ISSN (e): 2250 – 3005 || Volume, 08 || Issue,10|| October – 2018 ||
International Journal of Computational Engineering Research (IJCER)

EEG Channels Selection Based On Higher Order Statistics For BCI Application

Shah Chintan Pankajbhai[1], Dr. V.A.Shah[2]
[1]Assistant Professor, Biomedical Engg. Dept., Govt. Engg. College, Gandhinagar –382028,Gujarat,India.
[2]Professor and Head,Dharamshi Desai University, Nadiad –387001,Gujarat,India.
Corresponding Author: Shah Chintan Pankajbhai

ABSTRACT:*EEG signals are generally measured on multiple locations for Brain Computer Interface (BCI) application as well as other modalities. This high number of channel count increases complexity of algorithm additionally gives discomfort to patient. In order to resolve both problems we have proposed an approach based on higher order statistics like kurtosis, Skewness and moment. Based on them we find channels of highest deviation during onset of activity. The selected channels classification accuracy compared with Linear Discriminant Analysis, Artificial Neural Network and Support Vector Machine on publically available database on physionet website.*

KEYWORDS:*EEG, Brain Computer Interface, Channel Selection, kurtosis, Skewness, Moment, Linear Discriminant Analysis, Support Vector Machine, Artificial Neural Network, Statistical Features*

Date of Submission: 15-11-2018 Date of acceptance: 29-11-2018

I. INTRODUCTION

Brain Computer Interface provides means to communicate and control devices without use of muscular activity. It is used for amputee patients who have lost their hand or legs. Thus it is very important to provide noninvasive media for them since incretion of electrodes requires costly surgery and can lead to infections. In noninvasive Brain Computer Interface application we require to determine movement from scalp Surface EEG measurements. The EEG electrodes for measurement placed across several location on scalp. The location determined by montage system used. This montage system determines number of electrodes placed in particular region of scalp. We cannot use lower number of electrodes because it provides insufficient information. If we use large number of channel than it will increase redundant channel and noise in signals leads to reduction in performance of BCI plus increases BCI cost. It also increases setting time of the BCI and such large number or electrode creates discomfort to patient. To improvise performance under such circumstances we require to choose effective channel from total number of channels. Various channel selection algorithm are available from which several are discussed here.

TianLan et. al. have proposed channels selection algorithm based on mutual information maximization [1]. They have used linear Independent Component Analysis (ICA) with one dimensional entropy to estimate mutually independent features. They provide ranking to individual channel by selecting one channel at a time. During raking phase they estimate mutual information between features in selected EEG channel and its assigned class. This way they rank channel and finally based on ranking 10 channels were selected. And classification was done by KNN method and KDE method. MahnazArvanehet. al. have used sparse common spatialpattern (SCSP) algorithm for channel selection[2].They find subject dependent channels that provides highest accuracy by sparsifying common spatial filters within a constraint of classification accuracy. They have compared their results with commonly used channels Cz, C3 and C4 with channel selected by their approach. For classification they have used SVM and the hyper-parameters were obtained by a grid searchusing cross validation on the training data.

MahnazArvanehet. al. in another paper used Decision tree to select best channel set which produces highest detection accuracy[3]. In the paper author have reviewed different channel selection methods like Fischer Criteria, Mutual Information, Support Vector Machine and Common Spatial Pattern. In their proposed approach they have filtered EEG data using different filter bank. Followed by Feature selection based on Mutual Information. In their approach they have used Classification and Regression Tree (CART) as decision tree. And form ranking purpose they have used apruning process based on the overall accuracy of the tree is applied to rank the features.Yuan Yanget. al. have implemented subject specific channels selection by using Fisher Discriminant Analysis [4]. They used time segment optimization with channel location for data measurement.It projects high-dimensional data onto a directionand performs a linear classification in this one-dimensional

