RESEARCHERID
THOMSON REUTERS

[Shah* *et al.*, 6(11): November, 2017]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

BRAIN COMPUTER INTERFACE IMPLEMENTATION USING CLUSTERING TECHNIQUES

Shah Chintan [*1], Dr. V.A.Shah [2], P.M.Pithadiya[3]

[*] Biomedical Engineering Dept., GEC-Gandhinagar, India

DOI: 10.5281/zenodo.1049502

ABSTRACT

Brain Computer Interface (BCI) generally utilizes non-invasive EEG signals in order to detect intended movement. Normally all EEG electrode receive brain activity on scalp surface which is superimposition of different brain activity. As the number of channel is high we need to reduce it in order to properly detect movement related activity. In this work we have utilized different clustering algorithms to reduce insignificant channels. We have extracted features using statistical methods Principal Component Analysis (PCA). And the classification is done using unsupervised learning algorithm in order to improve generalization capability. The proposed approach reduces false positive rate and it also shows different person have different activity region.

KEYWORDS: Brain Computer Interface, k-means clustering, hierarchical clustering, Principal Component Analysis.

I. INTRODUCTION

Brain Computer Interface (BCI) is generally provides communication of human subject to control things using computer. The development of such system is for use of severally disable or amputee people with any limb. This kind of persons generally lacks normal functionality and in order to restore movement we require to understand intentions by using noninvasive means. EEG can be best suited for such purpose because it provides high temporal resolution, maximum sensitivity toward brain activity and noninvasive nature [1-3]. The classification of EEG channel with particular movement is challenging task since signals we get on scalp surface are superimposition of all brain activity [4-5]. Since the Brain dynamics are subject specific so generalized and robust algorithm is required for properly classification. That means in order to produce subject specific algorithm, we require channel location on scalp surface for each subject [6, 11]. In literature we found following major approaches applied. For feature extraction we have mainly two major approaches: 1) Time- Frequency domain feature [4], 2) statistical features [14]. Whereas for classification we have 1) Linear discriminants, 2) non-linear classifiers, 3) supervised and unsupervised neural network and 4) statistical classifiers.

Recent studies related to our work have utilized newer approaches which are worth mentioning here. Spyrou et.al. Have used alpha and beta frequency band to determine sensory-motor rhymes. Five frequency band's power averaged over time period to be used as features followed by regularized linear logistic regression (rLLR) classifier [6]. Whereas M Blokland et.al have used same method of Spyrou et.al. With different frequency parameter setting [7]. Myrden et.al. Have used spectral power of each frequency with 15 electrode 10/20 system and SVM and LDA as a classifier to determine mental state of subject rather than their movement [8]. Lu et.al. Have utilize wavelet package decomposition method as feature extractor and for classification Deep Belief Network based on restricted Boltzmann machine [9]. Uehara et.al. Have uses common spatial pattern as well as tangent space mapping to generate feature set and compared classifier performances of LDA and SVM [10]. Lin Wu et al. used sub-band CSP to generate feature for different frequency band and uses fuzzy integral with particle swarm optimization (PSO), which can regulate subject-specific parameters [11]. They have utilized the technique for Motor Imaginary (MI) data. Even Sang-Hoon et al. have developed algorithm for MI using regularized CSP for 4-40Hz band. They have used Fisher's Linear Discriminant (FLD) to extract feature from different sub-band and these all features than used to generate PCA features [12]. Blankertz et al. have utilized spatial pattern for features and used shrinkage linear discriminant analysis for better performance on subject specific data in order to reduce high dimentionality [13].

For this paper we have used EEG Motor Movement/Imager dataset from Physionet [14-15]. The data contains 64 channel sampled at 160Hz. The recording have event marker and consists of activity like raising right hand

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

or left hand and eye close or open. From above discussion it is apparent that we first require to choose channels significant for individual. We also neglected channels falls onto occipital lobe and frontal lobe since these parts are far away from sensory motor region. We have used k-means and hierarchical clustering algorithm to identify mostly effective channel for individual. After words features needs to selected from effective channels. For our purpose we have chosen statistical feature extraction method Principal Component Analysis. Then comes choice of algorithm, so in next section we discuss proposed algorithm in detail.

