EEG Signal Processing for Seizure Detection using STFT and DWT

Zain Manasia
Department of Computer Science
Columbia University
New York, NY 10027
zm2401@columbia.edu

***Abstract* – Epilepsy is a neurological disorder characterized by large-scale electrical activity in the central nervous system. This activity can be quantified using electroencephalography (EEG). Traditional seizure detection methods are done using neurologists measuring amplitude spikes for a set duration of time. I propose a method to quantify seizure activity using short time Fourier transform and discrete wavelet transform to analyze these multi-resolution signals.**

Index Terms* – *EEG, seizure detection, STFT, DWT

I. Introduction

Epilepsy is a neurological disorder that is characterized by disturbed electrical rhythms in the central nervous system. This condition affects roughly 1% of the world's population and causes symptoms such as loss of motor control or consciousness. This arrhythmic brain activity can be quantified using Electroencephalography (EEG), a test conducted by physicians to measure electrical activity in the brain via electrodes placed locally on the scalp or intracranially. EEG is the preferred method to measure brain activity clinically due to its cost, effectiveness, and ability to be processed as a digital signal for easy quantification. Seizures come in many different forms such as Grand Mal, Absence, Tonic, and Myoclonic, all of which manifest differently in the patient. The EEG produces multi-channel signals representative of brain activity in amplitude versus time. Seizure activity is classified clinically as two times baseline amplitude brain activity that lasts for a duration greater than five seconds. The seizure signal can be broken down into four phases. The portion before the first seizure and after the last are called pre-ictal and post-ictal, respectively. Ictal and inter-ictal represent intervals between seizure activity. Like previously mentioned, clinical seizure detection relies primarily on analysing amplitude, but in research, the EEG signal can be analysed further. Signals can be analysed using time, frequency, wavelet, and empirical mode decomposition (EMD) domains. In this paper, I propose a method to analyse EEG signals using short time Fourier transform (STFT) and discrete wavelet transform (DWT) to gain higher signal resolution for predicted seizure activity.

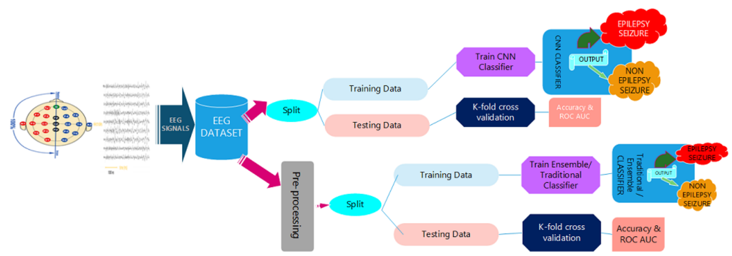

Fig 1: Visual diagram explaining EEG processing pipeline from collection to classification

II. Formulation and State of the Art

We can first discuss some methods of EEG signal processing that allow for optimal seizure detection. Time domain signal analysis describes how a signal varies over time. As previously mentioned, this method is most commonly used in the clinical setting. The signal is fragmented into time windows, known as epochs for the analysis. Alotaiby et al. proposed an analysis method by estimating the histograms of multichannel EEG signals. The collected signal was split into ten second epochs with five histograms measured from each segment. Histogram bins were assigned using a predefined criterion. EEG signal data was trained through this model, and then histogram bins were used as a classifying agent in deciding whether the epoch was ictal or non-ictal. This model was then tested on 309.9 hours of EEG data from five adult patients. The results indicated a specificity of 98.58%. [8]Runarsson and Sigurdsson designed a study to detect seizures using amplitude difference and time difference between peaks in the EEG signal. Two histograms were created, one of the amplitude differences and one for the time separation. The feature classification of seizure activity was estimated from varying values of amplitude difference and time separation with comparison to the pre-labeled data set. The data was clustered using a support vector machine (SVM) algorithm, generating a feature vector comprised from the above listed criteria. This method was able to achieve a specificity of 90%. The major problems encountered using this method were that the signal processing and extraction could only occur at when the half-wave peaks were found. Once this half-wave representation was found, this was a very fast clustering algorithm for time domain analysis.

