EEG Simulation Hardware for Realistic Seizure, Preseizure and Normal Mode Signal Generation.

Mohamed Shakir[1], Uvais Qidwai[2], Aamir Saeed Malik[1], Nidal Kamel[1]

[1]Neuro-Signal Processing Group, Centre for Intelligent Signal and Imaging Research Universiti Teknologi PETRONAS, Tronoh, Perak, Bandar Seri Iskandar, P.O 31750, Malaysia.

shakirmes@gmail.com, aamir_saeed@petronas.com.my, nidalkamel@petronas.com.my

[2] Computer Engineering Department, Qatar University, P.O 2713, Qatar

uqidwai@qu.edu.qa

Abstract. Unlike the ECG and EKG simulators which are very commonly used for these applications, there is a big need for seizure related EEG simulator. Having a hardware system that can be used instead of a real patient to generate realistic EEG signals is still in research phase. Here a framework is presented that can be used to realize EEG simulator in a pseudo-embedded form. This implies that the analog output depends on real patient data. The proposed hardware simulator will enhance researchers and hardware validators to simulate, validate and test their detection algorithms forehand, as well as for clinicians to use this system for training as well as for academic exercises. By utilizing significant spectral contents of real patient data, a simulated signal can be reproduced any time and can be modified for the seizure and pre-seizure cases by utilizing the model coefficients identified through standard ARMA system identification technique. A novel work has been done in producing simulated data based on empirical models of the real waveforms. Such a simulator will be very helpful in EEG related research since all the initial algorithms can be tuned to the controlled data first before going to the actual human subjects. Unlike the commercial ECG simulators, to the best of our knowledge, there is no such commercially available system that can be used for such research tasks. With controlled data types, healthy/normal, seizure and pre-seizure classes, tuning of algorithms for detection and classification applications can be attained. The model has been validated and tested with respect to accuracy of correct regeneration, false prediction rate, specificity, sensitivity and false detection rate.

Keywords: Bio-Signal simulator, EEG simulator, seizure, Pre-seizure, System Identification, ARMA model

1 Introduction

Seizures are brain disorders in which cells of the brain nervous system starts malfunctioning. As a result, it may generate the abnormal electrical signals that can cause a momentary malfunctioning of the human brain, resulting in change or complete loss of awareness. Normal message passing between the brain cells is thus temporarily disrupted and brain's messages are hindered by this disruption. The actual brain messages get mixed together and hence vanished. These seizures can be very dangerous and occurs with different frequency and some days can be very crucial. Detecting these signals can be done in clinical settings by using a number of modular techniques including Electroencephalography (EEG), Magneto encephalography (MEG) and the functional MRI (fMRI) which are the main neuroimaging modalities used for detecting seizures. This has led to the research that has resulted in many computational methods in clinical environments and prediction algorithms for EEG epileptic seizures. A real challenge is to have a detection system that can predict the seizures so that the patient can take appropriate safety measures. Different techniques have been used for prediction and detection of seizures in literature using Non-parametric Nonlinear Autoregressive (NNA) modeling of nonlinear dynamics of the brain activity [1], stochastic models [2], Principal Component Analysis (PCA) and Independent Component Analysis (ICA) [3], Time-Frequency analysis [4], several soft and hybrid computing techniques [5], etc.

However, from many findings, the generation of Electrocardiogram (ECG) cardiac abnormalities [6-9], is dependent on empirical algorithms and also on large amounts of similar dataset. This is not really a problem in ECG signals since many cardio-simulators are available commercially [10]. Thus, the EEG hardware simulator with EEG signal model for generation of variety of EEG signals that encompass various types of signatures as needed for the application such modeling of seizure or prediction of seizures is the motivation of this literature.

2 Method and Design

The proposed EEG bio-signal hardware simulator is intended for training as well as research applications where a user can repeatedly generate signals of normal (N-wave), pre-seizure (PS-wave) or seizure (S-wave) instances. In terms of EEG signal processing applications, this system can be used in order to design or develop detection and classification algorithms, other signal processing techniques, and other data mining applications for specific EEG signature recognitions.