II. MATERIALS AND METHODS

Database

Data base is taken from Physionet is contributed by Gerwin Schalk using BCI2000 instrumentation system [14]. Subjects performed different motor/imagery tasks while 64-channel EEG were recorded using the BCI2000 system. Each subject performed 14 experimental runs: two one-minute baseline runs (one with eyes open, one with eyes closed), and three two-minute runs of each of the four following tasks:

1. Subject opens or closes his right or left fist corresponds to target appears on screen.
2. Subject imagines opening or closing of his right or left fist corresponds to target appears on screen.
3. Subject opens or closes both fist or feet corresponds to target appears on screen.
4. Subject imagines opening or closing of both fist and feet corresponds to target appears on screen.

From above we have used data from first task which is of opening and closing of fist either on right side or on left side depends on input. Using this configuration we get 64 channel data but all the channel will not be utilized since different brain part performs different activity. That means we can remove non-significant channels which do not falls on motor sensory part of brain. As shown in fig.1 we will utilize highlighted 15 channels.

FC1	FCZ	FC2	C3	C1	C_Z	C2	C4	C6	CP3	CP1	CP_Z	CP2	CP4	CP6
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Table 9.Channel list with number (for further reference)

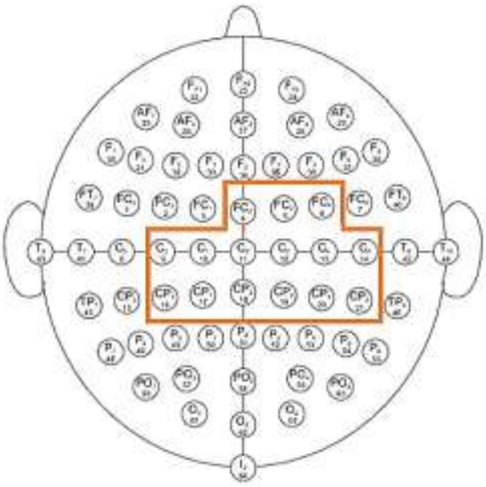


Fig.1 10/10 electrode system with motor sensory region selected

Clustering Algorithms

Clustering analysis is done to find out similarity or distinction between an object from set with other objects in set to identify groups present in the set. Greater homogeneity with in group and greater variance between groups creates better clustering. As we know that person to person spreading of volume conduction is different effective electrode will also differ. That's why clustering is better suited for identification of mostly distinct channel. So subject specific location can be found out from available 15 channels. Other advantage of clustering is that it is unsupervised in nature so we don't have to provide class label. It develops model from object to provide labels to unlabeled objects and create classes. We have utilized hierarchical clustering and k-means clustering to produce distinct classes which gives channel location with distinct activity. From available all clustering method we have used K-means and hierarchical. Both method have advantages and disadvantages which is: 1) K-means clustering requires prior knowledge of K (number of clusters), whereas in hierarchical clustering you can stop at whatever number of clusters we wish. 2) K-means starts with a random choice of cluster centers,

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

therefore it may yield different clustering results on different runs of the algorithm. 3) K-means algorithms complexity is linear whereas hierarchical clustering it is quadratic function. So, for better performance we decided to use hierarchical clustering to define cluster 1st. afterwards we will use this no of cluster (we get using Hierarchical cluster) and perform K-means clustering. At the end we will get one or two channel than we extract feature using PCA in order to get different feature for left side or right side movement.

Hierarchical Clustering

There are two basic approaches for generating a hierarchical clustering:

1. Agglomerative
2. Divisive

Agglomerative will start with the points as individual clusters and, at each step, merge the closest pair of clusters. Whereas in Divisive Start with one, all-inclusive cluster and, at each step, split a cluster until only singleton clusters of individual points remain. Human brain has ability to partition data using agglomerative or inclusive hierarchies. That's why we have used agglomerative clustering method for our algorithm and used Euclidian distance for linking next point to nearby cluster.