[10]Yoo et al. proposed a method integrating a system on chip (SoC) interface, using the electric input to the device as features. The device had eight channels for input, and two classifier modules. The SoC was tested on 309.9 hours of EEG data from five adult patients and used the SVM classification method as well. In this algorithm, the feature vector was constructed by segmenting the signal, giving much higher accuracy values, and in turn a much more accurate feature vector. This approach was able to achieve a specificity of 84.4%. [3]Dalton et al. discusses a similar method to Yoo et al. but a more generalizable algorithm that can be used on devices but also networks. The algorithm extract seizure features such as mean, variance, zero-crossing rate, entropy, and root mean square from time domain signal data. Seizures were also detected using a gyroscope and a magnetic sensor for physical features. This algorithm was able to achieve a specificity of 84%.

The limitations of time domain analysis are addressed by frequency domain analysis. Frequency domain analysis describes the frequency spectrum in a magnitude versus frequency domain. This method allows for valuable spectral data to be extracted, which can elucidate different properties of epileptiform activity. A signal can be transformed to the frequency domain using Fourier transform. Fourier transform is a mathematical transform that coverts a signal in the time domain to the frequency domain by measuring the signal in every possible cycle and returns the components of the cycle such as amplitude, offset, and rate. The benefits of using frequency domain analysis for EEG signals is that this method is much better for direct feature extraction since the signal can be decomposed into sinusoidal signals with different frequencies. [4]Mursalin et al. discussed an approach combining both time and frequency domain analysis methods. Time domain features such as mean, median, skewness, standard deviation, kurtosis, etc. were collected. Frequency domain features such as mean, median, and standard deviation of wavelet coefficients were collected as well. This data was then applied to an improved correlation-based feature selection algorithm to gain more accurate features and then ran through a random forest classification algorithm to sort by these features. This method provided significantly greater specificity compared to normal correlative methods.

STFT is a signal processing method which applies a Fourier transform to a subsection window of data in a given time, which allows for the signal to be mapped into a two-dimensional function of frequency and time.

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{\frac{-j2\pi kn}{N}}$$

Fig 2: Mathematical representation of short time Fourier transform

The benefits of using STFT over Fourier transform are that it applies well to constantly changing signal characteristics and can consider a small part of the signal at each instance. The spectral mapping can be achieved by taking the magnitude squared of STFT to analyze the signal. This provides important features for seizure classification such as seizure onset. The issue that arises is that an increase in interval or window size decreases temporal resolution and decreasing the size of the interval will decrease the spectral resolution. This poses an issue as lower resolution leads to higher inaccuracy in prediction, so wavelet analysis can be applied to solve this problem.

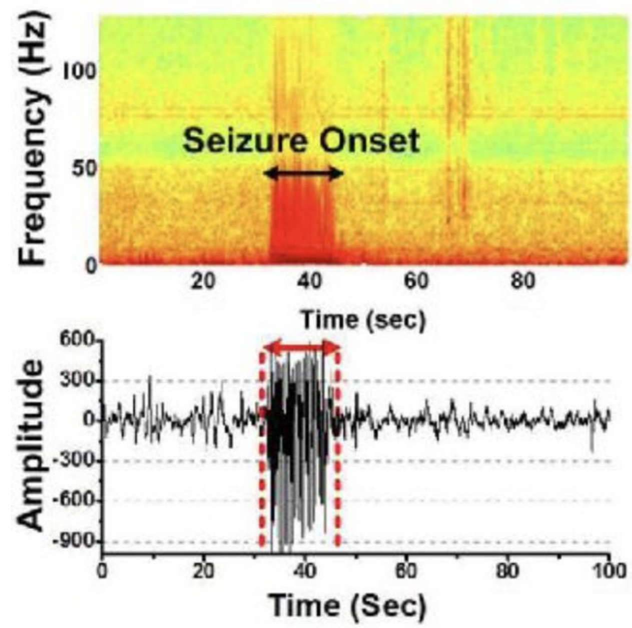

Fig 3: STFT spectral mapping of seizure onset in adult EEG recording

DWT can help solve the problem with resolution from STFT while also being able to provide information on the signal in time and frequency domains simultaneously. Using wavelet analysis, the provided signal is converted to a scaled and transformed mother wavelet, which is better for locating abnormalities within the signal itself. There are many examples of Mother wavelets such as Biorthogonal, Coiflets, Daubechies, Discrete Meyer, and Haar. While using DWT with a specific basis function, the signal is passed through a high pass filter and a low pass filter for difference and smoothing respectively. This functionality is unique to the DWT.

