Harikumar Rajaguru et al /J. Pharm. Sci. & Res. Vol. 8(4), 2016, 214-219

ISSN:0975-1459
Journal of Pharmaceutical Sciences and Research
www.jpsr.pharmainfo.in

Weighted KNN Measures for Epilepsy Classification from EEG signals utilized in Telemedicine Applications with a PSO Based Reduced PAPR and BER Analysis

Harikumar Rajaguru[1], Sunil Kumar Prabhakar[2]
[1]Department of ECE, Bannari Amman Institute of Technology, Sathyamangalam, India
Email: harikumarrajaguru@gmail.com
[2]Department of ECE, Bannari Amman Institute of Technology, Sathyamangalam, India
Email: sunilprabhakar22@gmail.com

Abstract:
Electroencephalograph (EEG) is nothing but the collection of electrical signals of brain. EEG contains the most significant information about the activities of the brain. In this paper, the detection and classification of epileptic seizures in EEG signals is done with the help of Fuzzy Mutual Information (FMI) and Weighted KNN Classifier. Initially, the dimension of the EEG is reduced with the help of Fuzzy Mutual Information and then it is transmitted through a 2 x 2 Differential Space-time Block Coded (DSTBC) System. The DSTBC system is incorporated with a Particle Swarm Optimization (PSO) - Based Peak to Average Power Ratio (PAPR) Reduction Technique in order to obtain a reduced PAPR and Bit Error Rate (BER) at the receiver side. At the receiver, Weighted KNN measures is employed as a Post Classifier to classify the epilepsy risk levels from the EEG signals Thus the signals can be easily transmitted with the help of the system developed and at the receiver the signals can be classified easily, thereby enabling the doctors with an added advantage and aiding in the telemedicine application. The performance measures are analyzed in terms of specificity, sensitivity, time delay, quality values, accuracy, performance index measures, PAPR and BER.
Keywords: EEG, epilepsy, FMI, Weighted KNN, DSTBC

1. Introduction

One of the most common neurological disorder is the epilepsy which occurs in brain [1]. Due to a number of unrelated conditions such as consequences resulting from stroke, high fever, electrolyte-imbalances and toxicity, epilepsy can be caused and triggered. Being a chronic nervous disorder, epilepsy is characterized by recurrent seizures and duration may vary from a short lapse of muscle jerk to prolonged convulsions [2]. Seizures are generalized into two categories, namely partial and general. If the seizures are produced from a localized area in the brain, then it is called as partial seizure and if the seizure involves the entire brain accompanied by loss of consciousness, then it is called as a generalized epileptic seizure [3]. Both these types of epileptic seizures can affect the individuals at any age. The nature of these seizures, severity level and timing situations in which they come is a great social difficulty for the epileptic patient. The people around the epileptic patients are sometimes overprotective and restrict the patient from doing most enjoyable activities and hence such a condition leads to isolation and depression for the patient. People suffering from epilepsy cannot have stable jobs and a peaceful marriage life. Monitoring the activities of the brain through EEG is a technique and used for the analysis, classification and diagnosis of neurological diseases especially epilepsy [4]. EEG monitoring systems produce an enormous amount of data and so this entire visual analysis is not possible. With the utilization of computers and automated systems, recognizing electroencephalographic changes have become easier. The development of automated devices helps us so much because of the increased use of prolonged EEG recordings used for the proper evaluation and treatment of neurological disorders and to prevent the possibility of the misreading information.