The basis of the proposed simulator is real EEG data from real patients that were obtained from children's hospital at Boston. In Figure 1, the simulator can directly map the values from the patient data files into analog voltages by utilizing the power of LabVIEW environment and to implement it into an embedded Data Acquisition System (DAQ). The user can view these signals in real time as if these are being obtained through the EEG probes. Alternatively, the user may select to have signals which are generated using model. The advantages of simulators are that they can be used repeatedly; many times as needed for any test applications. In this literature, the model of the EEG signals is obtained using Autoregressive Moving Average (ARMA) process. With the capability of LabVIEW Data Flow program, the spike wave and the multi-spike wave of the EEG morphology are reconstructed [11] from the text file and corresponding voltages of the EEG signals are generated physically using National Instruments is a USB Data Acquisition device. This multifunctional device has high performance digital I/O, control channels, counter/timers and analog measurement.

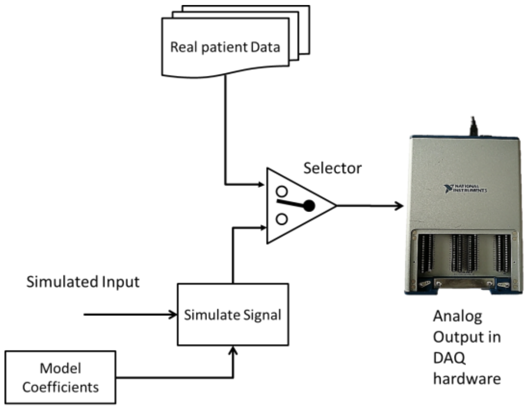


Fig. 1 Block diagram representation of the proposed system.

In case of real time EEG acquisition of data, we can directly wire the channel signal to the analog input for post processing including classification and identification of seizure types. This device also supports more accurate triggering, advanced timing, synchronization and high speed bidirectional real time streaming of signal and data by using NI-STC3 protocol technology. This protocol supports dedicated digital Input-Output and analog subsystems with independent timing engines for parallel digital I/O and analog to execution at different frequency synchronization with the help of 32-bit counters. Using the LabVIEW environment, a dedicated biomedical tool kit [12] is developed for the simulation of EEG from various file formats.

The LabVIEW program has two modes of operations: (i) read existing normal, pre-seizure and seizure patient waveform signals from an offline database (ii) simulate normal, pre-seizure and seizure waveform from the model. In either case, the generated waveform is sent to a USB-DAQ device, NI X Series DAQ, from National Instruments, for analog generation of the data. In case of real time EEG acquisition of data, we can directly wire the analog output channel signal to the receiving module for post processing including classification and identification of seizure types, etc.

The limitation of this DAQ of having only 4 analog output channels can be avoided by added analog signal multiplexers. However, due to trivial nature of this extension, it is not presented here and more focus is placed on the actual modeling and signal generation for one channel.

3 Direct Generation of EEG Signal

The seizure start points and end points have been clearly described in the data sets given by the CHB-MIT database. This database, collected at the Children's Hospital Boston, consists of EEG recordings from pediatric subjects with intractable seizures. Subjects were monitored for up to several days following withdrawal of anti-seizure medication in order to characterize their seizures and assess their candidacy for surgical intervention. The data used in this work for modeling the Seizure patients are from the CHB-MIT EEG scalp data repository. This free database consists of EEG data of 22 patients. The database consists of 129 files with one or more seizure with one or more pre-seizure incident. The file format has an extension of '.edf'. Here the sampling rate of the signal per second is 256. And the resolution is 16 bit.