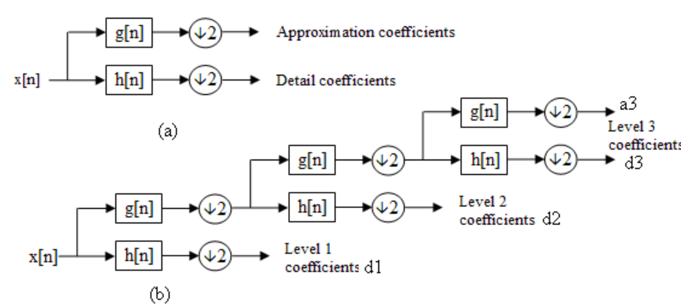

Fig 4: (a) Analysis wavelet filter banks (b) DWT decomposition using 3-stage approximation and detail coefficients

The Haar DWT is a suitable method for EEG signal processing because for each time series, detail coefficient vectors are produced. This provides information as to where and what abrupt changes in the signal occur. This in turn proves valuable information as components to detect points of change, which is useful for marking seizure activity.

III. Research Approach and formulation

In EEG signal processing, the most common spectral decomposition method is the fast Fourier transform as it provides a clear method to transform raw EEG-data from time-domain to frequency-domain, by a series of sinusoids. This process allows for lengthy and noisy data to be conveniently plotted in a frequency power-spectrum, allowing for easier detection of hidden features.[6] The authors employ the use of two transforms: discrete wavelet transform (DWT) and the Hilbert transform. DWT has many advantages over a FFT because it offers a time-frequency domain decomposition, while FFT does not capture location data (time-domain) and uses variable windows to produce good time-frequency localization. DWT decomposes a signal into detail (D_1) and approximation (A_1) sub-bands by passing the data through a low-pass (h[n]) filter with an impulse response of *h*, and a high-pass (g[n]) filter with an impulse response of *g*, convoluting the two[4]. The outputs describe the A_1 and D_1 coefficients of the transform, detailed in Equations 1 and 2 below.

$$A_1 = x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k] * h[2n - k] \quad (5)$$
$$D_1 = x[n] * g[n] = \sum_{k=-\infty}^{\infty} x[k] * g[2n - k] \quad (6)$$

The Hilbert transform, also tested in this study, is an important tool to calculate abrupt features of the signal and reconstruct the imaginary parts of the signal[3]. The Hilbert transform of a signal x(t) is the convolution of the signal *x(t)* with $1/\pi t$ to produce a signal H(x)(t). This transform is highly regarded for its use in feature detection and classification using amplitude and phase. The Hilbert transform of a signal H[x(t)] is defined in Equation 3 below.

$$H[x(t)] = x(t) * 1\pi t = \int_{-\infty}^{\infty} x(\tau)t - \tau d\tau \quad (7)$$

The authors gathered their EEG data from the Bonn Hospital Department of Epileptology dataset. They used two subsets (A and B) of healthy adult scalp EEG data and one subset (E) of epileptic adult intracranial data to process and train their kNN algorithm. The data was digitized at 173.61 samples per second using a 12-bit analog-digital converter. A band pass filter from 0.53 to 40 Hz was also applied to pre-process the data. The pre-processed EEG data was then fed through a two-stage DWT decomposition scheme and the Hilbert transform was then applied to collect the imaginary parts of the signal. The output, feature vectors, were collected after signal processing. These feature vectors were then statistically analyzed to reduce the dimensionality of the feature vectors, priming them for introduction into the kNN algorithms. The mean, maximum, minimum, standard deviation, and average power of the coefficients were calculated for each subsets A, B, and E of data. The data from each transform was then pipelined to kNN, a supervised machine learning algorithm used in simple clustering tasks. After clustering, two clusters were formed, and Euclidean distance was used to determine nearest neighbors from a central point. The clustering was conducted between A-E, and B-E subsets of data, each containing 100 EEG sets.

