PRE-PROCESSING, FEATURE EXTRACTION AND CLASSIFICATION OF EEG SIGNALS FOR BRAIN COMPUTER INTERFACE

Shubhangi Gupta[#1], Jaipreet Kaur Bhatti[#2]
[#1]Student, Department of Electronics and Communication, Gyan Ganga College of Technology, Rajeev Gandhi Technical University, Jabalpur, India
[#2]Asst. Professor, Department of Electronics and Communication, Gyan Ganga College of Technology, Rajeev Gandhi Technical University, Jabalpur, India

Abstract: Brain Computer Interface establishes an artificial communication channel between human brain and external world. It enables the patient of complete body paralysis to control assistive applications such as movements of robotic arm or a wheel chair by using the power of human thoughts embedded in Electroencephalogram signals. In present work, a methodology is proposed for pre-processing, feature extraction and classification of Electroencephalogram signals for implementation of Brain Computer Interface. Motor Imagery signals for different mental activities are recorded and filtered out for noise reduction and artifact cancellation in the first step. In the second step, useful features are calculated from processed Electroencephalogram signals to prepare the feature vector. This feature vector is given as an input to the Artificial Neural Network classifier for training purposes in the third step. Appropriately trained classifier algorithm can further be used for translating Electroencephalogram signals to meaningful commands in real time.

Index Terms-- Brain Computer Interface (BCI), Electroencephalogram (EEG), Support Vector Machine (SVM), Classification

I. INTRODUCTION

Human-computer interaction has been a topical research area since the birth of the computer era. Methods of computer interaction have progressed rapidly over the years for various vital applications. Today there exist varieties of innovative technologies that allow humans to interact with computers for the applications such as control or communication. The most of the efforts have been dedicated to the design user friendly and efficient system to produce comfortable means of communication with external world. In order to implement the Interfaces such as voice recognition, gesture recognition and other technologies based on physical movement have received enormous research attention over the years and successful examples of these technologies are being rolled out commercially as a consequence.

The Brain-Computer Interface (BCI) is a communication channel that enables subject to control assistive applications and appliances through thoughts only. The area of cognitive neuroscience has been prompted by technological advances such as Electroencephalography (EEG), Magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI). EEG technique is a noisy technique of brain activity recording. However, it is the most useful technique of brain activity recording as it is cheap, non-invasive and carries direct functional correlations with high temporal resolution. The main goal of any BCI research is to develop an artificial communication channel for severely disabled people of locket-in syndrome. For those patients who have lost all voluntary muscle control, referred to as locked-in syndrome [1]. The Locked-in syndrome can be caused diseases such as amyotrophic lateral sclerosis (ALS) [2], brainstem stroke, mitochondrial disease, spinal-cord injury, traumatic-brain injury [3]. Although, these patients are completely paralyzed and unable to speak, they are cognitively intact and alert and thus require various means to communicate. It is marked in the studies that about one million people are suffering from locked-in syndrome worldwide. It is this motivation that has inspired researchers to explore the possibility of harnessing the intact brain signals of these people as a means of communication.

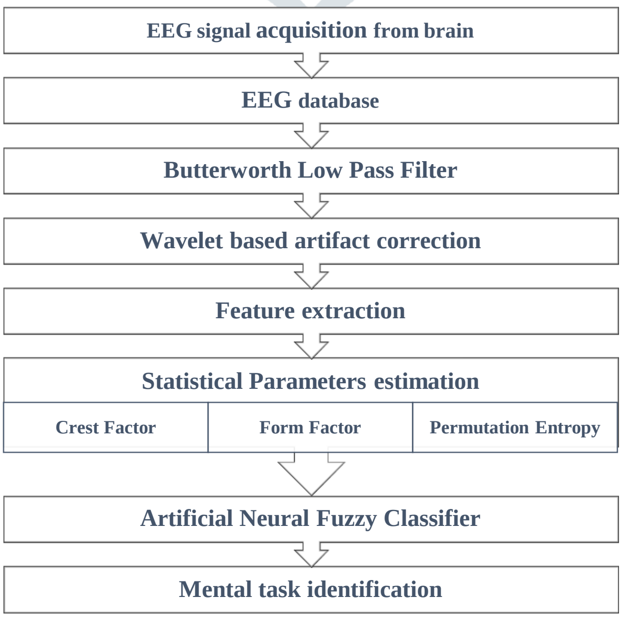

Fig. 1 Methodology proposed for EEG signals pre-processing, feature extraction and classification

