A Review on EEG Signal Analysis

Pooja J. Kulkarni
Assistant Professor
Department of Electronics and Telecommunication Engineering
Pillai HOC College of Engineering and Technology, Rasayani, India

Abstract— Electroencephalogram (EEG) is a tool which is used to monitor the electrical activity of the brain. EEG signals identify brain simulations accurately so that we can get correct information from the brain. This information can be used to classify with different mental tasks. It is very much useful in brain-disease diagnosis and assessment of various brain diseases. The raw data collected from the electrical activity generated in the brain undergoes through various steps, namely: signal acquisition, data preprocessing, feature extraction, and classification. This paper presents the different methods and algorithms which are used in each of the above mentioned processes.

Keywords: Electroencephalogram, Acquisition, Pre-Processing, Feature Extraction, Classification

I. INTRODUCTION

The recording of electrical activities associated with human brain in called as Electroencephalogram. EEG is the rhythmic potential generated by individual neuron in brain when excited or stimulated to produce action potential. It is used to detect different brain abnormities. During the test, the electrodes with the shape of tiny metal discs are placed onto the scalp of the patient during the test. These metal discs detect the small electrical charges from the brain cell activities. The signals generated dues to small electrical charges are of very low amplitude hence they are amplified up to certain level and shown as a graph on the screen. There are five types of EEG signals based upon their frequency ranges viz alpha, beta, gamma, theta, and delta. They don't have any specific function. There are no hard-and-fast rules about their functions but survey tells that higher the frequency of the wave, the more alert and awake you are.

The Table 1.1 gives you brief description about all five waves.

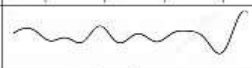
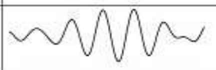
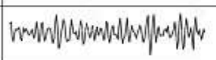
Name of the Brainwave	Frequency Range in Hz	Description	Diagrammatic Representation
Delta	0.5 to 4	In combination with other waves is typically for intuitive concentration, Unconscious waves, Deep Sleep without dreams	
Theta	4 to 8	During the peak movements, creativity and spirituality, Subconscious waves, paradoxical sleep phase state	
Alpha	8 to 13	Thoughts about something peaceful and relaxation	
Beta	13 to 30	During the internal active commentator or criticism, logical and mindful thinking	
Gamma	30 to 100	Strong focusing and concentration, High activity of the information processing	

Table 1.1: Different Waves generated in the brain

Depending upon your work that you are doing or what you are thinking or feeling at a particular moment, some waves are more dominant at that moment. To recognize such brain signals first we have to analyze EEG. The figure below shows various steps that are involved in EEG signal Analysis.

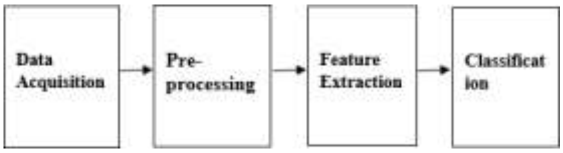


Fig. 1.1: Steps involved in EEG Signal Analysis

II. DATA OR SIGNAL ACQUISITION

Acquiring the EEG signal in a proper way is one of the important aspects in EEG Signal Analysis. From a technical perspective, it means that the signal quality should be optimal so that brain computer interface can work very effectively. And from an end-user perspective, it means with maximum comfort and without taking any extra inconveniences like washing the hair etc., the system should work efficiently.

There are three different EEG acquisition systems available. These systems differ in the type of electrode they used i.e.(gel-, water-, and dry- based), the amplifier technique, and the data transmission method. It is found that water-based system has the lowest short circuit noise level, the hydrogel-based system has the highest P300 spelling accuracies, and the dry electrode-based system causes the least inconveniences.

III. PRE-PROCESSING

In preprocessing step noise and artifacts presented in the raw EEG signals must be determined to reduce their impact on the feature extraction stage. In this step also, it is crucial to decide frequency and channel from EEG as it is produced from numerous electrodes [2] – [4]. Although EEG is intended to record cerebral movement, it likewise records electrical exercises from places other than the cerebrum. The recorded movements which are not originated from cerebral are called as artifacts. There arevarious ways of removing artifacts and noise from captured signals. Some of the frequently used techniques are given below.