The results of this experiment demonstrated that both transforms were able to effectively filter and extract features from the EEG recordings. The Hilbert transform performed with 100% and 100% accuracy for the A-E and B-E clusters respectively, while the DWT performed with 100% and 96% accuracy for the A-E and B-E clusters respectively. Due to the discrepancy in the DWT results, it can be inferred that the Hilbert transform produces more distinct features, correlating to the higher rate of success in the kNN[1]. This study provides evidence that the Hilbert transform is equally as good, if not better than DWT in detection of epileptic seizures in some cases, as DWT is a commonly used transform in EEG feature extraction. This study does face the problem that the dataset is very homogenous and is not considerate of the wide gamut of EEG data that is used not only clinically but in research settings as well. Commonly used EEG data comes from human adult patients, pediatric patients, neonates, as well as mice, rats, and primates. There is a large variability in EEG signal phenotypes when analyzing such a broad spectrum of data, so there is a possibility that the Hilbert transform does not effectively produce as distinct features in some EEG types as it did in these adult human recordings. This can be seen as a direction to continue this study, as many types of EEG data are available for researchers and scientists to use and analyze.

Following a similar procedure from the literature above, for this project, I will be using STFT and DWT via Haar wavelets for analysis of the data. The data that will be used for this project is from the UC Irvine Machine Learning repository, consisting of adult EEG recordings from epileptic and non-epileptic patients. The first process is to import and restructure the dataset and visualize the EEG signal. The data will be reshaped from 11500 x 179 to 500 x 4094 and visualized. The next process will be to perform the Haar Transform to obtain detail coefficients in a matrix. The Haar transform is a 1-D transform with a level of 5 and approximation and detail coefficients. Haar wavelet transform is a method for representing a signal or a function as a sum of waves with different frequencies. It is a discrete wavelet transform, which means that it operates on a discrete set of data points rather than a continuous function. The Haar wavelet transform works by decomposing a signal into a series of wavelets, each of which corresponds to a different frequency range. The wavelets are constructed using a scaling function, which determines the overall shape of the wavelets, and a wavelet function, which determines the details of the wavelets. The scaling function is typically a smoothing function, while the wavelet function is a waveform that oscillates rapidly. The benefit of using this case is that it is computationally efficient and can be used to analyze sudden changes in the signal. The resulting data is then stored in a matrix. Each level of the detail coefficients serves as the output of a high-pass filter of the previous level's approximation coefficient. After each level, there is some level of decimation, so multi-level details represent signal decomposition at all resolution levels. Next, the variance changepoints are calculated at each DWT level. The variance change points provide information needed to plot the signal, which can then be used to compare to the original signal.

IV. Experiment and Results

For the experiment, the data was first obtained and visualized for verification of valid EEG signal data. The data is then reformatted and relabeled to fit the signal processing framework. The data is then passed through STFT function to obtain the correct features for seizure detection and formatting. The data is then passed through the DWT function using Haar wavelets to obtain the variance changepoints and for better signal resolution for seizure detection.

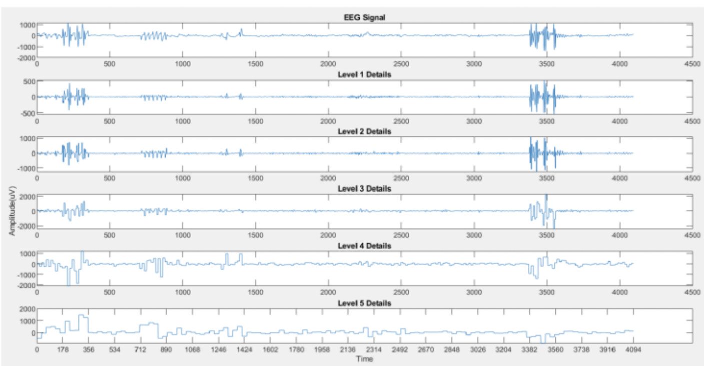

Fig 8: Multi-level detail coefficients from after application of STFT and DWT

The data is then examined to make prediction of seizure activity. The predicted seizures are then compared with the true seizure data to validate their accuracy.

