Int. J. Pharm. Sci. Rev. Res., 38(1), May – June 2016; Article No. 24, Pages: 144-148 ISSN 0976 – 044X

Research Article

Clinical Health Care for Long Distance using Matrix Factorization and Mahalanobis Based Sparse Representation Measures for Epilepsy Classification from EEG Signals

Harikumar Rajaguru[1], Sunil Kumar Prabhakar[2*]
[1]Department of ECE, Bannari Amman Institute of Technology, Sathyamangalam, India.
[2]Department of ECE, Bannari Amman Institute of Technology, Sathyamangalam, India.
***Corresponding author's E-mail:** sunilprabhakar22@gmail.com

Accepted on: 29-03-2016; **Finalized on:** 30-04-2016.

ABSTRACT

Around 1% percent of the human population in the world is affected by epilepsy. Because of the abnormal electrical patterns and discharges occurring in the brain, recurrent seizures occur. The mass activities of the neurons are recorded with the help of Electroencephalograph (EEG) placed on the scalp of the brain. EEG serves as a good companion to physicians in order to diagnose a lot of neurological disorders including epilepsy and stroke. For the diagnosis of epilepsy, detection of the epileptic seizures in the EEG signals forms a vital step. Earlier neurophysiologists used to inspect the EEG signals usually to differentiate between normal and abnormal EEG. With visual inspection, the detection of epilepsy was time consuming and costly too. As a result, a lot of approaches to detect the seizures automatically came into existence. Also the EEG recordings are lengthy and hence it is too difficult to process the entire data and so dimensionality reduction techniques are implemented to reduce the size of the EEG data. In this paper, the size of the entire EEG data is reduced with the help of Matrix Factorization (MF) Concept. The dimensionality reduced values are then transmitted through the Differential Space Time Block Coded Multiple-Input Multiple-Output System (DSTBC MIMO-OFDM). As a high Peak to Average Power Ratio (PAPR) is found for the DSTBC MIMO-OFDM System, Selective Mapping (SLM) Technique is incorporated to reduce the PAPR and Bit Error Rate (BER). At the receiver, a modified form of Mahalanobis Based Sparse Representation Classifier (M-SRC) is employed to classify the epilepsy from EEG signals. The bench mark parameters analyzed here are Specificity, Sensitivity, Time Delay, Quality Value, Performance Index, Accuracy, PAPR and BER.

Keywords: EEG, epilepsy, SRC, DSTBC MIMO-OFDM, EM, PAPR, BER.

INTRODUCTION

EEG involves the electrical signal recording and analysis which is generated by the brain[1]. For diagnosing neurological disorders related to epilepsy and to monitor it, EEG serves as a boon. Due to the excessive discharging from different groups of brain cells, recurrent and transient disturbances occur suddenly in the mental functions. From the scalp, the epileptic EEG is measured and is indicated with high amplitude and periodic waveforms which are usually synchronized[2]. The demand for highly skilled and patient doctors is very difficult to find nowadays. As EEG is a pretty complex and time taking operation, it usually requires visual screening by such doctors. This leads to prolonged diagnosis time and hike in medical expenditures and unnecessary delay in starting the treatment. Therefore automated techniques are given more importance to detect and classify the seizures[3]. A lot of related works has been proposed in literature. Acir detected the epileptic form events in EEG automatically with the help of three stage procedure based on Artificial Neural Networks[4]. Petrosian showed the complexity and recognition of different pre-ictal EEG patterns[5]. Exarchos utilized a data mining based approach for the EEG transient event detection and classification[6]. Iasemidis discussed the long term prospective on-line real time seizure prediction[7]. Niederhauser detected the seizure precursors from depth EEG by utilizing a sign periodogram transform[8]. Klaus Lehnertz showed the prediction of seizures by nonlinear EEG analysis[9]. The detection of singularity value of a character wave in an epileptic EEG signal by means of a wavelet was done by Chen[10]. The Recurrent Neural Networks employing Lyapunov exponent for the EEG signal classification was performed by Gulera[11]. The intelligent identification and classification of epileptic seizures by means of using wavelet transform was performed by means of Najumnissa and Shenbaga Devi[12]. The paper is organized as follows: In section 2, the materials and methods are illustrated and discussed. In section 3, the PAPR and BER analysis of DSTBC MIMO-OFDM System is written in detail. In section 4, the EM based SRC is discussed followed by the results and discussion in section 5. Finally the paper is concluded followed by the references.

MATERIALS AND METHODS

In the materials and methods section, the acquisition of the EEG signals is discussed initially. Then the need for dimensionality reduction is explained followed by the usage of Matrix Factorization as a dimensionality reduction technique.

Acquisition of EEG Data

For our experimental work, the raw EEG data was collected from 20 patients who are suffering from epilepsy at Sri Ramakrishna Hospital, Coimbatore. The data was collected in European Data Format (EDF). The recordings were done for nearly thirty minutes and the recorded signals were divided into epochs which have two second duration. The changes in the activity of the signal can be traced just with 2 second duration. For every patient, the total number of channels required for recording was 16 and it was done under 3 epochs. As the maximum frequency of the signal is around 50 Hz the sampling frequency is set as 200 Hz as the sampling frequency should be twice greater than the maximum frequency. There are about 400 values obtained for each and every epoch and each sample corresponds to the amplitude value of the signals. Totally for twenty patients around 25, 000 samples are present and hence it is difficult to process with this huge data set. So with the advent of dimensionality reduction techniques, it has become easier to reduce the dimensions of the entire EEG dataset. The block diagram is shown in Figure 1.