A. Independent Component Analysis (ICA):

In ICA, the artifacts and the EEG signals are separated by considering that they are independent. During the elimination of artifacts, data is preserved. This algorithm is used to break the EEG data into components which are spatial-fixed and temporal independent. This algorithm is efficient when it comes to computation. Its performance is good when substantial amount of data is to be decomposed [1]. Decomposition of the signals require more computation. ICA performance found to be much better than classical Principle Component Analysis (PCA) in many applications. ICA is particularly good at removing ocular artefacts from EEG especially when the amplitudes of ocular signals and brain signals are comparable.

B. Common Average Reference (CAR):

Noise present in the EEG signal can be the common activity. CAR eliminates the noise by removing the common activity from the position of that we are interested in. The presence of the artifacts in EEG signal decreases the signal to noise ratio. In this method, having calculated the mean of all electrodes, noise is eliminated by removing the mean from all the electrodes. CAR outperforms all referencing methods and shows best classification accuracy results. The main causes of problems in calculating the averages in referencing methods are incomplete head coverage of EEG electrode arrays and finite sample density. In a study [5] it was found that CAR, specified with the current brain activity, accomplishes rate of identification of the states of mind that is second to none, when a user's brain activity is better balanced and he/she is more conscious of his/her mental state. It directly concerns the brain activity; hence this method acts as its natural filter.

C. Spatial Laplacian:

Surface Laplacian of the skull gives the estimation of current density entering or leaving the scalp through the skull. Only the outer shape of the volume conductor is considered and any details of volume conduction are not required [6]. It can efficiently eliminate ocular movements during the signal acquisition. Visual inspection is needed for large artefacts ranging from 50 microV (>50 microV) and by considering the shape of the artefacts the gradients of activities are obtained [7]. According to a study, the users do not properly control their brain activity at the very beginning of training process. This stage requires a more objective method. Spherical spline approach-based Surface Laplacian achieves best classification rates at this stage [5].

D. Adaptive Filters:

Adaptive filter is a computational device used for mathematical operations. It shows the interrelation between the input/output signals of the adaptive filter iteratively. It has ability to self-adjust the filter coefficients according to an adaptive algorithm. It works by changing signal attributes according to the specific feature of the signals under consideration [6]. When the signal and noise are overlapping, other filters may remove the signal of interest. Adaptive filters are used to avoid this problem.

IV. FEATURE EXTRACTION

After preprocessing of the EEG signal, different techniques are used to extract different features from EEG as discussed below.

A. Wavelet Transform:

P.Jahankhani and K.Revett have described that [8] wavelet transforms (WT) is suitable for non-stationary signals by analyzing various transient events in biomedical field They have also proposed that WT has advantage over spectral analysis. Wavelet Transform is proved to be an effective method for time frequency representation because of its important feature of providing accurate frequency information at the low frequencies and accurate time information at the high frequencies. The signals obtained in biomedical field always contain high frequency components with short time period and low frequency components with long time period. Thus this property is highly used in biomedical applications. The WT provides multi-resolution analysis of non-stationary signals.

B. Independent Component Analysis:

W.Zhou and J.Gotman have proposed [9] independent component analysis (ICA) in determination of source signal localization of independent components using single dipole model. The ICA aims to find linear representation of non-gaussian data. It is based on the linear transformation of the recorded EEG signal. ICA is a computational technique which finds the hidden factors that underlie sets of random variables. It is given by the following formula

$$x_1(t)=a11s1+a12s2$$
$$x_2(t)=a21s1+a22s2$$

where $a11$, $a12$, $a21$ and $a22$ are parameters and signals are given by $s_1(t)$ and $s_2(t)$ using the recorded signals x1(t) and x2(t). It is capable of producing subcomponents of multivariate signal.

ICA defines generative model for observed multivariate data which is typically given as large database of samples. For the decomposition of ongoing EEG signal ICA is used. This process separates mixed measured signals into an independent source [10] [11] it is used in the several applications. EEG data recorded at multiple scalp sensors are linear sums of temporally independent components which are arises from spatially fixed or overlapping brain areas. The propagation time delays are negligible. The eye artifacts and EEG are independent as their generating mechanisms are completely different. Because of all these reasons, this method is mainly used in Blind Source Separation. This method helps to identify the independent signals and also noise separation from brain signal.

C. Autoregressive Modeling:

AR model can be called as the representation of type of random process in different areas like signal processing, statistics etc. Its output variable depends on its previous values linearly. The method is used to classify imagined letters from EEG signal. The information retrieved from the imagined writing is used to find the writing disorders.

