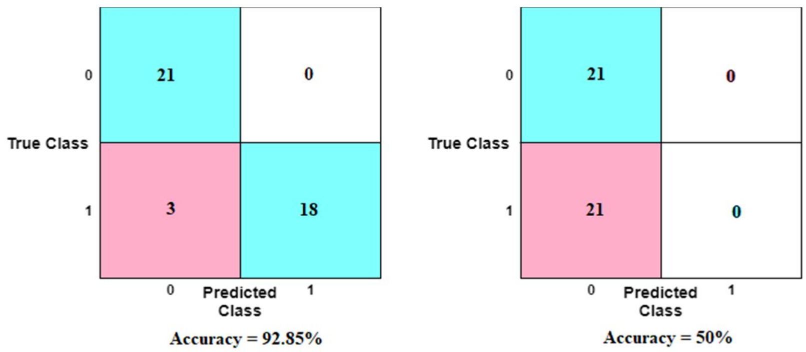

Figure2. Confusion matrix for 10 channel subset (left) and 64 channel set (right)

IV. CONCLUSION

This paper proposes an effective channel selection based on EEG statistics. Finding channels relevant to a particular motor imagery task reduces the computational complexity involved in a BCI system originally comprising a large number of electrodes. The channel subsets are selected based on the Inv CV metric. The working frequency is divided into 6 sub bands and the channels are selected from all the bands with the largest metric. These reduced channel sets proved to have better efficiency as compared to the original 64 channels. A subset of 10 channels gave maximum accuracy for every classifier model used. The accuracy obtained for 64 channels is nearly in the range of 50-52.5%, while the reduced set with 10 channels give accuracy in the range of 69-92.85% based on the classifier model. The QDA classifier yielded high performance accuracy of 92.85%. The reason for difference in classification accuracy for different classifiers can be studied as a future work.

REFERENCES

[1] B. J. Baars, N. M. Gage, “Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience”, 2nd edn. Elsevier, 2010.

[2] Thiyam, D.B.; Cruces, S.; Olias, J.; Cichocki, A. Optimization of Alpha-Beta Log-Det Divergences and their Application in the Spatial Filtering of Two Class Motor Imagery Movements. Entropy 2017, 19, 89.

[3] Turky Alotaiby, Fathi E Abd El-Samie, Saleh A Alshebeili and Ishtiaq Ahmad, "A review of Channel Selection Algorithms for EEG signal processing", EURASIP Journal on Advances in Signal Processing, 2015, pp.2-21.

[4] Yang H, Guan C, Ang KK, Phua KS, Wang C, “Selection of effective EEG channels in brain computer interfaces based on inconsistencies of classifiers”, 36th Annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2014, pp. 672–675.

[5] Z. Qiu, J. Jin, H.K. Lam, Y. Zhang, X. Wang, and A. Cichocki, “Improved SFFS method for channel selection in motor imagery based BCI”, Neurocomputing, vol. 207,2016, pp. 519–527.

[6] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals, 2003, pp.215-220.

[7] Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R. BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. IEEE Transactions on Biomedical Engineering, 2004, pp.1034-1043.

[8] Jian Kui Feng, Jing Jin, Ian Daly, Jiale Zhou, YugangNiu, Xingyu Wang, and Andrzej Cichocki, “An optimized channel selection method based on multifrequency CSP-rank for motor imagery-based BCI system”, Hindawi Computational Intelligence and Neuroscience, 2019, pp. 1-10.

[9] Vikram Shenoy Handiru and Vinod A. Prasad, "Optimized bi-objective EEG channel selection and cross-subject generalization with brain computer interfaces", IEEE Transactions on Human-Machine Systems, 2016, pp. 1-10.

[10] Kai Keng Ang, Zheng Yang Chin, Haihong Zhang and Cuntai Guan," Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface," 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, 2008, pp. 2390-2397.