1.1 Related Works:

A lot of works has been described in the literature for the EEG signal classification. The detailed analysis of EEG records for an epileptic patient utilizing wavelet transforms was done by Adeli et.al [5]. With the help of logistic regression and Artificial Neural Networks, the seizure detection was done automatically in EEG by Alkan et.al [6]. To classify the epilepsy from EEG signals, a Radial Basis Function Neural Network model was developed by Aslan et.al [7]. The usage of adaptive neuro-fuzzy inference system for classification was developed by Guler and Ubeyli [8]. With the help of Fast Independent Component Analysis and ANN, Kocyigit et.al classified the EEG records [9]. With the help of discrete wavelet transform and Approximate Entropy, Ocak detected the epileptic seizures automatically [10]. The epileptic seizures detection using the probability distribution based on equal frequency discretization was represented by Orhan et.al [11]. In the intra and extracranial EEG, recurrent Neural Network based prediction of epileptic seizures is done by Petrosian [12]. The detection of seizure activity in EEG by ANN was done by Pradhan et.al [13]. The epileptic detection using time-domain and frequency-domain features using ANN was proposed by Srinivasan et.al [14]. By combining Eigen vector methods and multiclass SVM's, the analysis of EEG signals was reported by Ubeyli [15].

2. MATERIALS AND METHODS

2.1 Acquisition of EEG data

Based on the standard International 10-20 system, by placing the electrodes on the scalp of the epileptic patients, the EEG is recorded. Totally 16 channels of EEG are recorded simultaneously for both the referential montages. The EEG data taken for the study is obtained from 20 different epileptic patients who were in the examination and treatment process in the Department of Neurology, Sri Ramakrishna Hospital, Coimbatore, India. The method employed here is bipolar method and it is employed in European Data Format (EDF). With the help of neurologist, each record with most distinct features had been obtained. The total number of artifacts considered here is of four types namely, motion artifacts, chewing, eye blinks and Electromyogram (EMG) artifacts. Just in order to differentiate the spike categories of waveforms, the inclusion of artifacts was done. A particular segment of EEG data is selected to train and test the component structures of the signals. The EEG recordings are done for a period of thirty minutes and each EEG record is continuous therefore it is divided into epoch of two second duration. To ignore the unwanted repetition and to trace the important changes in the activity, a two second signal is more than enough. The maximum frequency of the EEG signal is 50 Hz and therefore the sampling frequency is considered to be twice more than the maximum frequency and therefore 200 Hz is chosen. Every sample value indicates its respective amplitude values of the signal and therefore for a single epoch, totally 400 values are obtained. Therefore for our entire data, there are totally 25,600 sample values. To process this, it is quite a hectic task. So with the help of Fuzzy Mutual Information as a dimensionality reduction technique, the dimensions of the EEG data are reduced. Figure 1 shows the Block Diagram of the Work