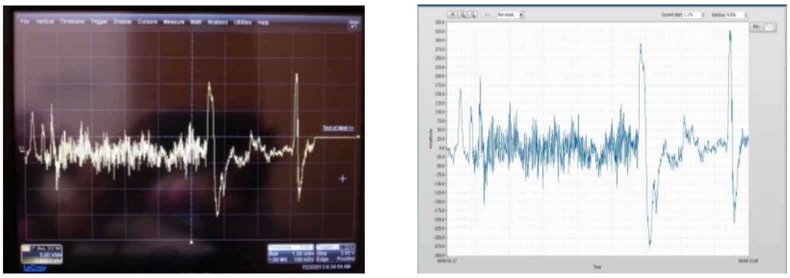
Fig. 2 EEG signal which consist of Normal, Seizure and Pre-Seizure Signals from the probe channels (FP1-F7), (T7-FT9), (P7-T7), (FP2-F8), (F4-C4), (FP2-F4),(C3-P3),(PF-01), (T7-P7) and (F7-T7) from CHB-MIT dataset.

Before sending the test signal to the NI-DAQ hardware, the signal should be checked in the graph to select the start point of the seizure occurrence, manually by entering the values or parsing from a text file which contains an array of the file name, time (ms) and the waveform data. Figure 2 shows a various zones of the EEG signals where the simulator can generate the various segments of the EEG signal from the database with the actual generated signal as voltage output from the analog output port of the DAQ module. Probe channels showed in Figure 2

((FP1-F7), (T7-FT9), (P7-T7), (FP2-F8), (F4-C4), (FP2-F4),(C3-P3),(PF-01), (T7-P7) and (F7-T7)) shows maximum seizure activities. Using statistical characteristics of the EEG signal, the characteristics of the pre-seizure and seizure signals can be classified into respective zones [13].

With respect to this, the user can calculate pre-seizure, seizure and normal EEG instances and corresponding portions of the datasets can be selected and sent to the hardware module. Figure 3 shows some of these signals as seen on an oscilloscope.

As can be seen in the Figure 3, the two waveforms are exactly the same thus proving a good duplication of the stored data values to be used in real-time or on-the-fly type techniques that require true signals in the form of voltages just like the actual EEG probe system. The result is shown for only one channel but it can be enhanced for up to 14 channels with the addition of some digital to analog hardware with the digital outputs of the DAQ module. However, this extension is quite trivial and all other procedural steps will remain the same.



(a) (b)

Fig. 3 The EEG signal segment (a) actual output from the DAQ module as viewed in an oscilloscope, and (b) display of the waveform from the database.

4 Modeling of EEG Signal

A linear dynamical system can be considered conceptually to represent the EEG signal generator (Figure 4). The EEG signal u(t) is the input variable and e(t) is the disturbances/noise, which generates y(t), the output signal. Here, the disturbance/noise refers to (1) probe noise, (2) Input noise, (3) Process noise, and (4) Uncertainties.

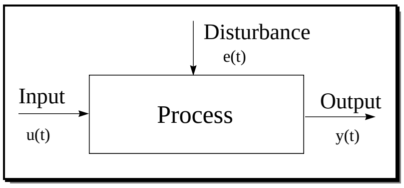


Fig. 4 Ideal Linear Dynamic Model.

In order to obtain a linear dynamical model from the data, a specific structure and order of the dynamic system should be known as a start. Actual model coefficients are then calculated as part of the system identification procedure. By exciting the system with some persistently exciting input, experiment for the identification can be done either by specially selected with respect to the nature of the system, or the usual inputs to the system. The input-output signals are observed over an interval of time and are fitted with the selected structure in order to identify the model coefficients. With respect to some of the classification data, the parametric model can be used to generate same type of EEG classes. For this, a suitable model should be determined beforehand, usually a definite order linear difference equation. Secondly, to determine the parameters, a particular statistical method is selected for the linear model structure coefficient identification. The validation of the obtained model is done to check if the model represents approximate actual model or the system. If it does not meet the actual model, then a more complex structure model is selected and correspondingly the parameters are then re-identified. The new model is again validated for the closeness with the actual system, and so on. The procedure is repeated until a satisfactory degree of tolerance is achieved.