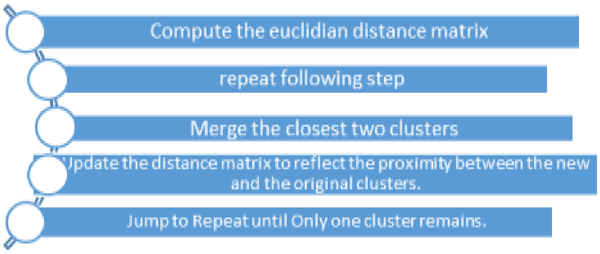


Fig.2 algorithm for agglomerative hierarchical clustering

Major shortcoming of this algorithm is locally optimizes an objective function of distance and computation load high. As a remedy to situation we can use K-means clustering as a next step to attain global objective function.

K-means Clustering

K-means method is basic method where initially user specifies number of classes and same number of centroid randomly initialized. Each point in cluster than assigned to nearest centroid and collection of such points are termed as cluster. Then update centroid based on each point in cluster. Repeat the process until we get stable centroid value. The algorithm is shown in figure below.

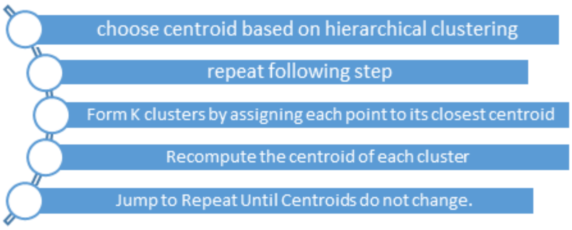


Fig.3 algorithm for K-means clustering

In fourth step the recomputed centroid can vary depends on data and goal of clustering. And here we can assign an objective function that can determine closeness of point to the centroid. Objective function is taken as sum squared error (SSE), where Euclidian distance between centroid and each point is calculated in Euclidian space. Quality of clustering assured when Sum square of all the distance for each class is minimized.

$$SSE = \sum_{i=1}^{k} \sum_{X \in C_i} dist(C_i, X)^2 \qquad (1)$$

where dist is standard Euclidian distance between two points. And center that minimizes SSE is nothing but the mean value which is given below,

$$C_i = \frac{1}{m_i} \sum_{X \in C_i} X \qquad (2)$$

Step 3 forms clusters by assigning points to their nearest centroid, which minimizes the SSE for the given set of centroids. Step 4 recomputes the centroids so as to further minimize the SSE. There are several proximity function available which we have used to compute mostly affected channels.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Principal Component Analysis

Principal Component Analysis (PCA) is statistical process which can give weightage to different features from subject specific EEG. This weightage are in accordance to activity performed by user. The channels we find for both activity are same in almost all the cases so we use only affected channel to determine feature related to activity in that channels only. Since PCA gives Eigen value in return we can put a threshold on Eigen value so we can limit the number of features for classification. That means irrelevant features have less Eigen value so we can remove irrelevant features.

(i) Mean of the database image is calculated in order to remove common features from database.

(ii)

$$m = \frac{1}{M}\sum_{i=1}^{M} x_i \qquad (3)$$

Where x_i is pixels of images

(iii) Remove common features from database by subtracting each image from mean. Images are called as mean centered.

$$w_i = x_i - m \qquad (4)$$

(iv) To form covariance matrix data matrix must be multiplied by data matrix dash (transpose of data matrix).

$$C = WW^T \qquad (5)$$

(v) Calculate Eigen value and Eigen vector from covariance matrix.

$$WW^T(W d_i) = \mu_i(W d_i) \qquad (6)$$

Where wd_i is Eigen vector and μ_i is Eigen value

(vi) Sort the Eigen vector from high to low according to their corresponding Eigen values.