space.The optimal projection is found by maximizing the separation between two classes. They have used Time domain parameters (TDPs) are a set of broadband (i.e. 8-30Hz) EEG features defined in the time domain.
Alejandro Gonzalezet. al. have used combination of multi-objective hybrid real-binary Particle Swarm Optimization (MHPSO) with Fisher Discriminant Analysis[5]. The objective function of MHPSO is to find EEG channels subset and classifying parameters which maximizes classification accuracy. Here algorithm fine tunes channels selection using Particle Swarm Optimization method. The most common channels were selected from all the subjects and classification accuracy tested on those selected channel for all those subjects. HaijunShanet. al. have utilized Multiple frequency-spatial synthesized features from each channels and using IterRelCen algorithm they find effective channel[6]. IterRelCen is based on relief algorithm and modified for matching two aspects change in target sample selection process and iterative computation in order to achieve robust feature selection procedure. The proposed algorithm tested on three different Motor Imagery datasets and used Multiclass Support Vector Machine. TurkyAlotaibyet. al. have provided review of different channels selection algorithms [7]. They performed channels selection for different applications like Seizure detection/prediction, motor imagery classification, mental task classification, emotionclassification, sleep state classification, and drug effects diagnosis. They suggested that different applications have different requirements of channel selection they have used Signal processing tools such as time-domain analysis, power spectral estimation, andwavelet transform have been used for feature extraction. They have also evaluated filtering, wrapper, embedded, hybrid, andhuman-based techniques for channels selection.
From above literature we can conclude that channel selection is very important task for any application. For BCI application reduction in channel not only improves efficiency of algorithm but also reduces discomfort to subject. For channels selection statistical method can be a better tool compared to frequency based tools. Because frequency component will provide movement present in data but will not provide which particular limb movement is currently happening. So through statistical method we can detect difference in limb movement related data for EEG. In our work we have proposed a higher order statistical method to detect movement related differences and select channels to maximize classification accuracy. In proposed work we have used Kurtosis, Skewness and Moment for channel selection. Following section discusses database used for work, proposed methodology and results with discussion.

II. DATABASE USED

The database used here is publically available on physionet website [8]. We have used EEG Motor Movement/ Imaginary Database (EEGMMIDB). This database is developed and contributed by Gerwin Schalk and his team. The database contains recording of total 109 normal subjects with actual and imagined movement. The signals were recorded by 10/10 system with 64 total EEG channels. All the subjects have performed total six movements as follows:

- Eyes open
- Eyes close
- Left hand/ right hand actual movement
- Left hand/ right hand imagined movement
- Both fists/ both feet actual movement
- Both fists/both feet imagined movement

Above all movements are recorded in two sets with numerical encoding in dataset. The signals are recorded in accordance to queuing process where subject alternatively asked to rest and perform any above activity. The queuing process displayed on screen with instructions and subject performs task until instruction disappears on screen. The database is encoded by event marked with T0, T1 and T2. T0 means resting state with no movement, T1 means either left hand movement actual/ imagined or both fists movement actual/ imagined and T2 means either right hand movement actual/ imagined of both feet movement actual/ imagined. For our analysis we have considered all the 64 channels available in dataset.

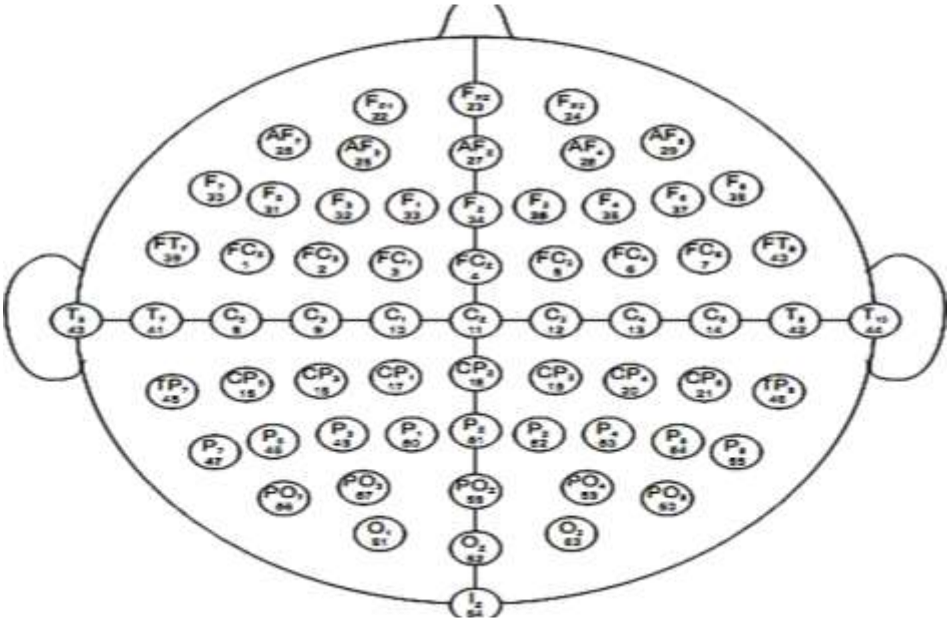


Fig. 1 Electrode Locations

In our analysis the data is pre-filtered with 12 - 30 Hz band in order to locate beta activity of brain parts. The sampling rate is 160Hz and activity performed for 3 seconds brief time window. This data are extracted based on event marker and epochs are generated based on activity performed. We have focused on left or right hand actual movement only.