The amplitude and frequency of EEG signal changes while performing different mental tasks. These features can be captured and extracted using modeling techniques. AR model has been widely used for EEG analysis. It can be assumed that the real EEG can be approximated by AR process. AR is able to present an expression of a frequency domain signal characteristics [14]. AR model has problem to estimate the models' parameters when the measured EEG signal has limited length. It is required to have accurate values for prediction order and its coefficient values to model EEGs by using an AR model. High prediction order does not split the true peaks in the frequency spectrum and low prediction order combines near peaks in frequency domain [15].

D. Principal Component Analysis:

Principal Component Analysis is very effective technique to analyze the data and get the different patterns from it. Data compression can be used in PCA for projecting higher dimensional data to lower dimensional data. It reduces artifacts induced by Transcranial magnetic stimulation

(TMS) Hence it also reveals TMS evoked potential. PCA converts set of observations of possibly correlated variables into set of values of linearly uncorrelated variables by using orthogonal transformation. PCA is well established multivariate data technique which finds the direction of data. It is called as a statistical method of second order. PCA method decomposes covariance matrix Φ of mean zero with N*p observed data matrix. As $\Phi = LP^T$ where L=principal component score matrix and P=(P1,..Pp) is leading matrix of PCA. It finds signal similarity in signals to improve good accuracy for signal classification [35].

E. Empirical Mode Decomposition:

EMD method is used to decompose any complex into a finite and very small number of components. These components are the orthogonal basis of the original signal and they are called as intrinsic mode function (IMF). Amplitude and frequency components of the original signal can be obtained from IMF. Huang, et al. (1998) has proposed the Empirical Mode Decomposition and application of Hilbert transform which is known as Hilbert-Huang Transform. EMD is the fundamental part of HHT. It extracts time-frequency information from a nonlinear and non-stationary data signal. R. Oweis and W. Abdulhay have introduced [17] EMD method to recognize the seizure and Non-seizure artifacts. M.Chen and D. Mandic have developed [18] EMD method for qualitative assessment of EMD process. EMD is an emerging technique in biomedical signal analysis [19]. The principle of this method is decomposition of EEG signal into limited number of IMF. The process is carried out in two steps: IMF is obtained by using EMD algorithm. Applying HHT to the result of above step, the instantaneous frequency spectrum of initial sequence is obtained. Sifting process is used until final constant residue is obtained. As the number of IMF increases, the corresponding data becomes smoother. EMD also acts as filer because adjacent IMF has lower frequency than previous one.

V. CLASSIFICATION

Signal classification is the last stage in EEG signal analysis. Different algorithms and classifiers are used here to obtain accuracy in the defined system The primary goal of classification is to portray a limit between the classes. The classifier can be straightforward as solving a threshold for features or extra sophisticated, for example, machine learning algorithms. Some of the classifiers are discussed below.

A. Support Vector Machine (SVM) classifier:

Support vector mechanism gives better performance in classification process. Ii can be called as a mature classification algorithm. SVM has minimum computational complexity as compared to KNN, MLP etc. SVM has solid a great mathematical and theoretical base and it is not a black box method. SVM parameters are easier to handle than that of Deep learning. Signals are identified using Gabor transforms. Relevant features and energy ratios are extracted. Discriminant ratios are advanced to augment among the four sorts of HFOs and artifacts. At that point, the identified occasions are ordered utilizing a multiclass SVM. Cross-wavelet transform is used to extract six features from EEG signals based on RBF Kernel-based SVM. The extracted features are utilized to prepare SVM performing binary classification [22] [23].

A new technique based on edge weight strategy to decide concealed data from the cerebrum EEG signal. The EEG individual information is fed into the SVM preparing model if the individual has a strange epilepsy signal, joined with anomalous information before retraining and learning [24]. EEG channel is partitioned into four fragments, then into small clusters. The statistical features are extracted from clusters. These features vector is fed to an undirected weighted network. Classification is made using four Tunable-Q Wavelet transform (TQWT).

B. Decision Tree (DT) classifier:

Principle component analysis (PCA) based on graph entropy (PCA-GE) is used to select the optimal channels. These selected channelsare classified using the J48 decision tree to predict if the person is alcoholic or not. In another algorithm sparse principal component analysis (SPCA) is used to decrease the data dimensions and then fed to soft decision tree (SDT) classifier to predict a seizure or seizure free.