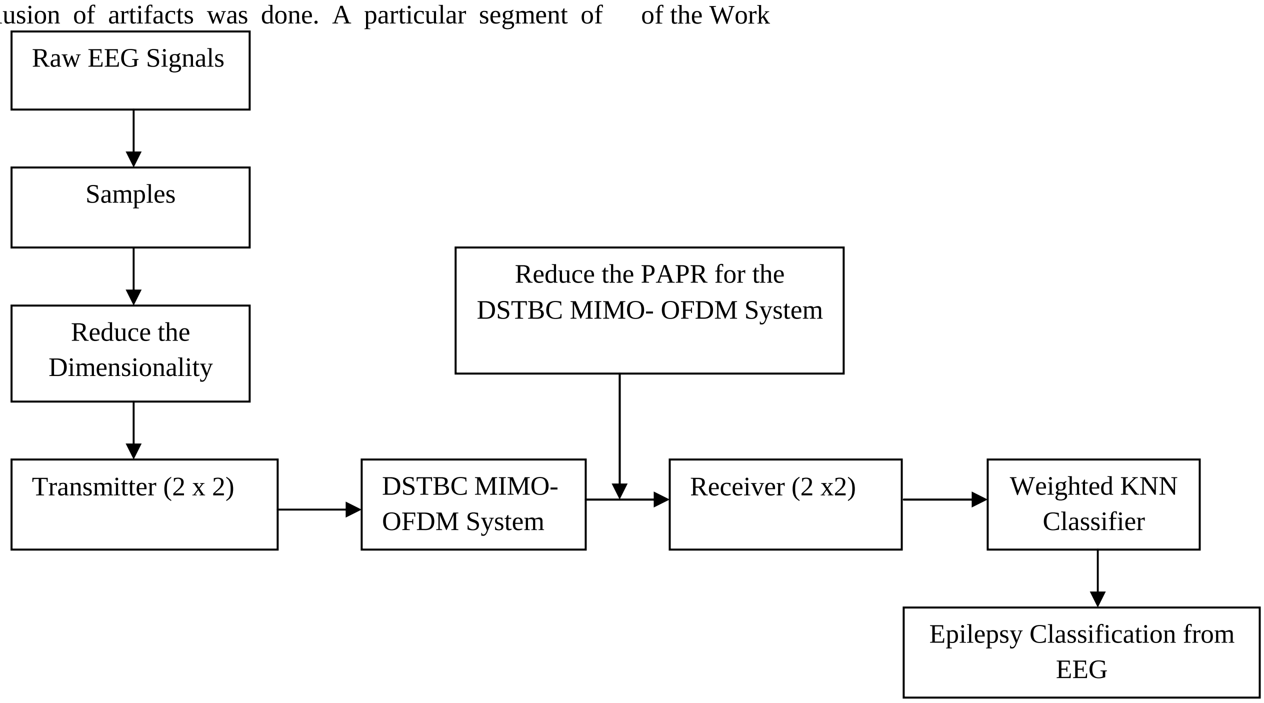

Figure 1. Block Diagram of the Work

2.2 Dimensionality Reduction using FMI

It is a filter method where the irrelevant features can be easily reduced. Enrichment of the mutual information is done using the fuzzy concept. Initially the discretization process is done and the number of clusters is assigned. The membership function of the fuzzy set is constructed using triangular membership function. According to [16], the fuzzy entropy is then calculated using class degree as follows

$$FFE(C/f) = \sum_{\tilde{A} \in V} \frac{S_{\tilde{A}}}{S} FE(\tilde{A})$$

The entropy of class C is calculated as follows

$$H(C) = -\sum_{i=1}^{n} P_i \log_2 P_i$$

The normalized Fuzzy entropy measure is then calculated as follows

$$NFMI(C, f) = \frac{H(C) - FFE(C/f)}{\min\{H(C), FFE(C/f)\}}$$

3. TELEMEDICINE APPLICATION IMPLEMENTATION

A Multiple Input Multiple Output (MIMO) system comprises numerous antennas at both the transmitter and receiver side [17]. To combat signal fading and to attain diversity gain, MIMO systems are used often. MIMO when combined with Orthogonal Frequency Division Multiplexing (OFDM) serves as a key technique for high speed wireless communication networks thereby improving the diversity gain [18]. When space time block codes are used with MIMO OFDM, then it is called as STBC MIMO OFDM system [19]. The coding is done only in the time domain (i.e.) OFDM symbols. The Differential STBC are highly useful in the context of wireless communication because they have no necessity to understand and identify the channel impairments present in the receive side for the sake of decoding the signal [20]. For this DSTBC MIMO OFDM System, a Particle Swarm Optimization (PSO) [21] based PAPR Reduction is performed to produce a low PAPR and BER.

An ordinary communication system incorporating space-time block coding technique with just two transmit antennas and one or more receive antennas is considered for the analysis. The information blocks of symbols in the transmitter side are passed to the next unit called differential space-time block encoder, where two symbols are embedded in each block. The code words of length $M = 2$ is generated by the space-time block encoder where M signifies to the total number of transmit antennas. The OFDM Modulator and the radio frequency (RF) front-ends obtain these code words and then it modulates the useful information onto the carrier frequency. On the constructed receiver side, up to N receiver antennas can be efficiently made use of for reception probably. The RF signals are completely down-converted and digitized in the RF front-ends and then finally passed to the differential space-time block decoder unit followed by the OFDM Reed Solomon demodulator unit. The interpretation of the received signals is done by the space-time block decoder and after that the received signals are obtained and generated for estimates as the transmitted information symbols, which are again provided simultaneously in blocks of two symbols.