III. PROPOSED METHODOLOGY

In proposed approach we have used filtered signals as suggested in previous section. Here we measure higher order statistics of each channel for all the subjects. We have measured Kurtosis, Skewness and 5thorder Moment for each channel. Since the channels are already filtered we get Event Related Synchronization (ERS). Following Figure shows Proposed Methodology.

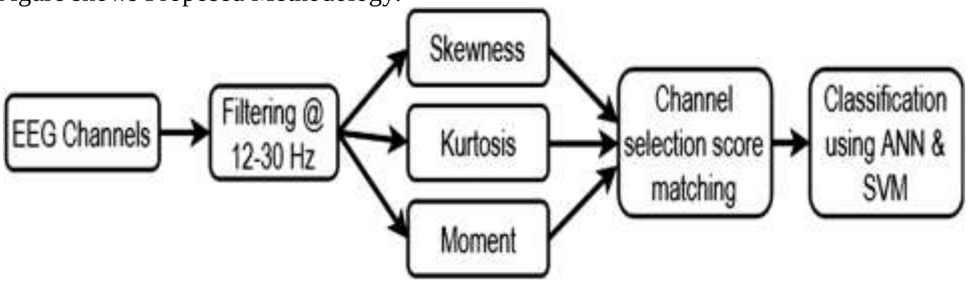


Fig. 2 Proposed Methodology

After generating feature vector of all the channel we perform score matching of each subject followed by classification by ANN and SVM.

Kurtosis:

Kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. That is, data sets with high kurtosis tend to have heavy tails, or outliers. Data sets with low kurtosis tend to have light tails, or lack of outliers. The kurtosis can be calculated as give in following equation.

$$k = \frac{\sum_{i=1}^{N}(\mathrm{d}_i - \overline{D})^4 / \mathrm{N}}{S^4} \quad (1)$$

Where d_1, d_2… $d_n$is data point, D is mean of sample distribution, N is total data points and S is standard deviation of channel data.

Skewness:

Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point.The skewness for a normal distribution is zero, and any symmetric data should have a skewness near zero. Negative values for the skewness indicate data that are skewed left and positive values for the skewness indicate data that are skewed right. By skewed left, we mean that the left tail is long relative to the right tail. Similarly, skewed right means that the right tail is long relative to the left tail. If the data are multi-modal, then this may affect the sign of the skewness. It can be calculated by following equation.

$$sk = \frac{\sum_{i=1}^{N}(d_i - \overline{D})^3 / N}{S^3} \quad (2)$$

Where d_1, d_2… d_n is data point, D is mean of sample distribution, N is total data points and S is standard deviation of channel data.

5th order Moment:

As with variance, skewness, and kurtosis, these are higher-order statistics, involving non-linear combinations of the data, and can be used for description or estimation of further shape parameters. The higher the moment, the harder it is to estimate, in the sense that larger samples are required in order to obtain estimates of similar quality. The 5th-order moment can be interpreted as measuring "relative importance of tails versus center (mode, shoulders) in causing skew" (for a given skew, high 5th moment corresponds to heavy tail and little movement of mode, while low 5th moment corresponds to more change in shoulders).

$$m_k = E(\mathrm{d} - \mu)^{\mathrm{k}} \quad (3)$$

Where E is expected value, μ is mean of sample and k is order of the system

Normalization of data:

As we measure higher order statistics the score or value we get have different scale so in order to provide common score selection criteria to all subjects. We normalize data across all subjects according to following equation.

$$norm = \frac{(d_i - \min(\mathrm{d}))*100}{\max(\mathrm{d}) - \min(\mathrm{d})} \quad (4)$$

Using all above equations we have generated following results. We have separated data epochs according to left side movement and right side movement. For individual higher order statistics are calculated for both sides' epochs. But both side of movement we got same score for same individual. But in case of different individuals the score is different due to non-linearity present in anatomical structure of individual.