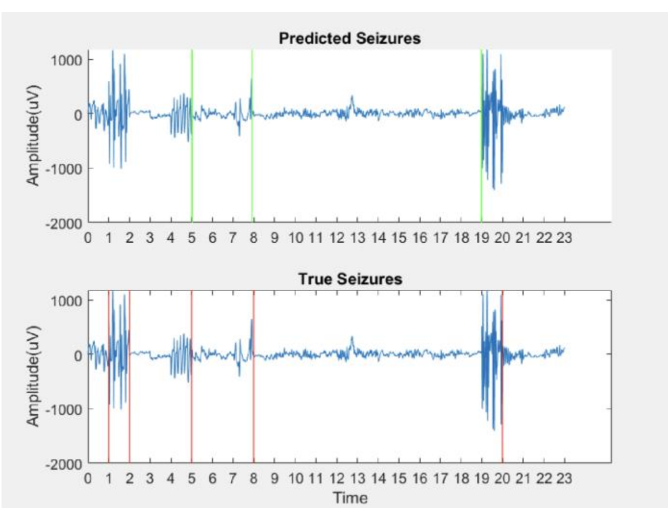

Fig 9: Comparison of Predicted Seizure with True Seizure marking changes in the signal

V. Conclusion and Future Work

EEG seizure detection methods using STFT, and DWT are usable methods to extract features for automated detection methods. The MATLAB program used properly processes the EEG signals, while applying STFT and DWT using Haar wavelets. The reasoning behind these two methods is explained as clear features and resolution. STFT was applied to observe the EEG signal in time and frequency domains. And DWT using 1-D Haar wavelets was applied to obtain detail coefficients for getting variance changepoints. Using this data, predicted seizure locations are able to be produced.

It can be seen that predicted seizure locations are fairly similar to true seizure locations. Future progress of this project can include using a CNN to quantify seizure prediction metrics with a larger data set.

References

[1]Alotaiby, T. N., Alshebeili, S. A., El-Samie, F. E. A., Alabdulrazak, A., and Alkhnaian, E. (2016) Channel selection and seizure detection using a statistical approach. 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA).

[2]Baldominos, A., and Ramon-Lozano, C. (2017) Optimizing EEG energy-based seizure detection using genetic algorithms. 2017 IEEE Congress on Evolutionary Computation (CEC).

[3]Dalton, A., Patel, S., Chowdhury, A. R., Welsh, M., Pang, T., Schachter, S., Olaighin, G., and Bonato, P. (2012) Development of a Body Sensor Network to Detect Motor Patterns of Epileptic Seizures. IEEE Transactions on Biomedical Engineering 59, 3204–3211.

[4]Mursalin, M., Zhang, Y., Chen, Y., and Chawla, N. V. (2017) Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier. Neurocomputing 241, 204–214.

[5]Niknazar, M., Mousavi, S. R., Vahdat, B. V., and Sayyah, M. (2013) A New Framework Based on Recurrence Quantification Analysis for Epileptic Seizure Detection. IEEE Journal of Biomedical and Health Informatics 17, 572–578.

[6]Polat, H., and Ozerdem, M. S. (2016) Epileptic seizure detection from EEG signals by using wavelet and Hilbert transform. 2016 XII International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH).

[7]Rana, P., Lipor, J., Lee, H., Drongelen, W. V., Kohrman, M. H., and Veen, B. V. (2012) Seizure Detection Using the Phase-Slope Index and Multichannel ECoG. IEEE Transactions on Biomedical Engineering 59, 1125–1134.

[8]Runarsson, T., and Sigurdsson, S. On-line Detection of Patient Specific Neonatal Seizures using Support Vector Machines and Half-Wave Attribute Histograms. International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC06).

[9]Tessy, E., Shanir, P. M., and Manafuddin, S. (2016) Time domain analysis of epileptic EEG for seizure detection. 2016 International Conference on Next Generation Intelligent Systems (ICNGIS).

[10]Yoo, J., Yan, L., El-Damak, D., Altaf, M. B., Shoeb, A., Yoo, H.-J., and Chandrakasan, A. (2012) An 8-channel scalable EEG acquisition SoC with fully integrated patient-specific seizure classification and recording processor. 2012 IEEE International Solid-State Circuits Conference.

[11]Zhang, X. (2021) Wavelet Filter Banks Using Allpass Filters. Wavelet Theory.