(vii) Multiply data matrix with sorted Eigen Vector to get Eigen space.

(viii) Project each data onto the feature space to get feature vector of each EEG data.

These projected data will yield Eigen values which we use as feature for classification purpose. Classification we have done with Linear Discriminant Analysis (LDA).

Linear Discriminant Analysis

LDA is supervised method which projects higher dimensional space data into lower dimensional space using a transformation function. The transformation function is found out by LDA preserves class structure of high dimensional space in low dimensional space. In discriminant analysis two scatter matrices are used as objective function, within-class (S_w) and between-class (S_b).

$$S_w = \sum_{i=1}^{k}\sum_{x\in\pi_i}(x - mi)(x - mi)^T \qquad (7)$$

$$S_b = \sum_{i=1}^{k} n_i(mi - m)(mi - m)^T \qquad (8)$$

Where $m_i = \frac{1}{n_i}\sum_{x\in\pi_i} x \quad mean\ of\ ith\ class$

And $m_i = \frac{1}{n}\sum_{i=1}^{k}\sum_{x\in\pi_i} x \quad global\ mean$

In above equation π_i is classes and n_i is data point of from ith class. An optimal transformation would maximize trace (S_b) and minimize trace (S_w). Eigen decomposition method will applied to solve following equation which gives us transformation function G. results will be discussed in next section.

$$\max\{trace((S_w)^{-1}S_b)\}\ and\ \min\{trace((S_b)^{-1}S_w)\} \qquad (9)$$

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

III. RESULTS AND DISCUSSION:

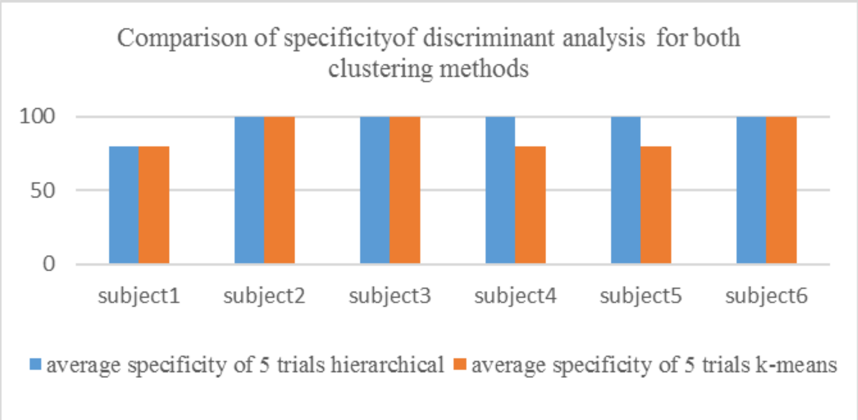
Subject	Left side movement channel	Right side movement channel
1	11,4	11,3
2	2,3	2,3
3	10,4	10,3
4	15,5	15,4
5	2,3	2,3
6	2,3	2,3

Table 2. Chart of effective channels found using Hierarchical method

Subject	Left side movement channel	Right side movement channel
1	2,3,4,5	2,3,4,5
2	2,3	2,3
3	10	10
4	15	15
5	2,3	2,3
6	2,3	2,3

Table 3. Chart of effective channels found using K-means method

The above tables' shows result of hierarchical and K-means clustering, where we get channels maximally affected for individual. These results are for average of five trials of each individual. As it is apparent that K-means is more stable as compared to hierarchical clustering. Now using this channels (for hierarchical 2-channels and k-means 1-channel is used) as input to PCA and we will get feature vector. From feature vector of each data we used 1st eleven features because their corresponding Eigen-values are maximum. The result comparison for classification accuracy of hierarchical cluster and k-means clustering method using discriminant analysis is in shown below chart.



Specificity means false classified as false of left side movement classified as left side movement. Here as it is apparent from chart that five out of six subject we get full specificity for hierarchical method. On the other side K-means method yields lower specificity rate as compared to hierarchical method.