IV. RESULTS AND DISCUSSION

Fig. 3 Score of all 64 channels

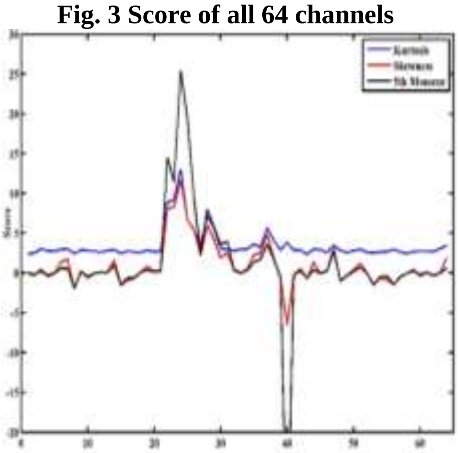


Fig. 4Kurtosis based Channel Selection

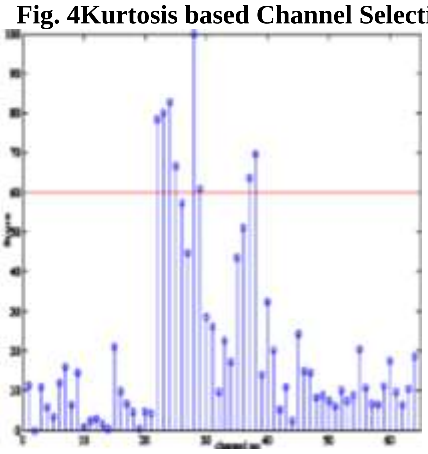


Fig. 5 Skewness based Channel Selection

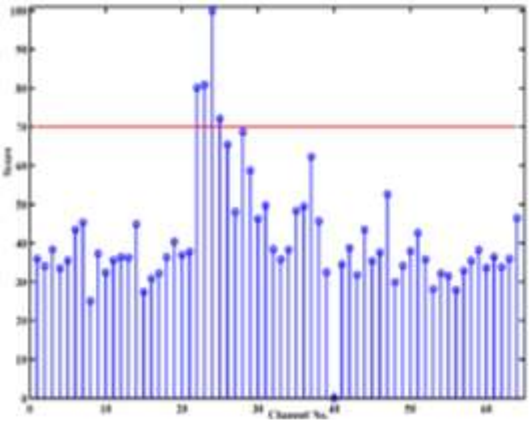


Fig. 6Moment based Channel Selection

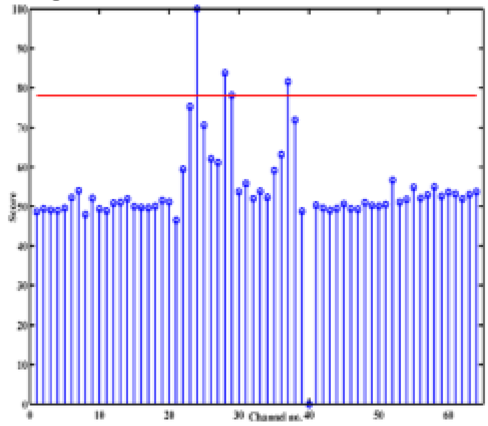


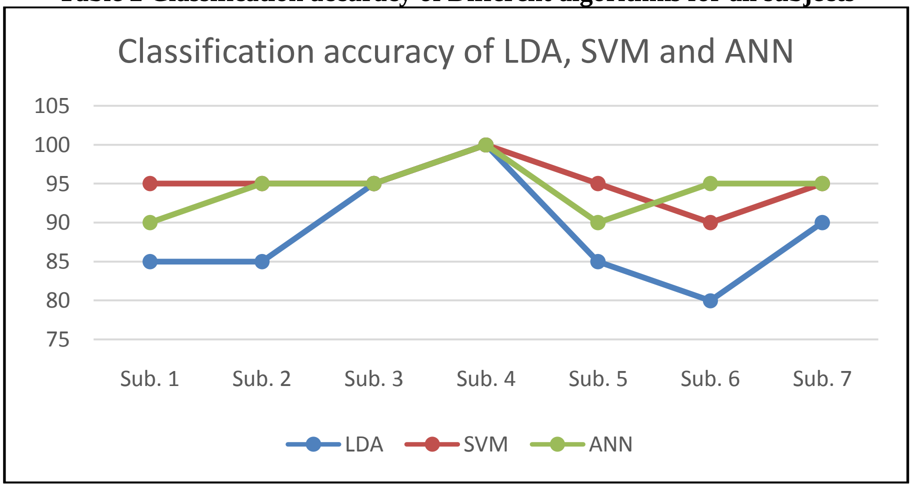
Fig. 3 is plotting of higher order statistics features versus channel number. Here 5th moment feature having highest values compared to Skewness and Kurtosis but due to normalization it is not visible. The higher order statistic almost shows same trend except for 5th moment of 40th channel. In same manner we gathered data for all subjects. This data is average of five trials of each subject so average EEG signals of a person is selected. Fig.4 is plotting of normalized kurtosis value to channel number. This is channels selection for one individual where threshold for channel selection is shown by red line. 60% threshold is take for channel selection of kurtosis feature for whole database. In same manner we have determined channel for Skewness based and Moment based features. Here 70% threshold is set for Skewness for whole database as shown in fig. 5. 78% threshold is set for Moment for whole database as shown in fig. 6. This way following channels selected based on average score of all features (Skewness Kurtosis and Moment).