3.1 PAPR in DSTBC MIMO-OFDM

In each and every N_T parallel OFDM transmitters, a particular block of D distinct complex-valued carriers say, $A_{\mu,v,} \mu = 1,..., N_{T,} v = 0,...D-1,$ is transformed into its respective time-domain using the Inverse Discrete Fourier Transform, that is,

$$a_{\mu,k} = \frac{1}{\sqrt{D}} \sum_{v=0}^{D-1} A_{\mu,v} . e^{j2\pi kv/D}, \mu = 1,..., N_{T,} k = 0,..., D-1.$$

On the combination of the time-domain samples into a specific vector $a_\mu = [a_{\mu,0,...,} a_{\mu,D-1}]$, the respective correspondence is written as $a_\mu = IDFT\{A_\mu\}$. Generally in almost all the wireless applications, all frequency-domain samples $A_{\mu,v}$ are expected to be obtained from similar constellation points with variance σ_a^2.

Since the carriers are statistically independent, the time-domain samples $a_{\mu,k}$ are complex Gaussian distributed in an approximated sense. This leads to a very high peak-to-average power ratio as

$$PAPR_\mu = \frac{\max_k |a_{\mu,k}|^2}{E\{|a_{\mu,k}|^2\}} = \frac{\max_k |a_{\mu,k}|^2}{\sigma_a^2}$$

As a versatile performance measure, the probability that the PAPR of an OFDM frame exceeds a given threshold, i.e., that the squared magnitude of at least one sample over the N_T antennas and D time steps is larger than tolerated: $|a_{\mu,k}|^2 > PAR_0 \sigma_a^2$ is carried out in literature. With this consideration, the Complementary Cumulative Distributive Function (CCDF) of the PAPR $P_r\{PAPR > PAPR_0\}$, clipping probabilities can be determined easily.

In MIMO-OFDM, since $N_T D$ instead of D time-domain samples are present and the CCDF of the PAPR[15] is represented mathematically as follows [22]

$$P_r\{PAPR_{MIMO} > PAPR_0\} = 1 - (1 - e^{-PAPR_0})^{N_T D}$$

3.2 PAPR Reduction Technique using PSO Based Active Constellation Extension Algorithm

For the standard Active Constellation Extension algorithm, PSO is applied and the PAPR and BER is analyzed.

Step 1: The program is started and then the input bits are generated randomly

Step 2: The serial data is converted into parallel data

Step 3: The input signals are modulated using appropriate modulation schemes like DQPSK. In DQPSK, the symbol information is encoded as the phase change from one symbol period to the next rather than as an absolute phase. In this case, the receiver has to detect phase changes and not the absolute value of the phase, which avoids the need for a synchronized local carrier.

Step 4: Meanwhile, the parallel data obtained is also directly computed using N point Inverse Fast Fourier Transform (IFFT)

Step 5: Then clipping process is done and it is converted into frequency domain. The out of band carriers are removed and then the Active Constellation Extension (ACE) algorithm [23] is applied

Step 6: The ACE algorithm can be initialized by selecting the parameters namely the target clipping level and the number of iterations, denoted by i.

Step 7: The initial target clipping level is assumed as A and then the clipping level is computed.

Step 8: The Particle Swarm Optimization is applied for further optimizing the clipping level before transferring to the anti-peak signal

Step 9: The clipping signal is transferred to the anti-peak signal subjected to the ACE constraint.

Step 10: Increase the total number of iterations counter to ten.

Step 11: The threshold value is calculated and then it is checked that whether PAPR>threshold value

Step 12: As a final step, the Complementary Cumulative Distribution Function (CCDF) plot versus probability of PAPR is computed and the plot is drawn

Step 13: The bit error rate is also computed for the DSTBC MIMO-OFDM system.