Usually an Input-Output system such as the one shown in Figure 4 can be represented as the following difference equation:

$$[1 + a_1 z^{-1} + \cdots + a_{na} z^{-na}]y[z] = [b_0 + b_1 z^{-1} + \cdots + b_{nb} z^{-nb}]u[z] \qquad (1)$$

Here orders of the model are represented by the parameter ‘*na*’ and parameter ‘*nb*’. y[z] is the output and u[z] is the input at sample time [k]. Thus *na* essentially represents the number of the poles and *nb* the number of zeros. This can be written in a simplified manner as:

$$A(z)y(z) = B(z)u(z) \qquad (2)$$

Here the polynomials with the delay operator z^{-1} are represented by A(z) and B(z) and are given by:

$$A(z) = 1 + a_1 z^{-1} + \Lambda + a_{na} z^{-na}$$
$$B(z) = b_0 + b_1 z^{-1} + \Lambda + b_{nb} z^{-nb} \tag{3}$$

With the input-output measurements being available, an empirical system structure can be formulated as follows for y_k:

$$y_k = [y_{k-1} \quad y_{k-2} \quad \cdots \quad y_{k-na} \quad u_k \quad u_{k-1} \quad \cdots \quad u_{k-nb}] \begin{bmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_{na} \\ b_0 \\ b_1 \\ \vdots \\ b_{nb} \end{bmatrix} \tag{4}$$

Hence the parameters which are to be identified can be given as the following vector:

$$\theta = [a_1 \quad a_2 \quad \Lambda \quad a_{na} \quad b_0 \quad \Lambda \quad b_{nb}]^T \tag{5}$$

Usually in a gray box identification problem, the structure is known and order is estimated using various criteria. However, due to the complex nature of the EEG signals, the first problem was to identify the persistently exciting input source that could be used as a common source to generate the desired output. In many other disciplines, this can be achieved since the nature of input source is known or can be estimated beforehand. But for the EEG signal case does not fall into this scenario and hence a different approach was adapted.

- Using the real data from real patients (22 different patients with seizure history and clinical data collected during the seizure episode) the segments pertaining to the seizure occurrences were indicated by the neurologists and were isolated as the required seizure waveforms (S-waves). A segment of 20 seconds before that was marked as pre-seizure (PS-wave) while the segment of 20 seconds before the PS was marked as normal EEG (N-wave). If the window segment is below 20 seconds or set too low, the artifact present shows as false positives [14] and to find if dominant peak exists in the power spectral density , if the user is using this model/ simulator for seizure detection.
- Each waveform was then converted into its 1024 point Fourier spectrum with a sampling frequency of 265 Hz (standard clinical practice).
- Hence an Fs/2 spectrum provides all the frequency components present in that signal. Utilizing only significant components (up to 10 components), with

their corresponding frequencies and amplitudes, a new source signal is generated by summing these components for the given time length. The output of this step is shown in Figure 5.

- Next, the actual EEG segment signal is utilized as output signal of the gray box while the generated signal from previous step as the input to this box, an Autoregressive with exogenous (ARX) model was approximated.
- In order to get the appropriate model order, a brute force approach was utilized and many different combinations were tested for *na* and *nb* in the ranges of $3 \leq na \leq 15$, and $1 \leq nb \leq na$. Here orders of the denominator and the nominator are represented as '*na*' and '*nb*' respectively for the ARX transfer function. The condition for *nb* ensures the normal rational nature of the transfer function which is the preferred form due to many stability and implementation properties. The Validation for this model is performed at the end section after the generation of s-wave, ps-wave and n-waves.
- For each order, a model is identified and new simulated output is calculated. Then the Mean Squared Error (MSE) is calculated between this simulated and the actual EEG output waveform.