C. Logistic Regression (LR) classifier:

Logistic regression predicts the probability of an outcome that can only have two values. It also based on the covariance matrix. Independent component analysis (ICA) is used to reduce the dimension of the EEG signal, and then logistic regression is used for classification. Another algorithm makes use of lifting-based discrete wavelet transform (LBDWT) to capture transient features from the signal. The load is reduced using classical wavelet transform (CWT). The coefficients of LBDWT are fed to the inputs of two classifiers; these classifiers are used LR and ANN [25]

D. Linear Discriminant Analysis (LDA) classifier:

LDA analysis is mainly used to classify two class problems. LDA enables us to predict the membership of a sample of a class which is predefined by a set of features. This method is used to project original data on a more reduced space so that we can get the well separated features of different classes.

LDA is simple, robust, and produce accuracy as good as compared with the complicated method. LDA usually used for dimension reduction. It based on a covariance matrix, which searches for a linear combination of a variable that separated two classes efficiently [27]. The extracted features line length and energy are fed to several classifiers such as Quadratic discriminant analysis (QDA), K- Nearest Neighbour (KNN), and Linear discriminant analysis (LDA) are used for classifying the EEG signals to seizure or seizure-free.

E. K-Nearest Neighbors (KNN) classifier

k-Nearest Neighbor classifier is a non-parametric approach. It stores all available cases which classifies a given data point according to the majority of its neighbors. The algorithm follows two steps, i) finding the number of nearest neighbors by useing distance metrics like Euclidean distance and ii) classifying the data point into particular class using first step. It chooses nearest k samples from the training set. After that it takes majority vote of their class such that k should be an odd number so that it can avoid ambiguity if any.

KNN is widely used in statistical estimation and pattern recognition as it provides good classification accuracy [28]. Segmentation and selection (SnS) with root mean square (RMS) are used to extract feature vector from EEG signal and then classified using KNN.

The table below shows a summary of the different techniques used for classification and their performance in terms of accuracy sensitivity, and specificity.

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
KNN [25]	96.37	94.12	98.62
LDA	88.98	80.36	97.6
LR	96.03	95.22	96.84
SVM	97.24	96.14	98.34
DT	97	96.2	97.53

Table 5.1: comparisons among different classifiers (data base- bonn university, germany)

VI. Conclusion

This paper gives complete review of basic four steps involved in EEG signal Analysis i.e. Data Acquisition, Pre-processing, Feature Extraction and Classification.

In pre-processing, the signal is filtered to extract the EEG from recorded data and throughout any unwanted artifacts. Feature extraction is used extract a specific data from the signal to use these data as an input for the classification stage. Finally, the paper focused on the different EEG classification techniques and different classifiers in Bonn University, Germany dataset. Hence an accurate method can be chosen for better results based upon applications,