Table 1 Selected Channel for all subjects

Subject	Channels Selection		
	Skewness	Kurtosis	Moment
subject-1	FC4, T7, T10	FCz, FC2, FC4, C5, FC6, T7, TP7	FC4, T7, T5, T10, TP7
subject-2	FCz, FC4, FC6, Cz, C2	FCz, FC2, FC4	FCz, FC2, FC5
subject-3	FC4	FCz, FC2, FC4, T7, T10, TP7	FCz, FC2, FC4, T7, T10, TP7
subject-4	FCz, FC2, FC4, FC6, C5	FCz, FC2, FC4, FC6, FT8, T7, C5,	FCz, FC2, FC4, FC6, C5
subject-5	FCz, FC2, FC4, FC6, T7, C5	FCz, FC2, FC4, T7, C5, C6	FCz, FC2, FC4, T7, C5, C6
subject-6	FCz, FC2, FC4, FC6, C5	FCz, FC2, FC4, FT8, FC6, T7,C5	FCz, FC2, FC4, FT8, FC6, T7,C5
subject-7	FCz, FC4, FC6, C5	FCz, FC2, FC4, FC6, T7, C5, T10	FCz, FC2, FC4, FC6, T7, C5, T10

In this table all red marked channels are selected based on common channels in each higher order statics.
Based on above table all the channels are selected from database according to individual selection. We have implemented LDA, SVM and ANN (pattern net). This algorithm is implemented on MATLAB. The results are shown in below table.

Table 2 Classification accuracy of Different algorithms for all subjects



From above chart it is clear that SVM provides highest accuracy compared to other two method. The reason for highest accuracy of SVM is due to small amount of data set used for training of algorithm. So for limited training sample SVM will always retain highest accuracy. This accuracy can be improved for other methods by increasing size of training data.

V. CONCLUSION

The proposed method reduces channel number so as to increase accuracy, reduce setup time and less discomfort to patient. Higher order statistical features forms the basis for such channel reduction. Further improvisation of algorithm could be addition of statistical analysis like PCA to further reduce dimensions of selected channel. The channels yielded by this method have some repetition in dataset so it is possible to populate more number of electrode in that region.

REFERENCES

[1]. TianLan, D. Erdogmus, A. Adami, M. Pavel and S. Mathan, "Salient EEG Channel Selection in Brain Computer Interfaces by Mutual Information Maximization," 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, 2005, pp. 7064-7067.

[2]. M. Arvaneh, C. Guan, K. K. Ang and C. Quek, "Optimizing the Channel Selection and Classification Accuracy in EEG-Based BCI," in IEEE Transactions on Biomedical Engineering, vol. 58, no. 6, pp. 1865-1873, June 2011.

[3]. MahnazArvaneh, Cuntai Guan, Kai KengAngand Hiok Chai Quek, “EEG Channel Selection UsingDecision Treein BrainComputer Interface,” Proceedings of the Second APSIPA Annual Summit and Conference, pages 225–230,Biopolis, Singapore, 14-17 December 2010

[4]. Yuan Yang, OlexiyKyrgyzov, Joe Wiart, Isabelle Bloch, “Subject-Specific Channel Selection For Classification of Motor Imagery Electroencephalographic Data,”38th International Conference on Acoustics,Speech, and Signal Processing, May 2013, Vancouver, Canada. Pp.1277-1280, 2013.

[5]. Alejandro Gonzalez, Isao Nambu, HaruhideHokari, and YasuhiroWada, “EEG Channel Selection Using Particle Swarm Optimization forthe Classification of Auditory Event-Related Potentials,” The Scientific World Journal, vol. 2014, Article ID 350270, 11 pages, 2014.

[6]. Haijun Shan,HaojieXu, Shanan Zhu and Bin He, “A novel channel selection methodfor optimal classiication in diferent motor imagery BCI paradigms,” BioMedical Engineering on Line, 14, Pp. 93-111, 2015.

[7]. TurkyAlotaiby, Fathi E Abd El-Samie, Saleh A Alshebeiliand Ishtiaq Ahmad, “A review of channel selection algorithms for EEG signal processing,” EURASIP Journal on Advances in Signal Processing, 66, 2015.

[8]. 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 51(6):1034-1043, 2004.

Shah Chintan Pankajbhai "EEG Channels Selection Based On Higher Order Statistics For BCI Application " International Journal of Computational Engineering Research (IJCER), vol. 08, no. 10, 2018, pp 13-18