The accuracy BCI systems mainly rely on accuracy and efficiency maintained during recording, preprocessing, feature extraction and classification of brain electrical activity. The cortically generated EEG is often contaminated by non-cerebral artifact sources which include ocular movements, eye blinks, Electromyogram (EMG) and Electrocardiogram (ECG) artifacts. Apart from these disturbances, other unknown sources of noise such as instrumentation noise, electrical disturbances and external electromagnetic activity have higher impact on EEG recordings in large extent. When such artifacts are strongly present with desirable EEG activity, it makes difficult the analysis and information extraction from EEG signals. In order attain higher classification efficiency; researches have proposed many novel algorithms for preprocessing, feature extraction and classification. For preprocessing Independent Component Analysis (ICA), Common Spatial Pattern (CSP), and Principle Component Analysis (PCA) are common and mostly used algorithms. For extracting useful features from EEG signals, different researchers applied different approaches and employed different neuromechanism [4], like sensorimotor activity [5-7], Visual Evoked Potential (VEP) [8], Slow Cortical Potential (SCP) [9] P300 and others. SVM and Neural Networks (NN) are the most trusted and efficient classification algorithms and give comparable results in the terms of classification efficiency [10]. In present work, EEG signals for response to a metal tasks imagined motor activity, for left fist blink, right blink and both fist blink are studied and classified.

In present work, EEG signals recorded for imagination of three different fist movements left fist blink, right fist blink and both fists blink are considered. EEG signals are pre-processed using wavelet based artifact correction technique for obtained clean EEG signals and passed through Butterworth low pass filter for obtaining specific frequency band only. Low pass Butterworth filter ensures that noise associated with frequencies other above 30 Hz does not affect the study. Further, feature extraction and classification is carried out on the artifact corrected EEG signals. EEG signals recorded from C3 and C4 EEG channels have been used and classified in present study. For Classification of three classes of EEG data, Artificial Neural Fuzzy Classifier (ANFC) has been used. Fig.1 shows overall methodology presented in this work. The methodology proposed in present work is useful for extracting features, classifying imagined motor movement EEG signals using extracted features, and generating command with higher classification efficiency. EEG signal shows different classification efficiency for different set of mental tasks, this makes feasibility analysis necessary on EEG signals for higher classification efficiency. Proposed methodology is used for analyzing feasibility of EEG signals recorded corresponding to different mental tasks, for higher classification efficiency.

II. EEG DATABASE

EEG database used for this study is taken from PhysioNet ATM [11]. PhysioNet is an open source of biomedical signals. In this database, EEG signal records of 109 volunteers for similar mental activity are available. EEG signals were recorded by means of BCI2000 system [12] for executed and imagined motor movement activity. In this database EEG signals were measurement through 64 EEG channels and sampling was done at the rate of 256 Hz. The position of the electrodes was as per the international 10-20 system of electrodes placement [13]. Each volunteer has undergone for 14 experimental runs, two experimental runs for baseline eyes open and eyes closed and rest 12 experimental runs, each of two minutes for following four tasks:

1. The volunteers perform opening and closing of right or left fist when a target appears on the right or left side of the computer screen.
2. The volunteers imagine opening and closing of right or left fist when a target appears on the right or left side of the computer screen.
3. The volunteers perform opening and closing of both fists or both feet when a target appears either on the top or the bottom of computer screen.
4. The volunteers imagine opening or closing of both fists or both when a target appears either on the top or the bottom of computer screen.

In present work, the EEG data is divided into three categories or classes. In each class, the EEG data recorded for imagination related tasks is only considered. The EEG data used in this study can be categorized as follows:

1. Imagination of opening and closing of Left fist
2. Imagination of opening and closing of right fist
3. Imagination of opening and closing of both fists

The data recorded from C3 and C4 EEG channels is considered in this study. As C3 and C4 channel record the electrical activity takes place in motor imagery part of brain cortex, which is responsible for the planning, preparation and execution of any body movement. In this study, EEG signals recorded from one volunteer are studied. Fig.2 shows the sample of EEG signals recorded for different types of fist movements.