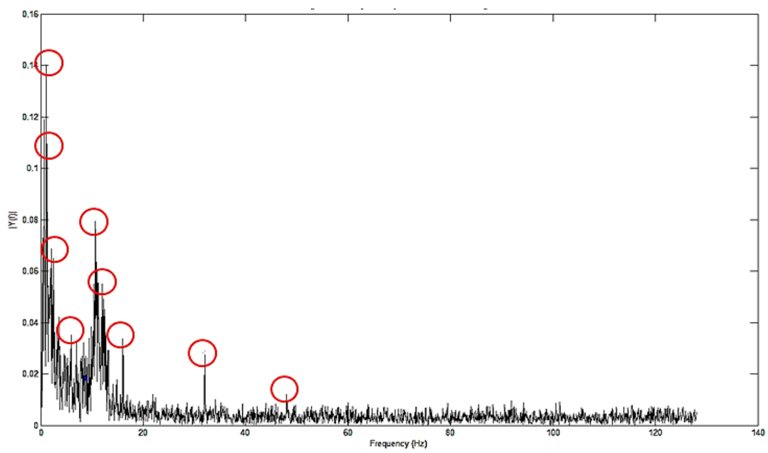


Fig. 5 Spectral selection from EEG signal, for the generation of generalized simulated signals. Circles represent the peaks of the prominent frequency components. [x-axis represents Frequency in Hz and y-axis about amplitude].

- Once all the combinations were tested, the model with the least value of the MSE was selected and the coefficients were saved along with the time segments from that specific data of the patient.
- Hence, this segment can now be re-generated any time for as many samples and for as long as needed by the user, this giving a very well suited freedom for signal length and samples needed in detection/classification algorithms for this class of signals.

Figure 6 shows a typical logic flow of the whole procedure:

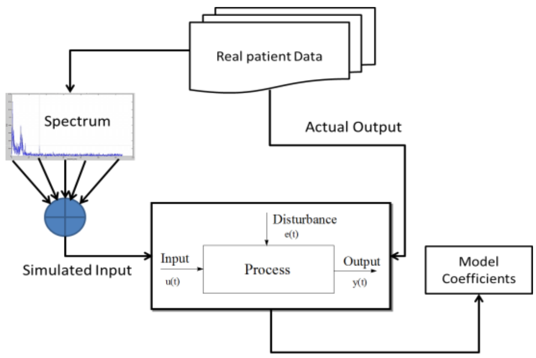


Fig. 6 Flowchart of the overall modeling process involved.

On the negative side, EEG signals are characterized by low signal-to-noise ratio, and non-stationary characteristics, which makes the processing of such signals for the extraction of useful information a challenging task. EEG data is generally considered to be non-stationary because its characteristics change with time, depending upon the mental states that are active at any given time instant. In order to handle this feature, one approach is to assume that over short time intervals the signal remains stationary. A batch processing algorithm is then applied to estimate the optimal parameters which best fit the measurements taken over each of these short time intervals. The performance of this criterion is tested by extensive Monte Carlo analysis [15], and it is also used to fit an ARMA model to real EEG data, indicating the presence of information bearing nonlinearities in the EEG signal.

5 Testing

Using the CHB-MIT EEG scalp data repository, several datasets were used for N, PS, and S waves. For the N-wave, most of the frequency components were quite similar, and hence a typical signal was selected to generate the simulated input signal. Once the input wave was formed, it was used with other types of waves to come up with model coefficients which could reproduce the desired signal when needed. A set of test images are shown in Figure 7 for the N-wave showing a very good match with the original signal as well as the appropriate coefficients of the model. Since both PS and S waves are critical, similar tests were also done on the Pre-seizure and Seizure signals (PS and S waves) and their final results are shown in Figure 8. In all of these cases, the y axis represents the normalized amplitude of the signals. X-axis is time axis with values in milliseconds. On applying equation (1), (2) and (3), the polynomials with the delay operator z^{-1} are represented by A(z) and B(z) and the following equation gives the resultant model:

