Multiresolution Analysis in EEG Signal Feature Engineering for Epileptic Seizure Detection

John Martin R. School of Computer Science Bharathiar University Coimbatore. India Sujatha S. Department of Computer Science Bharathi Women's College (Autonomous), Chennai, India Swapna S. L. Faculty of Computer Science Jazan University Kingdom of Saudi Arabia

ABSTRACT

In biomedical engineering, many attempts are being reported over the years for automated diagnosis of various brain disorders by classifying EEG (Electroencephalography) signals. Various machine learning algorithms are adopted to address different scenarios in EEG classifications. Feature engineering is playing a vital role in order to enhance the classification efficiency particularly in signal processing applications. This paper elucidates multi-resolution analysis (MRA) of feature engineering and demonstrates how the distinctive features are being engineered in wavelet domain. The implementation results are placed in the form of feature distribution diagrams and provide clear indications in feature selection for epilepsy seizure detection through classification.

General Terms

Machine Learning, Signal Processing

Keywords

Wavelets, Feature Engineering, Electroencephalography, Discrete Wavelet Transform, Multiresolution Analysis.

1. INTRODUCTION

Neurological disorders are very common and more than one percentage of the world population is suffering from brain diseases like epilepsy. Identifying brain disorders is sill challenging even by measuring the brain electrical activity through EEG. As the EEG signals are non-stationary, time of the particular event is crucial in identifying disorders. Many automated detection systems are proposed for diagnosing brain diseases like epilepsy [1]. Machine learning is playing major role in automated diagnosis of brain disorders as it is involving large volume of EEG data. Extracting features for machine learning classifiers is inevitable to attain accuracy in predictions.

Despite the fact that the wavelet domain based feature engineering is an ideal method of feature extraction and selection in EEG signal processing, it is also an effective tool for preprocessing the EEG signals which will ease the feature selection process even without reducing the dimensionality of features at least in some of the classifiers. In many EEG classification problems, wavelet based statistical features are extracted without preprocessing the signal data. Discrete Wavelet Transform (DWT) based EEG signal preprocessing is a more effective through MRA analysis. This paper reveals the effectiveness of DWT by processing the EEG signals for identifying distinctive features for classification problem of epileptic seizure detection.

This paper analyses with EEG signal data processing techniques with more emphasize on feature extraction in wavelet domain. A short briefing of EEG signal data is given in the next section. Section three elaborates various feature engineering approaches in time, frequency and wavelet domains for EEG classification. In section four, commonly used wavelet families for EEG feature analysis is discussed. Section five focused on DWT with threshold denoising for Multiresolution Analysis (MRA) of feature engineering with feature distribution diagrams of sample EEG data.

2. SIGNAL DATA2.1 Electroencephalography (EEG)

Applied electrophysiology techniques are important in diagnosing neurological disorders. The brain electrical activity is measured by using electroencephalogram (EEG) and its evoked potentials. The EEG signals are being generated by cerebral cortex.

Impulsive electrical fluctuations of voltage potential at the cortical surface are in the range of 100 to 1000 mV, but at the scalp are only 10 to 100 mV. Different locations of the cortex generate discrete potential fluctuations, which can also differ in the waking and sleep states, and also at stimulated states [2].



Figure.1 Sample of Normal EEG recording (Data points Vs Amplitude)

According to the clinical reference [2], in normal adults, the EEG electrical activity consists of sinusoidal (figure-1) oscillations occurring at 8 to 12 Hz called the alpha rhythm. The physical state of eye opening, mental activity, and drowsiness reflects the alpha rhythm. Human actions faster than 12 Hz will be the beta activity normally present over the frontal areas and may be especially prominent in patients receiving heart barbiturate or certain drugs. The electrical activity slower than 8 Hz is divided into delta activity (1 to 3 Hz) and theta activity (4 to 7 Hz).

Normally adults after the age of 60 years may show increased theta frequencies. Delta activities usually not present in adults on wakeup state but may show on sleep state. The amplitude theta and delta activities correlate closely with the depth of sleep. In new born children slow frequency EEGs are abundant, but these disappear progressively on maturation.