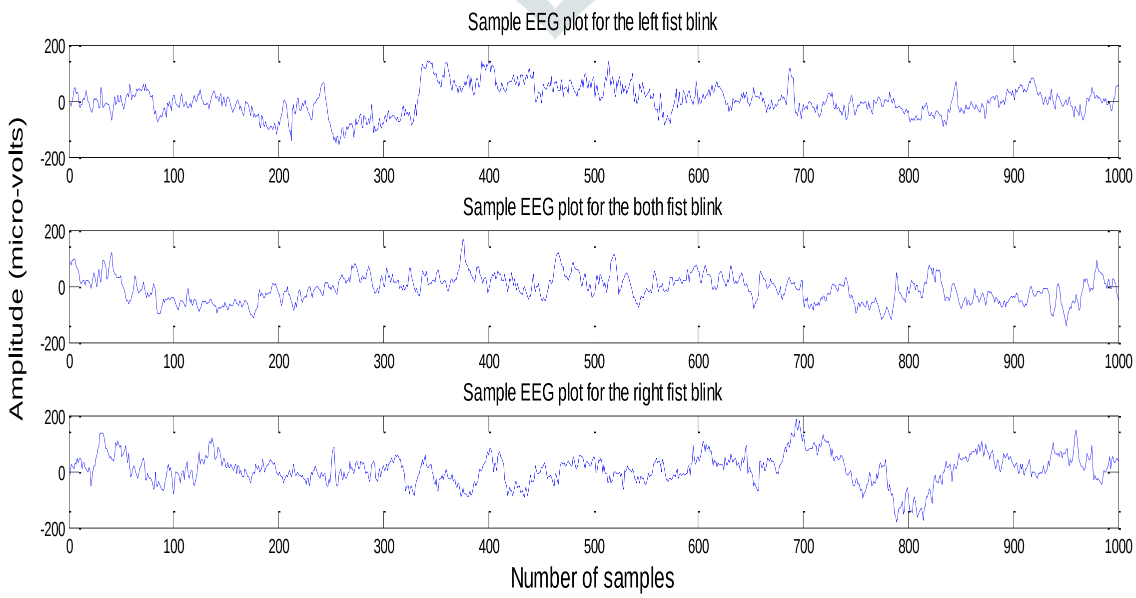

Fig. 2 Sample EEG plot for the different type of fist movements

III. PREPROCESSING OF EEG SIGNALS

3.1 BUTTERWORTH LOW PASS FILTERING

The EEG data recovered from C3 and C4 channel has been filtered out by implementing a low pass Butterworth filter, with cut-off frequency of 30 Hz, to reduce the noise associated with other frequency ranges in EEG signals. EEG signal divides in the following frequency bands 0.1- 3.5Hz (delta), 4-7.5Hz (theta), 8-13Hz (alpha), 13-30Hz (beta) and >30Hz (gamma) [16-17]. Each frequency sub band has its own importance for different mental states and brain functioning parts. As most of the essential information related to brain electrical activities concentrate below 30 Hz of frequency, so it becomes feasible to filter out rest of the frequency components.

Butterworth filter has been used for filtering EEG data, as it has monotonic and maximum possible flat frequency response in the pass band. Also transition bandwidth of the Butterworth filter reduces with the increase in filter order. In this work for sharp cut-off filtering, low pass Butterworth filter having filter order 10 is used.

3.2 EEG ARTIFACT CORRECTION

Electroencephalographic data is usually affected by artifacts produced from non-cerebral origins such as ocular movements, Electromyogram activity and other electrical noises. It is needed to remove artifacts from electroencephalogram activity prior to information extraction and classification of Electroencephalographic activity. In present work, an automated artifact correction technique is used for ocular artifact elimination and noise suppression from EEG signals of different fist movements. The proposed technique relies on the joined application of Independent Component Analysis (ICA) algorithm and Wavelet Transform (WT). This study is based on the fact that the ocular artifacts contribute high amplitude to the WT coefficients; hence suppression of WT coefficients yield ocular artifact corrected signals. This technique is applied on real EEG signals in order to remove any possible contamination due to non-cerebral sources. This WT transform based automatic artifact correction technique helps in error-free classification of different electroencephalogram signals for implementation of a Brain-Computer interface.