Step 14: Stop the program

4. WEIGHTED KNN AS A POST CLASSIFIER AT THE RECEIVER SIDE

The Nearest Neighbor rule (NN) is one of the simplest, oldest and traditional classifier used in the field of pattern classification [24]. If a set of training samples and query are given, the aim of NN algorithm is to find a point which matches very closely to the query followed by the assignment of its class label to the query. In case of a KNN algorithm, based on the majority role of its k-nearest neighbours, the query can be labeled in a training set.

Let $T = \{(p_i, l_i)\}_{i=1}^{N}$ denote the training set, where $p_i \in R^m$ is training vector and y_i is the corresponding class label. If a query x' is given, then its unknown class l' is assigned with the help of Euclidean distance functions and the prediction of the class labels of the query. In Weighted KNN (WKNN), the neighbours which are closer to each other are weighted more heavily and the neighbors which are far from one another are weighted less heavily. The weight w_i for the i^{th} nearest neighbor of the given p' is defined as follows

$$q_i' = \left\{ \frac{d(p', p_k^{NN}) - d(p', p_i^{NN})}{d(p', p_k^{NN}) - d(p', p_1^{NN})} \right\} if \quad d(p', p_k^{NN}) \neq d(p', p_1^{NN})$$

$$q_i' = 1 \;\; if \quad d(p', p_{k)}^{NN} = d(p', p_1^{NN})$$

Decided by the majority weighted noting, the classification result of the query is made as follows

$$l' = \arg\max_{y} \sum_{(p_i^{NN}, l_i^{NN}) \in T'} q_i' p \delta(l = l_i^{NN})$$

5. RESULTS AND DISCUSSION

5.1 PAPR Results

Initially the PAPR reduction for the Differential Space Time Block Coded-MIMO OFDM Systems are simulated according to the simulation parameters listed in Table 1. The PAPR reduction analysis is shown in Table 2 and BER Analysis is shown in Table 3.

Table 1 Simulation Parameters for PAPR Reduction in DSTBC MIMO-OFDM System

Modulation used	**DQPSK**
MIMO System analysed	2 x 2 MIMO-OFDM
Number of subcarriers	128
Maximum Iteration Count (POCS-ACE)	100
Maximum symbols loaded	1e5
Symbol rate	250000
No of time slots	2
Window function	Blackman-Haris
HPA Model	SSPA
No of frames	10
No of OFDM symbols/ frame	4
Bandwidth	5 MHz
Oversampling factor	4

Table 2 PAPR Reduction Analysis

Modulation Scheme used	PAPR reduction in dB (20 iterations)	PAPR reduction in dB (40 iterations)	PAPR reduction in dB (60 iterations)	PAPR reduction in dB (80 iterations)	PAPR reduction in dB (100 iterations)
DQPSK	3.9	4.1	4.3	4.6	5

Table 3 BER Analysis

SNR	**BER**
0	0.1797
1	0.1340
2	0.0938
3	0.0621
4	0.0375
5	0.0216
6	0.0112
7	0.0054
8	0.0022
9	0.0008
10	0.0004
11	0.0002
12	0
13	0
14	0
15	0
16	0
17	0
18	0