IV. CONCLUSION

The methodology adopted is subject specific channel selection which solves main key aspects like sensor location variation and volume conduction function variation. The task of algorithm become much easier by taking reduced data channels, increased robustness and reduced training time. So from this study we can conclude that by using statistical method like clustering which is also unsupervised in one sense, can reduce the overall burden of classification method. Generalization achieved by proposed approach which is also subject specific in nature. This method can further extended to imagined movement where subject which amputees can also benefitted.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

V. REFERENCES

[1] P. Aricò, G. Borghini, G. Di Flumeri, N. Sciaraffa, A. Colosimo and F. Babiloni, "Passive BCI in Operational Environments: Insights, Recent Advances, and Future Trends," in IEEE Transactions on Biomedical Engineering, vol. 64, no. 7, pp. 1431-1436, July 2017.

[2] J. Wolpaw and E.W. Wolpaw, Brain-Computer Interfaces Principles and Practice, Oxford University Press, 2012.

[3] B. Blankertz, S. Lemm, M. Sebastian Treder, S. Haufe, and K-R. Mller, "Single-trial analysis and classification of erp components - a tutorial,"NeuroImage, pp. 814–825, 2011.

[4] S. Saha, K. I. Ahmed, R. Mostafa, A. H. Khandoker and L. Hadjileontiadis, "Enhanced inter-subject brain computer interface with associative sensorimotor oscillations," in Healthcare Technology Letters, vol. 4, no. 1, pp. 39-43, 2 2017.

[5] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," Clinical Neurophysiology, vol. 113, pp. 767-791, 2002.

[6] L. Spyrou, Y. Blokland, J. Farquhar and J. Bruhn, "Optimal Multitrial Prediction Combination and Subject-Specific Adaptation for Minimal Training Brain Switch Designs," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 24, no. 6, pp. 700-709, June 2016.

[7] M Blokland, Yvonne & Farquhar, J & Mourisse, Jo & Scheffer, Gert Jan & Lerou, Jos & Bruhn, Jörgen, "Towards a Novel Monitor of Intraoperative Awareness: Selecting Paradigm Settings for a Movement-Based Brain-Computer Interface," PloS one. 7(9), 2012.

[8] A. Myrden and T. Chau, "A Passive EEG-BCI for Single-Trial Detection of Changes in Mental State," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 4, pp. 345-356, April 2017.

[9] N. Lu, T. Li, X. Ren and H. Miao, "A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 6, pp. 566-576, June 2017.

[10] T. Uehara, M. Sartori, T. Tanaka and S. Fiori, "Robust Averaging of Covariances for EEG Recordings Classification in Motor Imagery Brain-Computer Interfaces," in Neural Computation, vol. 29, no. 6, pp. 1631-1666, June 2017.

[11] S. L. Wu et al., "Fuzzy Integral with Particle Swarm Optimization for a Motor-Imagery-Based Brain–Computer Interface," in IEEE Transactions on Fuzzy Systems, vol. 25, no. 1, pp. 21-28, Feb. 2017.

[12] S. H. Park and S. G. Lee, "Small Sample Setting and Frequency Band Selection Problem Solving Using Subband Regularized Common Spatial Pattern," in IEEE Sensors Journal, vol. 17, no. 10, pp. 2977-2983, May15, 15 2017.

[13] S. Ge, R. Wang, Y. Leng, H. Wang, P. Lin and K. Iramina, "A Double-Partial Least-Squares Model for the Detection of Steady-State Visual Evoked Potentials," in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 4, pp. 897-903, July 2017.

[14] 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.

[15] 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. Circulation 101(23):e215-e220

CITE AN ARTICLE

Chintan , S., Shah, V. A., & Pithadiya, P. M. (n.d.). BRAIN COMPUTER INTERFACE IMPLEMENTATION USING CLUSTERING TECHNIQUES. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY*, *6*(11), 245-250.