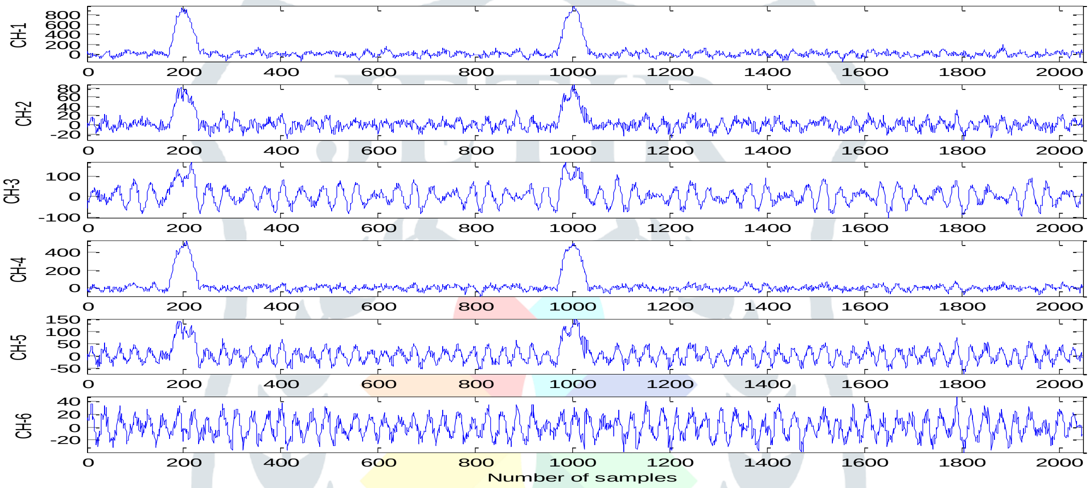

Fig. 3 A set of six channel simulated artifact contaminated EEG activity

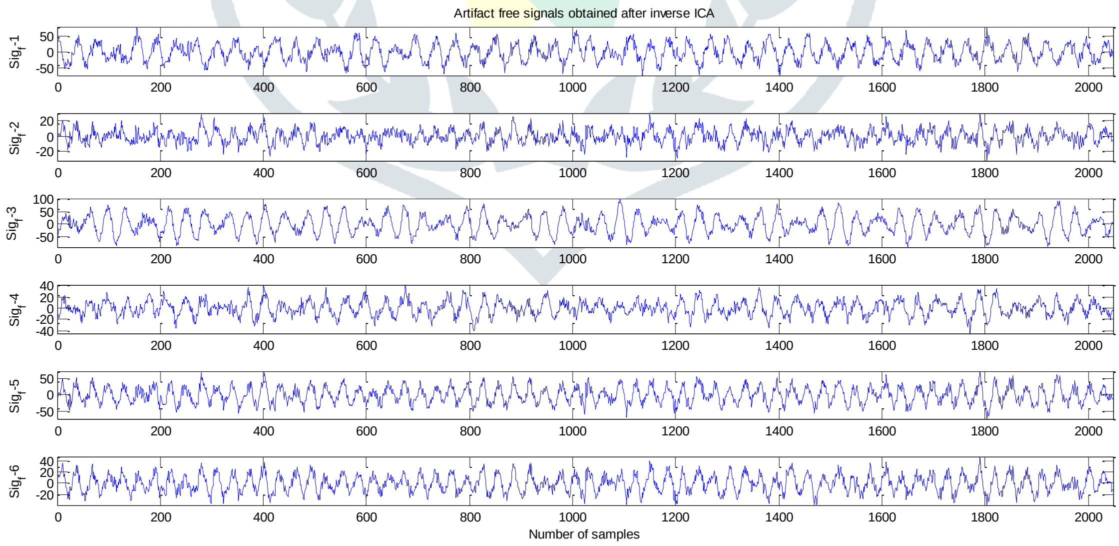

Fig. 4 Artifact corrected EEG activity obtained after using automatic artifact correction technique