Physiological changes in brain activity are assessed by using EEG. Many abnormalities in EEG are not specific but

suggestive in some conditions like epilepsy, herpes encephalitis, metabolic encephalopathy etc.



Figure.2 Sample of Epileptic Seizure EEG recordings (data points Vs amplitude)

Electrical discharges on epileptic conditions are usually spikes or sharp waves (figure-2) that occur interictally in patients with epilepsy and sometimes in persons who do not experience seizures but have a genetic tendency to epilepsy. In epileptic condition these discharges may be focal or generalized, depending on the seizure type.

2.2 EEG Database

In most of the experiments, the publicly available EEG database owned by the University Of Bonn, Germany [3] is used. The database consists of five sets (A, B, C, D, and E); each one containing 100 single-channels EEG segments with the duration of 23.6 seconds. Sets A and B were taken from healthy patients in eye open and eye closed states respectively. Sets C, D, and E were taken from five patients with epileptic conditions; where C and D measured in seizure-free intervals and set E only consists of seizure activity. Samples of all five EEG segments are illustrated in figure. 3.



Figure.3 Sample EEG signal segments A, B, C, D and E (data points vs. amplitude)

3. FEATURE ENGINEERING APPROACHES

Feature engineering is a process of moulding data in such a way to increase classification efficiency. It is a classical term used in machine learning which ensembles the processes like data preprocessing, noise reduction, feature extraction, feature selection and data dimensionality reduction. In EEG classification, different feature engineering techniques were adopted by the researchers to address different scenarios in machine learning models.

3.1 Time Domain Features

Time domain features are amplitude related statistical features such as energy, power, and mean, and variability, regularity are tested with variance, coefficient of variation and total variation. In time domain, the epileptic seizure events are detected by analyzing discrete time sequences of EEG segments. These kinds of analysis can be performed through histograms of the EEG time segments.

A simple time-domain seizure detection method presented by Runarsson and Sigurdsson and suggesting for estimating the histograms for two variables: the amplitude difference and time gap between peak values as well as minima, as the features [4]. The estimated values of the histogram bins are the features for classification of an EEG segments as a seizure or non-seizure.

Lina Wang, et al. developed a comprehensive approach of feature engineering by combining all three domain features including statistical features in time domain and obtained higher average accuracy [5]. Neural Network based BCI model proposed by Ankita Mazumder, et al., [6] used time varying adaptive autoregressive algorithm (TVAAR) for extraction of features in time domain. Changjian Yang et al. brought a fuzzy logic system using time domain statistical features for recognition of EEG signals [7].

Time domain approach of feature engineering only provides spatial features, but temporal information is missing. In the event of epileptic seizures, signal might have uncertain oscillations, thus the frequency component is required to classify seizures.

3.2 Frequency Domain Features

Frequency is one of the prime components used to measure the occurrence of the events at precise time. As EEG is nonstationary, different frequency bands are used to locate the events. When an EEG signal is represented by its frequency component and all related features are estimated in frequency domain, it is called as frequency domain analysis.

For EEG seizure detection Fourier transform magnitudes are being used as frequency domain features. Amjed S Al-Fahoum et al. used frequency domain features for comparing the performance of classification for epileptic seizure detection with features in wavelet domain [12]. Frequencymoment signatures are adopted in the work proposed by Khamis et al. for building patient-specific epilepsy seizure detection through classification [9].

As EEG signals are non-linear and non-stationary, there is a difficulty to characterize different activities of EEG signals with certain mathematical models. In order to address this issue, Acharya et al. proposed a method for the detection of normal, pre-ictal, and ictal conditions from recorded EEG signals [10]. They uses four entropy features namely, phase entropy 1 (S1), phase entropy 2 (S2), approximate entropy (ApEn), and sample entropy (SampEn), for discriminating features for ictal, pre-ictal, and inter-ictal activities.