$$\begin{aligned} A(z) = {} & 1 - 1.934 z^{-1} + 1.31 z^{-2} - 0.6087 z^{-3} + 0.3019 z^{-4} \\ & + 0.1604 z^{-5} - 0.3289 z^{-6} - 0.04284 z^{-7} + 0.3543 z^{-8} \\ & - 0.3067 z^{-9} + 0.2205 z^{-10} + 0.07694 z^{-11} - 0.1619 z^{-12} \\ & - 0.2488 z^{-13} + 0.3695 z^{-14} - 0.1516 z^{-15} \end{aligned} \tag{6}$$

$$\begin{aligned} B(z) = {} & 1.681 z^{-1} - 7.486 z^{-2} + 19.13 z^{-3} - 37.31 z^{-4} + 58.31 z^{-5} \\ & - 75.03 z^{-6} + 81.69 z^{-7} - 75.7 z^{-8} + 58.9 z^{-9} - 37.77 z^{-10} \\ & + 19.46 z^{-11} - 7.2 z^{-12} + 1.319 z^{-13} \end{aligned} \tag{7}$$

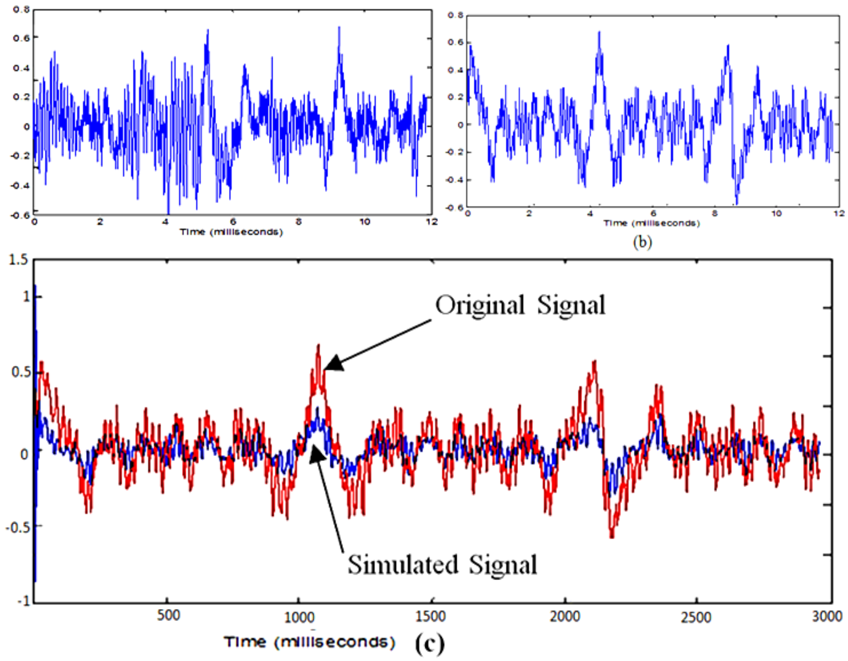
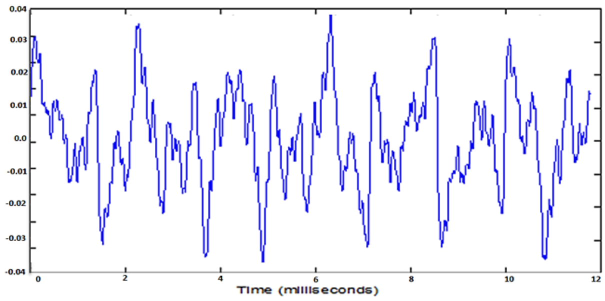


Fig. 7 Various signals for N-wave. (a) Original signal, (b) simulated excitation signal, and (c) simulated (blue) and original (red) signals.

Figure 7(c) shows the resultant simulated signals (blue color) along with the superimposed original signal (red color). Figure 8 shows the PS wave and S wave simulated by using this model. The N-wave, PS-wave and S-wave EEG signals cannot be distinguished directly by evaluating on the morphology of the wave, since the characteristics of these waves vary from patient to patient; therefore additional computational methods should be used in order to detect/distinguish between them in real-time.