References

[1] T.P. Jung, S. Makeig and C. Humphries, "Extended ICA Removes Artefacts from Electroencephalographic Recording", Advances in Neural Inf. Processing Systems, Cambridge, Vol. 10, 1998.
[2] S. Kappel, D. Looney, D. Mandic, and P. Kidmose, "Physiological artifacts in scalp EEG and ear-EEG," Biomed. Eng. Online, vol. 16, no. 1, p. 103, 2017.
[3] L. Wang, W. Xue, Y. Li, M. Luo, J. Huang, and W. Cui, "Automatic Epileptic seizure detection in EEG signals Using Multi-Domain Features Extraction and Nonlinear Analysis," pp. 1– 17, 2017.
[4] A. Savelainen, "An introduction to EEG artifacts,", pp. 1–67, 2010.
[5] Mourino, J & Millan, Jose del R. & Cincotti, Febo & Chiappa, S & Jan´e, Raimon & Babiloni, Fabio, "Spatial filtering in the training process of a Brain-Computer interface", 2001.
[6] Siyi Deng, William Winter, Samuel Thorpe, Ramesh Srinivasan, "EEG Surface Laplacian using realistic head geometry", Int. J. of Bio- electromagnetism, Vol. 13, No. 4, pp. 173-177, 2011.
[7] C. Tandonnet, B. Burle, T. Hasbroucq, F. Vidal, "Spatial Enhancement of EEG traces by Surface Laplacian Estimation: Comparison between Local and Global Methods", Clinical Neurophysiology 116 2005 pp.18–24.
[8] P. Jahankhani, V. Kodogiannis and K. Revett, EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks, IEEE 2006.
[9] A. Soetraprawata, Feature Extraction of EEG- P300 Signals Using Nonlinear Independent Component Analysis ,International Journal of Mechanical & Mechatronics Engineering IJMME-IJENS Vol.13 No.2.April 2013.
[10] F.Shiman, S.H.Safavi, F.M.Vaneghi, M.Oladazimi,M.J.Safari and F.Ibrahim, EEG Feature Extraction using Parametric and Non- Parametric Models,IEEE International Conference on Biomedical and Health informatics, 2012.
[11] I. Daly, R.Scherer and G.Putz, FORCe:Fully Online and automated artifact Removal for brain- Computer interfacing," IEEE Trans.On Neural Health system and Rehabilitation Engineering, 2014.
[12] W. Zhou and J. Gotman, Automatic removal of eye movement artifacts from the EEG using ICA and the dipole model, Progress in Natural Science 2009.
[13] A.Zabidi,W.Mansor, Classification of Imagined Writing from EEG Signals using Autoregressive Features,IEEE 2012.
[14] X.Ma, The research of brain-computer interface based on AAR parameters and neural networks classifier, IEEE 2011 Int. Conference on Computer Science and Network Technology.
[15] M. Hoodgar, A.mehranzadeh and F.Forootan, Proposing an Effective Feature Extraction Model for EEG Signals to Enhance Quality of Hand's Motion Detection, Journal of Academic and Applied Studies Vol. 3(2) February 2013.
[16] E.Braack, M.J.Putten, Reduction of TMS Induced Artifacts in EEG Using Principal Component Analysis, IEEE Trans. on Neural System and Rehabilitation Engineering, vol.21,no.3,May 2013.
[17] R.Oweis and E.Abdulhay, Seizure Classification in EEG signals utilizing Hilbert- Huang Transform, Biomedical Engineering Online 2011.
[18] J.Fleureau,L.Senhadji, Turning Tangent Empirical Mode Decomposition:A Framework for Mono and Multivariate Signals, IEEE Trans.on Signal Processing vol.59,no.3,Mar.2010.
[19] M.Chen,D.Mandic and M. Ungstrup, Qualitative Assesment Of Intrinsic Mode Functions Of Empirical Mode Decomposition, IEEE 2008.
[20] C.Park, D.Looney, D.Mandic, Time Frequency Analysis of EEG Asymmetry Using Bivariate Empirical Mode Decomposition,IEEE Trans. on Neural System and Rehabilitation Engineering,vol.19,NO.4,Aug 2011.
[21] R. Pachori, Research Letter Discrimination between Ictal and Seizure-Free EEG Signals Using Empirical Mode Decomposition,Hindwi Publication,2008.
[22] N. Jrad, A. Kachenoura, I. Merlet, F. Bartolomei, A. Nica, and A. Biraben, "Automatic detection and classification of High Frequency Oscillations in depth-EEG signals," vol. 64, no. 9,pp. 2230-2240, 2017.
[23] S. Sengupta, D. Chanda, A. Mitra, and S. Dutta, "Computer Aided Technique for Epilepsy Detection Using Crosswavelet Transform and RBF-Kernel Based Support Vector Machine," 2nd IEEE Int. Conf. Next

Gener. Comput. Technol. (NGCT - 2016), pp. 501– 505, 2016.
[24] C. Chen, Z. Liu, H. Li, R. Zhou, Y. Zhang, and R. Liu, “EEG Detection Based on Wavelet Transform and SVM Method,” Proc. - 2016 IEEE Int. Conf. Smart Cloud, SmartCloud 2016, pp. 241– 247, 2016.
[25] H. Rajaguru, “Non Linear ICA and Logistic Regression for Classification of Epilepsy from EEG Signals,” Int. Conf. Electron. Commun. Aerosp. Technol., pp. 577–580, 2017.
[26] A. Subasi and E. Erc, “Classification of EEG signals using neural network and logistic regression,” Comput. Methods Programs Biomed, vol. 78, no. 2, p. 87—99, 2005.
[27] E. Tessy, and S. Manafuddin, “Time Domain Analysis of Epileptic EEG for Seizure Detection,” international conference on next generation intelligent systems, 2016.
[28] V. L. Kaundanya, A. Patil, and A. Panat, “Classification Of Emotions From Eeg Using K- NN,” vol. 3, no. 2, pp. 103– 106, 2015.
[29] U. Awan, H. Rajput, G. Syed, R. Iqbal, and I. Sabat, “Effective Classification of EEG Signals using K-Nearest Neighbor Algorithm,” International Conference on Frontiers of Information Technology, 2016, pp. 120–124.