3.3 Wavelet (Time-Frequency) Domain Features

Time and frequency domain analysis of features have their own limitations in dealing with EEG signals. Even if the time domain analysis provides spatial feature components, the frequency content information is missing for EEG analysis. Frequency domain method may use temporal information for extracting features, but only after windowing the function. These limitations can be resolved by using time-frequency domain analysis.

Yuanfa W et al. notified in his research [11] for automatic detection of epileptic seizures that three levels DWT with Db4 wavelet efficiently performs three-class classification using multiclass sparse extreme learning machine. Amjed S et al. had performed a detailed study on the feature engineering with frequency and time-frequency domain methods for EEG classification. Widely used feature extraction methods such as TFD, FFT, EM, DWT and ARM are taken into consideration and concluded that each method has its own advantages and highly depends on the signal to be analyzed for the application [12].

A new proposal by Yang Li et al. suggests Multi-scale Wavelet Method is ideal to analyze time-frequency distribution of biomedical signals like EEG [13]. Similarly Oliver Faust et al. recommends wavelet based EEG processing for computer-aided epilepsy diagnosis after reviewing research outcomes reported in the recent past [14].

4. WAVELET FAMILY

Wavelet transform will be an effective time-frequency analysis tool for analyzing the ephemeral nature of EEG signals as it unifies different approaches towards bio-medical classifications.

Wavelet features are playing major role in evaluating transient events in biological signals. Variety of wavelet families is available for signal characterization. Few widely used mother wavelet families for EEG classifications include Harr, Daubechies, Symlet and Coiflet. Selection of suitable wavelet is crucial for the analysis of signals through wavelet transform. Based on the biomedical signal to be analyzed, the mother wavelet is chosen [14] [15].

As mentioned in [15], based on the classification accuracy and computational time obtained in the experiment, it was found that Coiflet of order 1(Coif1) is the best wavelet family for analysis of EEG signal as the support width of the mother wavelet function resembles that of the EEG signal and also has a compact filter length, thus reducing the processing time. This argument is being challenged by many researchers [16][17] and recommends Haar and, second and fourth order Daubechies (Db2, Db4) wavelets for signal preprocessing and feature extraction and were provided better accuracy in the recent classifications.

5. WAVELET TRANSFORM

Signal processing achieved milestones by the evolution of various signal decomposition methods due to their unique nature in processing signals. Several wavelet transforms namely Discrete Fourier Transform (DFT), Discrete Wavelet Transforms (DWT), Empirical Mode Decomposition (EMD, Single Valued Decomposition (SVD) and their variants were widely used for seizure detection and prediction applications and projected their efficiency in that applications.

Although many other time-frequency feature engineering approaches prevailing for signal processing such as EMD, SVD, ICA and PCA [18], DWT based wavelet feature analysis is identified as effective for time-frequency domain analysis due to its multiscale approximation feature.

Another highlight of DWT based feature engineering is that DWT is used for both signal noise reduction as pre-processing and feature extraction. According to Lina Wang et al. [5], multiresolution analysis (MRA) of feature engineering produced better EEG signal processing results. Wavelet Energy and entropy are considered as the prime features for wavelet analysis as reported by Yatindra et al. [8].

Wavelet domain feature engineering using DWT and its variants are used in combination with many machine learning classifiers for epileptic seizure detection. Few ideal research works reported since 2015, their methodologies and the outcomes are placed in the table (Table 1) for reference.

Many classification works were reported in the recent days using DWT but all the works are not ideal for analysis, as the datasets used is not clear. The few works mentioned here were used novel approaches for classification after wavelet analysis. The performances of the works were validated with public datasets obtained from University of Bonn, Germany.

From the data shown in the table.1, it is very well observed that the multiresolution analysis using DWT is dominating now in EEG signal feature engineering for epileptic seizure detection in combination with diversified classifiers. It is also noticed from the recent works that the SVM and FFNN classifiers are used most with DWT as feature engineering tool. Multiresolution analysis of signal processing and feature engineering is enhancing the classification accuracy considerably and is notable with recent works.