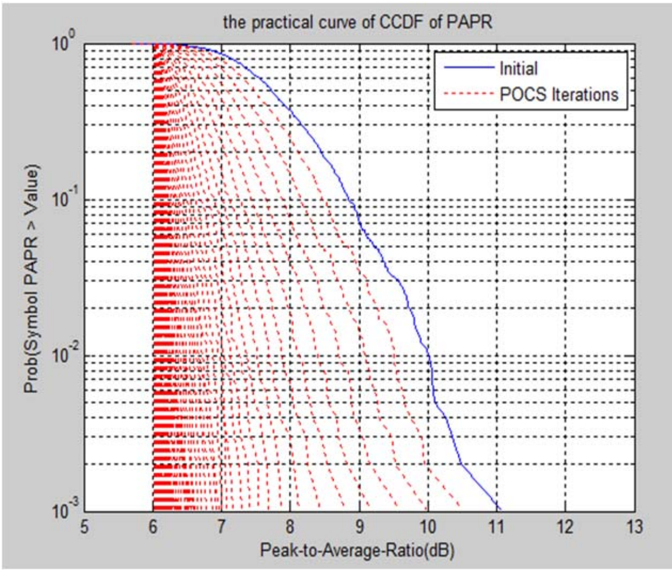

Figure 2 PAPR Reduction Curve for DSTBC MIMO-OFDM System for 100 iterations

5.2 Classification Results

For FMI as dimensionality reduction techniques and Weighted KNN as a Post Classifier at the receiver side, based on the Quality values, Time Delay and Accuracy the results are computed in Table 4 respectively. The formulae for the Performance Index (PI), Sensitivity, Specificity and Accuracy are given as follows

$$PI = \frac{PC - MC - FA}{PC} \times 100$$

where PC – Perfect Classification, MC – Missed Classification, FA – False Alarm,

The Sensitivity, Specificity and Accuracy measures are stated by the following

$$Sensitivity = \frac{PC}{PC + FA} \times 100$$

$$Specificity = \frac{PC}{PC + MC} \times 100$$

$$Accuracy = \frac{Sensitivity + Specificity}{2} \times 100$$

The Quality Value Q_V is defined as

$$Q_V = \frac{C}{(R_{fa} + 0.2) * (T_{dly} * P_{dct} + 6 * P_{msd})}$$

where C is the scaling constant,

R_{fa} is the number of false alarm per set,

T_{dly} is the average delay of the onset classification in seconds

P_{dct} is the percentage of perfect classification and

P_{msd} is the percentage of perfect risk level missed

The time delay is given as follows

$$\text{Time Delay} = \left[2 \times \frac{PC}{100} + 6 \times \frac{MC}{100} \right]$$

Table 4. Consolidated Average Values

Parameters	**Epoch 1**	**Epoch 2**	**Epoch 3**	**Average**
PC	94.37	94.58	94.37	94.44
MC	5.62	5.41	5.62	5.55
FA	0	0	0	0
PI	93.96	94.20	93.99	94.05
Sensitivity	100	100	100	100
Specificity	94.37	94.58	94.37	94.44
Time Delay	2.22	2.21	2.22	2.2
Quality Value	22.51	22.59	22.5	22.53
Accuracy	97.18	97.29	97.18	97.22

6. Conclusion

Thus the EEG signals not only represent the brain function but also the status of the whole body, i.e. a simple action as blinking the eyes introduces oscillation in the EEG records. Then, the EEG is a direct way to measure neural activities and it is important in the area of biomedical research to understand and develop new processing techniques. EEG signal pre-processing and post-processing methods could be considered as a “pattern recognition system” with focus on the classification algorithms. In this research the dimension of the EEG data was reduced using FMI and then the dimensionally reduced values are transmitted through a 2 x 2 DSTBC MIMO –OFDM System. Since the DSTBC MIMO OFDM has a high PAPR, PSO based ACE constellation algorithm was employed to reduce the PAPR and BER. It is observed that a PAPR value of about 5 dB is reduced. At the receiver side Weighted KNN Classifier is also placed to classify the epilepsy from EEG signals. The classification accuracy shows an average of about 97.22% with an average perfect classification rate of 94.44%. Future work may incorporate the possible usage of different dimensionality reduction techniques with various other types of classifiers for the perfect classification of epilepsy risk levels from EEG signals.

References

[1] Subasi, A. (2005b). Epileptic seizure detection using dynamic wavelet network. Expert Systems with Applications, 29, 343–355.

[2] Ogulata, S. N., Sahin, C., & Erol, R. (2009). Neural network-based computer aided diagnosis in classification of primary generalized epilepsy by EEG signals. The Journal of Medical Systems, 33, 107–112.

[3] Hazarika, N., Chen, J. Z., Tsoi, A. C., & Sergejew, A. (1997). Classification of EEG signals using the wavelet transform. Signal Processing, 59(1), 61–72.

[4] Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., & Elger, C. E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Physical Review E, 64(6), 061907.