In order to remove artifacts from EEG signals, first multichannel EEGs are separated into fundamental Independent Components (ICs) using MULTI-COMBI ICA algorithm. Once ICs of the EEG signals are obtained, the criterion of Fractal Sparsity is applied for automated detection of the artifactual ICs. Further, WT is performed on the identified artifactual ICs. Once, the WT coefficients are obtained for identified artifactual ICs, thresholding of the WT coefficients is performed in the next step. Thresholding of the WT coefficients is performed using soft thresholding approach. Further, the thresholded WT coefficients are processed by inverse WT algorithm to obtain clean ICs in time domain. Inverse ICA is performed on time domain clean ICs to reconstruct ocular artifact free EEG signals in final the step. In order to illustrate the efficacy of artifact correction technique, the application of the methodology is presented on the simulated artifact contaminated EEG signals. Fig.3 shows the six channel artifact contaminated EEG activity. Fig.4 shows the artifact corrected EEG activity obtained after automated artifact correction methodology.

IV. FEATURE EXTRACTION & CLASSIFICATION

Three features are extracted from the processed EEG signals for preparation of the feature vector. Once the feature vector is prepared, it is given to the classifier algorithm Artificial Neural Fuzzy Classifier (ANFC) for training and validation of the feature extraction methodology proposed in the present work. The features used in present work are demonstrated as follows:

1. Permutation entropy: It is defined as

   PE = $-\sum_{i=1}^{m} \pi_i \ln \pi_i$; here π represents permutation pattern

2. Form factor: It is the ratio of RMS value and average value of signal.

   Form factor = RMS value/Average value

3. Crest factor :It is ratio of peak value of signal to its RMS value

   Crest factor = Peak value/RMS value

4.1 Adaptive Neural Fuzzy Classifier (ANFC)

The neuro fuzzy classifier is an adaptive network based system in which the antecedent parameters are adapted with neural networks. This combined system with fuzzy logic qualitative approach and artificial neural network adaptive capabilities named as adaptive neuro fuzzy classifier. ANFC explicates a zero order surgeon fuzzy inference model in to the framework of a multilayer artificial neural (ANN) network with adaptive and non-adaptive nodes. ANFC is based on fuzzy rules and to initialize fuzzy rules k-mean algorithm is used. ANFC regulates the membership function and other antecedent parameter using scaled conjugate gradient (SCG) algorithm. Moller (1993) has described SCG is two time faster than the back propagation algorithm because of its super linear convergence rate [14].

For two inputs{x_1,x_2} and one output y fuzzy classification rule is defined as

If X_1 is A_1 and X_2 is A_2 then y is C_1 class

where A_1 and A_2 are the linguistic terms that are defined on feature space X_1 and X_2 and C_1 represents class label of the output y.

In this architecture each node in the same layer has the same node function. First Layer generates membership grade of each input to specified fuzzy region. In this layer for membership function (MF) bell shape, gaussian, triangular and trapezoidal functions can be used. Gaussian functions have less parameter and because of its parameters smooth partial derivatives it is utilized as MF. Gaussian MF is described as

$$\gamma_{ij}(x_{cj}) = \exp\left(-0.5\frac{(x_{cj}-\alpha_{ij})^2}{\delta_{ij}^2}\right)$$

where x_{cj} is the input variable and α_{ij} and δ_{ij} are the centre and width of the Gaussian function respectively. Next layer is rule layer which uses the membership values of input to calculate firing strength of fuzzy rules. So the θ_{ic} firing strength of the i^{th} rule is

$$\theta_{ic} = \prod_{j=1}^{N} \gamma_{ijc}$$

This layer describes the fuzzy rules for x_c sample. N is number of features. Third layer in classifier calculates the weighted outputs. Maximum firing strength of the rules decides output class. If the rule output weight for a class is biggest among the other class weight it means that the particular class region is controlled by that rule. Weighted output ϕ_{ck} for the c^{th} sample in the k^{th} class can be shown as

$$\phi_{ck} = \sum_{i=1}^{M} \theta_{ic}\omega_{ik}$$

where ω_{ik} denotes degree of association to k^{th} class that is controlled with the i^{th} rule, M represents number of rules. The next layer is known as normalization layer. Its function is to normalized network output because sometimes the summation of the weight can be larger than *1*.