		-			
Year	Related Works	Method	Wavelet Family	Classifier (s)	Class. Accuracy (%)
2015	Gajic D et al. [21]	DWT	Db4	Quadratic classifiers	99.0
2015	Juárez-Guerra E et al. [16]	DWT, MODWT	Haar, Db2, Db4	FFANN	99.26
2017	Lina Wang et al. [5]	DWT	Db4	KNN, LDA, NBLR, SVM	99.25
2017	Yuanfa Wang et al. [11]	LDWT	Db4	SELM	98.4
2017	Jesus Martinez-del-Rincon et al. [22]	DWT	Db4	BoW, SVM	94.68
2018	Tzimourta, Tzallas et al. [23]	DWT	Db4	SVM	Above 99

 Table 1: Summary of most relevant state-of-the-art works since 2015, uses DWT based Feature Engineering with Bonn,

 Germany EEG Database

5.1 Discrete Wavelet Transform (DWT)

In DWT, wavelet is decomposed into multiple levels using low-pass and high-pass filtering as illustrated in figure.4. The original EEG signal (0-64Hz) is initially decomposed into high (32-64Hz) and low (0-32Hz) frequency bands which embodies detail and approximation of the input signal in level one. Then low frequency at the first level approximation is further divided into high (16-32Hz) and low (0-16Hz) frequency bands which represents detail and approximation at level two. Similar way low frequency bands will be decomposed at each level till level four as shown in figure.5. Expansion into further levels of decomposition may provide more clarity to differentiate features so as to facilitate the optimization.



Figure.4 Wavelet decomposition and reconstruction



Figure.5 Four level EEG signal decomposition

5.2 Wavelet threshold Denoising

Wavelet threshold method is a technique in DWT for preprocessing non-stationary signals like EEG by denoising as defined in [19]:

$$\lambda = \sigma \sqrt{2 \log N} \tag{1}$$

Where λ is wavelet threshold and σ is standard deviation of the noise and *N* represents number of sample points. Here the word 'noise' is mentioned as a standard term used in signal processing, but in EEG signal processing noise is in the form of sharp waves which are not significant for identifying events in the EEG signals [20]. It is significant to remove some frequency bands appearing in the decomposed signal bands by using wavelet threshold method so that distinguished features can be obtained for further analysis. The original input signal and the corresponding signals obtained after DWT threshold denoising are given in figure.6 and figure.7 respectively.

In the denoised distribution, D1 sub band is treated as noise which contains high frequency signals which is not a relevant factor for epileptic detection. D2 sub band is treated by eliminating high frequency signals with lower magnitude using threshold value. The denoised EEG signal will facilitate to extract distinguishable features than original signal especially for epileptic event detection.



Figure.6 Original EEG Signal after 4-level decomposition



Figure.7 Denoised EEG Signal after applying wavelet threshold

5.3 Feature Extraction

Feature vectors are calculated from signal data using MRA analysis and the coefficients are retained for further processing. Very common time-frequency domain DWT based statistical features for classification include mean average value, standard deviation, energy and spectral entropy. The energy factor at each level of decomposition is calculated as:

$$EyD_{i} = \sum_{j=1}^{N} |D_{ij}|^{2}, \quad i=1,2,...,l$$
$$EyA_{i} = \sum_{j=1}^{N} |A_{ij}|^{2}$$
(2)

The spectral entropy is calculated as:

$$SEN_{i} = \sum_{j=1}^{N} D_{ij}^{2} \log(D_{ij}^{2}) \quad i=1, 2, ..., 1$$
(3)

Where, i= 1,2,...,l decomposition level, N represents number of detail or approximate coefficients at each level of decomposition. Mean factor is found using:

$$\mu_i = \frac{1}{N} \sum_{j=1}^{N} D_{ij} \qquad i=1,2,...,l \qquad (4)$$

The standard deviation at each decomposition level is given by:

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^{N} \left(D_{ij} - \mu_i\right)\right) \qquad i=1,2,\dots,l \qquad (5)$$

In *l* level decomposition, particular feature vector of the given input signal is FD_1 , FD_2 ,..., FD_1 , from detail frequency band and FA_1 from approximation of the last decomposition level as illustrated in figure.6.