[5] Adeli, H., Zhou, Z., & Dadmehr, N. (2003). Analysis of EEG records in an epileptic patient using wavelet transform. Journal of Neuroscience Methods, 123, 69–87.

[6] Alkan, A., Koklukaya, E., & Subasi, A. (2005). Automatic seizure detection in EEG using logistic regression and artificial neural network. Journal of Neuroscience Methods, 148, 167–176.

[7] Aslan, K., Bozdemir, H., Sahin, S., Ogulata, S. N., & Erol, R. (2008). A radial basis function neural network model for classification of epilepsy using EEG signals. The Journal of Medical Systems, 32, 403–408.

[8] Guler, I., & Ubeyli, E. D. (2005). Adaptive neuro-fuzzy inference system for classification. Journal of Neuroscience Methods, 148, 113–121.

[9] Kocyigit, Y., Alkan, A., & Erol, H. (2008). Classification of EEG recordings by using fast independent component analysis and artificial neural network. Journal of Medical Systems, 32, 17–20.

[10] Ocak, H. (2009). Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Systems with Applications, 36, 2027–2036.

[11] Orhan, U., Hekim, M., & Ozer, M. (2011). Epileptic Seizure Detection Using Probability Distribution Based On Equal Frequency Discretization. Journal of Medical Sytems. doi:10.1007/s10916-011-9689-y.

[12] Petrosian, A., Prokhorov, D., Homan, R., Dashei, R., & Wunsch, D. (2000). Recurrent neural network based prediction of epileptic seizures in intra and extracranial EEG. Neurocomputing, 30, 201–218.

[13] Pradhan, N., Sadasivan, P. K., & Arunodaya, G. R. (1996). Detection of seizure activity in EEG by an artificial neural network: A preliminary study. Computers and Biomedical Research, 29, 303–313.

[14] Srinivasan, V., Eswaran, C., & Sriraam, N. (2005). Artificial neural network based epileptic detection using time-domain and frequency-domain features. Journal of Medical Systems, 29(6), 647–660.

[15] Ubeyli, E. D. (2008). Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines. Computers in Biology and Medicine, 38(1), 14–22.

[16] Harikumar R, Sunil Kumar.P (2015), Fuzzy Mutual Information as a Dimensionality Reduction Technique for Epileptic Electroencephalography Signals , Research Journal of Applied Science, Engineering and Technology, Maxwell Scientific Publishers, Vol.10, No.9, 1035-1037.

[17] Lozano A et al (2005) , "High-SNR Power Offset in Multi-antenna Communication. *IEEE Tr.Comm*", pp.4134-4151

[18] Xiaodong Zhu et al (2012) , " Low-BER Clipping Scheme for PAPR Reduction in STBC MIMO-OFDM System" *Wireless Pers. Commun*, 65:335-346

[19] Tarokh V et al (2000) "Space-time block codes form orthogonal designs*", IEEE Trans. J.Sel.Areas Commun*., 18:7, 1169-1174.

[20] Ganesan G and Stoica P (2002) , "Differential modulation using space-time block codes", *IEEE Signal Process. Lett*., 9:2 , (2002) 57-60.

[21] Tanyawat Sanguanchue & Kietikul Jearanaitanakij (2012) 'Hybrid Algorithm for training feed-forward neural network using PSO-Information gain with back propagation algorithm', Ecticon 2012, 9th International conference 16-18th May Thailand, pp1-4

[22] Bassem et al (2007), "PAPR Reduction method for STBC MIMO-OFDM Systems using SLM in combination with sub band permutation", *Third International Conference on Wireless and Mobile Communications (ICWMC'07*) Guadeloupe.

[23] Krongold B and Jones D (2003), "PAR Reduction in OFDM via Active Constellation Extension", *IEEE Tr. Broadcasting*, pp.258-268.

[24] Vivencio D.P (2007), " Feature-weighted K-nearest neighbour classifier", Proceedings of the IEEE Symposium on Foundations of Computational Intelligence', pp: 481-486.