$$\eta_{ck} = \frac{\theta_{ck}}{\sum_{i=1}^{K}\theta_{cl}} = \frac{\theta_{ck}}{\sigma_c}, \sigma_c = \sum_{l=1}^{K}\theta_{cl},$$

where η_{ck} is the degree of cth sample that belongs to k^{th} class. K is the number of classes. Then, λ_c class label can be calculated by the maximum of η_{ck}

$$\lambda_c = \max_{k=1,2....K}\{\eta_{ck}\}$$

V. RESULTS

EEG signals, recorded from motor imagery of a subject for three different fist movements, have been used to train and test the efficiency of proposed methodology. As EEG signals recorded from different subjects are widely different. This makes it necessary to use EEG signals for same subject for training and testing the methodology. However, with the increase in training inputs to the machine learning techniques ANFC, classification efficiency increases. So, EEG data was five folded to increase training inputs corresponding to each class. The model was trained with the 54 set of features (86% of total feature vector) corresponding to each class. For testing the efficiency of complete model, 9 set of features (14% of the total feature vector) corresponding to each class was given to algorithm. Table (1) shows confusion matrix for classification of three classes of EEG data. From Table (1) it is clear that, the proposed model could efficiently classified between mental imagination activity for left fist and right fist using input feature vector obtained from C3 and C4 channels and classification efficiency was 100%. However model was unable to classify both fist blinking very accurately and the classification efficiency was only 66%. The overall classification efficiency of complete model for the imagination of three different fists movement was 88 %.

Table (1) Confusion matrix for Subject-1

Test set	Test set	Test set	Output
LFBI	RFBI	BFBI	Classification
9	0	0	Left Fist Blink Imagined (LFBI)
0	9	0	Right Fist Blink Imagined (RFBI)
1	1	7	Both Fist Blink Imagined (BFBI)

Table (2) Comparison table of methodology

References	Ahmed [15]	Yazdani et al. [16]	Iturrate et al. [17]	Ahirwar et al. [18]	Iturrate et al. [2]	Present Work
No. of channels /electrodes	8 electrodes	16 electrodes	16 electrodes	19 electrodes	16 electrodes	2 electrodes
Neuromechanism (Electrophysio-logical sources employed by a BCI user to generate control signals)	Sensorimotor rhythms (SMR)	Sensorimotor rhythms	P300	Sensorimotor rhythms	P300	Response to the Mental Task
Features Considered (Time or Frequency domain features)	Wavelet Transform (Time-Frequency features)	Discrete Fourier Transform and Common Spatial Patterns	Time domain patterns as features	Frequency Domain Features	Time domain patterns as features	Time-Frequency domain patterns as features
Classifier Used	Neural Network	Artificial Neural Networks and Support Vector Machine	Linear Discrement Analysis	-NA-	Linear Discrement Analysis	ANFIS classifier
Classifier Efficiencies	-NA-	66%	94%	-NA-	94%	Best efficiency obtained for two mental imagination tasks is 100%

Table (2) presents a brief comparison of present work with previous research findings. In present work, higher classification efficiency is achieved for a set of mental activity (Left fist blink and right fist blink) using time domain features of EEG data recorded over C3 and C4 channels only. This reduced the number of channels requirement and processing time as dimensionality of the data is reduced with the reduction in number of channels.

VI. CONCLUSION

The aim of the study was to implement a BCI with three controlling commands, using brain signal of imagination recorded only on C3 and C4 channel. As, less number of channels reduce the complexity associated with multiple channels and also computational time requirement. Input vector to ANFC classifier consist features, extracted from pre-processed EEG signals for imagination of three different fist movements left fist blink, right fist blink and both fist blink. Using 14 % of the total input vector for testing the model, SVM classifier classified three classes of EEG data with 88% of classification efficiency for C3 and C4 channel. SVM classified among two classes (left fist blink and right fist blink movement imagery activity) with 100% of classification efficiency. But classification efficiency for both fist blink movement imagery activity was 66% only. This model shows fairly good results for generation of two controlling commands using EEG signals for left fist blink and right fist blink movement imagery activity, hence can be used for application where two controlling commands are sufficient to regulate complete process.