6. RESULTS

Two sets (100 EEG segments of healthy patient, 100 EEG segments of epileptic patient during seizure) of EEG signals are analyzed using MRA by using Db4 wavelet function. Mean average value (MAV), standard deviation (SD) and average energy (AvEnergy) across different decomposition levels of 100-segment EEG signals of healthy patient is represented in distribution diagrams given in figure.8.



Figure.8 Feature distribution of 100-segment EEG signals of healthy patient a) Mean Average Value (MAV) b) Standard Deviation (SD) c) Average Energy

Figure.9 shows the distributed features of mean average value (MAV), standard deviation (SD) and average energy (AvEnergy) across different decomposition levels of 100-segment EEG signals of epileptic patient during seizure activity.



Figure.9 Feature distribution of 100-segment EEG signals of Epileptic Seizure a) Mean Average Value (MAV) b) Standard Deviation (SD) c) Average Energy

On viewing the average energy (AvEnergy) factor of healthy subject shown in diagram 8c, it clearly indicates that the D2, D3 and D4 frequency bands corresponding to beta, alpha and theta are significantly lower in magnitude which signals normal EEG. While comparing this normal EEG with epileptic seizure EEG illustrated in 9c, the D3 and D4 frequency bands have elevated magnitude value which are almost similar magnitude value with A4. It shows the abnormalities in EEG signal. In this way the distinctive features can be identified through MRA analysis of DWT.

7. CONCLUSION

From the distribution diagrams shown in figure-8 and figure-9, it is evident that the DWT based multiresolution analysis is an effective tool for identifying distinguishable statistical features and to prepare the feature set for classification. The features extracted through this analysis can be subjected to dimensionality reduction to attain accurate classification of EEG signals for Epileptic seizure detection. There are more prospects to employ diversified approaches for dimension

reduction and classification with this signal processing method.

8. ACKNOWLEDGMENT

We gratefully acknowledge the Department of Epileptology at the University Hospital of Bonn for providing public access to their EEG database.

9. CONFLICT OF INTERESTS

No potential conflict of interest was noticed by the authors.

10. REFERENCES

- Mingyang Li et al., "Automatic epileptic EEG detection using DT-CWT-based non-linear features", Biomedical Signal Processing and Control, Vol. 34, pp. 114–125, 2017.
- [2] Robert B. Daroff et al., Bradley's Neurology in Clinical Practice, Elsevier Saunders, 2012, Ch.32A.
- [3] EEG time series data source, Department of Epileptology at the University Hospital of Bonn, October 2012, Available: http://epileptologiebonn.de/cms/front_content.php?idcat=193&lang=3.
- [4] TP Runarsson, and S Sigurdsson, "On-line detection of patient specific neonatal seizures using support vector machines and half-wave attribute histograms", in proc. The International Conference on Computational Intelligence for Modelling, Control and Automation, Vienna, Nov 2005, pp. 673–677.
- [5] Lina Wang, et al., "Automatic Epileptic Seizure Detection in EEG Signals Using Multi-Domain Feature Extraction and Nonlinear Analysis", Entropy, *19*(6), 222, 2017.
- [6] Ankita Mazumder, et al., "A Back-propagation through Time based Recurrent Neural Network Approach for classification of cognitive EEG states", in Proc. 2015 IEEE International Conference on Engineering and Technology (ICETECH), India, 2015, pp. 1-5.
- [7] Changjian Yang, Zhaohong Deng, Kup-Sze Choi, and Shitong Wang, "Takagi–Sugeno–Kang Transfer Learning Fuzzy Logic System for the Adaptive Recognition of Epileptic Electroencephalogram Signals", IEEE Transactions on Fuzzy Systems, Vol. 24, Issue. 5, pp. 1079 – 1094, 2016.
- [8] Kumar Y, Dewal M L. and Anand R S, "Relative Wavelet Energy And Wavelet Entropy Based Epileptic Brain Signals Classification", Biomed. Eng. Lett, Vol. 2, pp. 147–157, 2012.
- [9] H Khamis, A Mohamed, S Simpson, Frequency-moment signatures: a method for automated seizure detection from scalp EEG. Clin. Neurophysiol. 124(12), 2317– 2327 (2013).
- [10] UR Acharya, F Molinari, SV Sree, S Chattopadhyay, KH Ng and JS Suri, "Automated diagnosis of epileptic EEG using entropies", Biomed. Signal. Process. Control, vol. 7(4), pp. 401–408, 2012.
- [11] Yuanfa Wang, Zunchao Li, Lichen Feng, Chuang Zheng, and Wenhao Zhang, "Automatic Detection of Epilepsy and Seizure Using Multiclass Sparse Extreme Learning Machine Classification", Computational and Mathematical Methods in Medicine, vol. 2017, Article id. 6849360, June 2017.