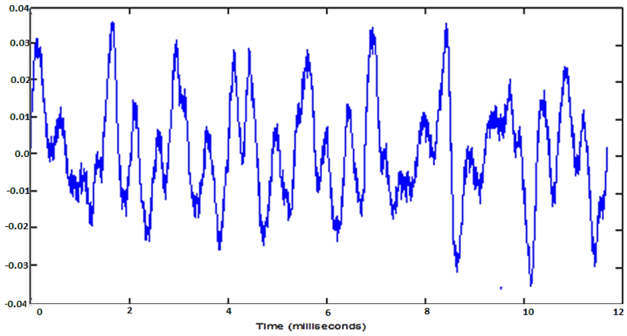


Fig. 8 Simulated Outputs for (a) PS-wave, and (b) S-wave with amplitude in y-axis and time in x-axis.

Since the EEG generated from this model is Polymorphic, and our verification is based on less computational algorithms, the generated signal is passed through a threshold-based decision on certain qualitative statistical measures, which made the classification possible using a simple look up table.

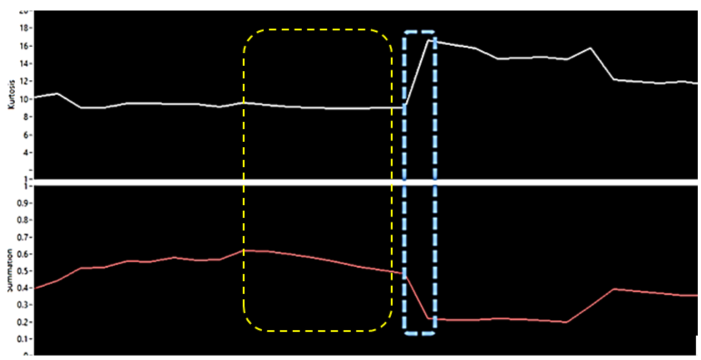


Fig. 9 The blue-dotted region shows S-wave class and yellow-dotted region shows PS-wave class characteristics based on Kurtosis and Summation on the y-axis respectively and time on x-axis.

Two features were isolated from the above list since they showed maximum differentiation of the two output classes and were found sufficient for the required classification and varification of the model. The EEG data set is processed by using the Kurtosis and Summation statistical tool. The property of the signal showed some distinguishable dynamics when it reached the point of PS-wave and S-wave respectively as in figure 9. The low frequency components of the EEG signal vanished or significantly reduced. At the same point, the Kurtosis of the EEG signal showed a spontaneous average rise in threshold as shown in Figure 9.

6 Results and Conclusion

In this paper, a strategy is presented in order to modify an existing Data Acquisition/Generation module to function as an EEG simulator. Such a simulator will be very helpful in EEG related research since all the initial algorithms can be tuned to the controlled data first before going to the actual human subjects.

With respect to the Table 1 results, this model has an overall 93.47% accuracy of correct regeneration of PS-wave and S-wave from the EEG data and from the model with less than 5% false prediction rate, Specificity of 93%, sensitivity of 93.97% and False Detection rate of 6.98%. In this work, statistical Rule-based classification is offered as an extension of verification of the model to achieve economically viable classification technique and to distinguish N-wave, S-wave and PS-waves.

TABLE I. MODEL VALIDATION RESULTS

Type	Positive	Negative
P-wave and PS-wave	*(True Positive)* 109	*(False Negative)* 9
N-wave	*(False Positive)* 7	*(True Negative)* 120

Unlike the commercial ECG simulators, to the best of our knowledge, there is no such commercially available system that can be used for such research tasks. The use of NI device and LabVIEW environment makes it very user friendly for anyone to develop a device tailored to their requirements. The data we have used is from actual patients and the system converts the text files for the data values into electrical voltages similar to what would be measured by the EEG probes. Further work is being done in producing more simulated data based on empirical models of the real waveforms. With controlled data types, healthy/normal, Seizure, and Pre-Seizure classes, the user can select the type and work on tuning their algorithms for detection and classification applications.