REFERENCES

[1] Vidal, J. J. "Real-time detection of brain events in EEG" Proceedings of the IEEE, Vol. 65, Iss. 5, pp. 633-641. (1977).

[2] Berger, H. and Gloor, P. "Hans Berger on the electroencephalogram of man : the fourteen original reports on the human encephalogram" (1969). Amsterdam; Barking: Elsevier. 'Electroencephalography and Clinical Neurophysiology' supplement; no.28.

[3] Bronzino, J. D. "The Principles of Electroencephalography" The biomedical engineering handbook. Vol. 2nd ed, pp. 201-212. (2000). Boca Raton, Fla.: CRC Press.

[4] Bashashati,A.,Fatourechi, M., Ward, R. K. and Birch,G. E.(2007)'A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals', J. Neural Eng. , Vol. 4, pp.R32–R57.

[5] Furtscheller, P. and Aranibar, G. (1977)'A Event-related cortical desynchronization detected by power measurements of scalp EEG',Electroencephalogr. Clin.Neurophysiol.,Vol.42, pp.817–26.

[6] Jasper, H. and Penfield,W. (1949) 'Electrocortiograms in man: effect of voluntary movement upon the electrical activity of the precentralgyrus', Arch. Psychiat. Nervenkr., Vol. 183, pp.163–74

[7] Ramoser,H., Muller, J. and Pfurtscheller,G. (2000)‘Optimal spatial filtering of single trial eeg during imagined hand movement, Rehabilitation Engineering’, IEEE Transactions on Neural Systemsand Rehabilitation, Vol. 8, pp. 441-446.

[8] Kubler,A.,Kotchoubey, B., Kaiser,J.,Wolpaw, J. R. and Birbaumer (2001) ‘Brain–computer communication: unlocking the locked’,Psychol. Bull., Vol.127, pp.358–75.

[9] Neumann,N.,Kubler, N.,Kaiser, J.,Hinterberger, T. and Birbaumer,(2003) ‘Conscious perception of brain states: Mental strategies for brain–computer communication’Neuropsychologia, Vol.41, pp.1028–36.

[10] Pfurtscheller, G. and Neuper, (2001)‘Motor imagery and direct brain–computer communication’,Proc. IEEE, Vol.89, pp.1123–34

[11] AL Goldberger , LAN Amaral, L Glass, JM Hausdorff, PChIvanov,RG Mark, JE Mietus, GB Moody , C-K Peng , HE Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals.” Circulation 101(23):e215-e220 [Circulation ElectronicPages; http://circ.ahajournals.org/cgi/content/full/101/23/e215] ;2000 (June 13).

[12] Schalk,G., McFarland, D.J., Hinterberger, T., Birbaumer,N., Wolpaw,J.R. (2004)‘BCI2000: A General-Purpose Brain-Computer Interface (BCI) System.’, *IEEE Transactions on Biomedical Engineering*, Vol. 51, pp. 1034-1043.

[13] Niedermeyer, E., Silva, F. L. D. (2004) ‘Electroencephalography: Basic Principles, Clinical Applications, and Related Fields’.

[14] Moller MF (1993) A Scaled conjugate gradient algorithm for fast supervised learning, *Neural Networks* 6 (4): 525-533.

[15] Ahmed, K. S. (2011) ‘Wheelchair Movement Control VIA Human Eye Blinks’, American Journal of Biomedical Engineering, pp. 55-58.

[16] Yazdani,N., Khazab, F.,Fitzgibbon,S.,Luerssen,M.,Powers, D. and Clark, C. R. ‘Towards a brain-controlled Wheelchair Prototype’, School of Computer Science, Engineering and Mathematics, Flinders University, Adelaide, Australia

[17] Iturrate,I., Antelis, J. and Minquez, J. (2009) ‘Synchronous EEG brain-actuated wheel chair with automated navigation’, IEEE Internation Conference on Robotics and Automation, pp. 2318-2325.

[18] Ahirwar, M.K. andLondhe, N.D. (2012) ‘Classification of 2D hand movement with power spectrum estimation’, International Journal of Biomedical Engineering and Technology, pp. 1-10.