- [12] Amjed S. Al-Fahoum and Ausilah A. Al-Fraihat, "Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains", ISRN Neuroscience, Volume 2014, Article ID 730218, Feb 2014.
- [13] Yang Li, Wei-Gang Cui, Mei-Lin Luo, Ke Li and Lina Wang, "High-Resolution Time–Frequency Representation Of EEG Data Using Multi-Scale Wavelets", International Journal of Systems Science, vol.48(12), pp. 2658-2668, 2017.
- [14] Faust O, U. Rajendra Acharya, Hojjat Adeli and Amir Adeli, "Wavelet-Based EEG Processing For Computer-Aided Seizure Detection And Epilepsy Diagnosis", Seizure, vol. 26, pp. 56–64, 2015.
- [15] Tapan Gandhi, Bijay Ketan Panigrahi and Sneh Anand, "A comparative study of wavelet families for EEG signal classification", Neurocomputing, vol. 74, pp. 3051–3057, 2011.
- [16] Juárez-Guerra E, Alarcon-Aquino V and Gómez-Gil P, "Epilepsy Seizure Detection in EEG Signals Using Wavelet Transforms and Neural Networks", in. Lecture Notes in Electrical Engineering, New Trends in Networking, Computing, E-learning, Systems Sciences, and Engineering, vol. 312, Elleithy K. and Sobh T. (eds), Springer, 2015.
- [17] Aminian, F., et al., "Electroencephalogram (EEG) signal classification using neural networks with wavelet packet analysis, principal component analysis and data normalization as preprocessors", In proc. Twenty-First MAICS 2010 Midwest Artificial Intelligence and Cognitive Science Conference, 2010, pp. 55-62.
- [18] Alotaiby et al., "EEG seizure detection and prediction algorithms: a survey", EURASIP Journal on Advances in Signal Processing, Vol. 2014, Issue. 183, Dec 2014.
- [19] Montefusco L and Puccio L, Wavelets: Theory, Algorithms and Applications, Academic Press, London, UK, 2014.
- [20] Williams J.W and Li Y, "A New Approach To Denoising ERG Signals-Merger of Translation Invariant Wavelet And ICA", Int. J. Biom. Bioinform, Vol. 5, pp. 130–148, 2011.
- [21] Gajic D, Djurovic Z, Gligonjevic J, Gennaro S.D and Gajic I.S, "Detection Of Epileptiform Activity In EEG Signals Based On Time-Frequency And Non-Linear Analysis", Front. Comput. Neurosci., Vol. 9, Issue. 38, 2015.
- [22] Jesus Martinez-del-Rincon et al., "Non-linear classifiers applied to EEG analysis for epilepsy seizure detection", Expert Systems With Applications, Vol. 86, pp. 99–112, 2017.
- [23] Tzimourta K.D, Tzallas A.T, Giannakeas N, Astrakas L.G, Tsalikakis D.G and Tsipouras M.G, "Epileptic Seizures Classification Based on Long-Term EEG Signal Wavelet Analysis". In proc. IFMBE, Singapore, Maglaveras N., Chouvarda I., de Carvalho P. (eds) Precision Medicine Powered by pHealth and Connected Health, Springer, 2018, vol. 66.