Acknowledgements

This publication was made possible by a grant from Qatar National Research Fund under its National Priority Research Program, for projects NPRP 09-292-2-113. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of Qatar National Research Fund.

References

1. Valdes P, Bosch J, Jiménez J C, Trujillo N, Biscay R., Morales F, Hernández-Caceres JL, and Ozaki T, The statistical identification of nonlinear brain dynamics:Progress report 1999. pp.278-284,Nova Science Publishing ,USA.
2. Fumikazu.M,Ruben.P, Pedro.A, Valdes.S, Tohru.O, A comparison of non-linear non-parametric models for epilepsy data. Vol.31, pp.41-57 Computers in Biology and Medicine, 2001.
3. Fumikazu.M, Pedro.A, Valdes.S, Eduardo.V, Jorge.B, Jobu.W, Hiroaki.M, Yoko.Y, Decomposing EEG data into space-time-frequency components using parallel factor analysis and its relation with cerebral blood flow,Vol. 4984 pp. 802-810 LNCS, Springer , USA, 2008.
4. Martínez-Montes E, Marquez-Bocalandro Y and Lage-Castellanos A, EEG Pattern Recognition by Multidimensional Space-Time-Frequency Analysis V Latin American Congress on Biomedical Engineering CLAIB 2011.
5. Stacey.W, Levan.Q, Mormann.F, Schulze-Bonhage A.What is the present-day EEG evidence for a preictal state? Vol.3 pp.243-251, Epilepsy Res. 2011.
6. Qidwai U, and Shakir M, Real-Time Patient monitoring system with Filter Bank and Fuzzy Classifier for Severe Cardiac Arrhythmia Detection and Notification System, 3rd International Conference on Environmental, Biomedical and Biotechnology, Singapore, Vol.3,pp.588-592, International Academy Publishing, USA 2013.
7. Qidwai U, Chaudhry J and Shakir M, Ubiquitous Monitoring System for Critical Cardiac Abnormalities, 3rd IEEE/EMBS Conference on Biomedical Engineering & Science, IEEE USA 2012.
8. Qidwai U, and Shakir M, Embedded System Design with Filter Bank and Fuzzy Classification Approach to Critical Cardiac Abnormalities Detection, pp.361-366 IEEE Symposium on Industrial Electronics and Applications, IEEE USA 2012.
9. Qidwai U and Shakir M, Filter Bank Approach to Critical Cardiac Abnormalities Detection using ECG data under Fuzzy Classification, pp. 405-412, Vol.5 (International Journal of Computer Information Systems & Industrial Management Applications 2013.

10. De Lazzari, C., G. Ferrari, Right Ventricular Assistance by Continuous Flow Device-A Numerical Simulation. Methods of Information in Medicine,Vol-46,pp-530-537 , Schattauer Publishers, Germeny, 2007.
11. Trumbower, R. D., & Enderle, J. D. (2003). Virtual instruments in undergraduate biomedical engineering laboratories. Engineering in Medicine and Biology Magazine, IEEE, 22(4), 101-110.
12. L. Ljung, System Identification: Theory for Users, 1987 ,Prentice Hall, USA
13. P. Mirowski, D. Madhavan, Y.LeCun and R.Kuzniecky , Classification of patterns of EEG synchronization for seizure prediction. Clinical neurophysiology: Journal of the International Federation of Clinical Neurophysiology. November 2009 (volume 120 issue 11 Pages 1927-1940
14. B. Boashash, M. Mesbah and P. Colditz , Newborn EEG seizure pattern characterisation using time-frequency analysis, Acoustics, Speech, and Signal Processing IEEE proceedings, Vol.2,pp.1041-1044,2001.
15. Simon G. Fabri, Kenneth P. Camilleri,Tracey.C, Parametric Modelling of EEG Data for the Identification of Mental Tasks. Biomedical Engineering, Trends in Electronics, Communications and Software, pp-367- 386, chapter 19, 2011