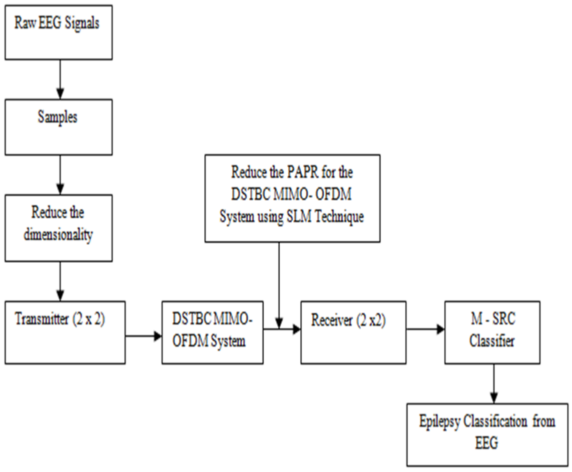

Figure 1: Block Diagram of the Paper

Matrix Factorization as a Dimensionality Reduction Technique

The dimensionality reduction is actually a pre-processing step required to reduce the dimensions of the data. In this paper, the usage of Variational Bayesian Matrix Factorization (VBMF) is utilized to reduce the dimensions of the data. The main goal of VBMF[13] is to approximate a particular unknown target $E(\in B^{D\times M})$ from its $'n'$ observations such that

$$A^n = \{A^{(i)} \in B^{D\times M}\}_{i=1}^{n}$$

If the assumption is made such that $D \le M$, a simple re-definition of the transpose E^T is made as E so that $D \le M$ holds. The most vital assumption of the matrix factorization is that E should always be a low-rank matrix. Assume that $C(\le D)$ be the respective rank of E. The matrix E can be decomposed as the product of $F \in B^{M\times C}$ and $G \in B^{D\times C}$ as follows

$$E = GF^T$$

Bayesian inference is very important in matrix factorization models which includes the case of conjugate Gaussian likelihood potentials.

DSTBC MIMO-OFDM System

A Multiple Input Multiple Output (MIMO) system has numerous antennas in both the transmitter and receiver side[14]. When MIMO is combined with OFDM, it paves the way for the next generation wireless systems and diversity gain is obtained and the signal fading is reduced[15]. With the advent of Space Time Block Codes, it can be incorporated with MIMO-OFDM System thereby becoming a STBC MIMO-OFDM System[16]. The channel impairments are generally known in STBC MIMO-OFDM system and it is not useful in fast moving environments and hence Differential STBC is of great use as there is no need to know about the channel impairments in order to decode the signal[17].

As the DSTBC MIMO OFDM System suffers from a high PAPR, Selective Mapping (SLM) Scheme[18] is incorporated here to reduce the PAPR and BER.

System Design and PAPR

A communication system with the Space Time Block Coding Technique with 2 transmit antennas and 2 receive antennas is considered for the analysis. From the transmitter side, the information blocks of symbols are passed to the next unit called DSTBC encoder, where each block embeds two symbols.

The space-time block encoder generates the code words of length $M = 2$, where M means the total number of antennas used in the transmitter side.

The OFDM Modulation unit and the RF frontends obtain the code words and then modulate the significant information onto the carrier frequency. On the receiver side, for the reception, upto N receiver antennas can be used efficiently. Using down-conversion unit, the RF signals are down-converted and digitized in the RF front-ends and then passed through the DSTBC decoder unit.

The received signals is interpreted with the help of STBC decoder and after the received signals are obtained it is generated as an estimate of the transmitted information symbols provided again as a block of two symbols simultaneously.

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

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

In MIMO-OFDM, since $N_T D$ instead of D time-domain samples are present and the CCDF of the PAPR is represented mathematically as follows[19]

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

PAPR Reduction Using Selective Mapping Technique

The steps to follow the SLM technique are as follows. The simulation parameters are used in Table 1. The figure 2 shows the typical block diagram where the PAPR can be minimized using Selective Mapping Scheme in the Space Time Block Encoded MIMO-OFDM System.

Table 1: Simulation Parameters

Modulation used	QPSK
MIMO System analyzed	2 x 2 MIMO-OFDM
Number of subcarriers	128
No of sub-blocks	4
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

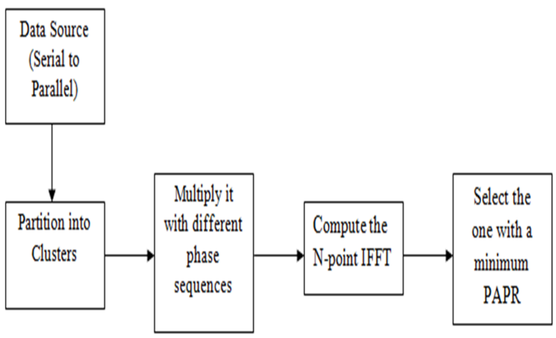

Figure 2: Block Diagram of Selective Mapping Technique

Here according to the Figure 2, the input data block say X=$[X[0], X[1],..., X[N-1]]$ is multiplied with B different phase sequences $B^u = [B_0^u, B_1^u, ..., B_{N-1}^u]^T$ where $B_v^u = e^{j\varphi_v^u}$ and $\varphi_v^u \in [0, 2\pi]$ for $v = 0,1,..., N-1$ and $u = 1,2,..., U$, which produce a modified data block $X^u = [X^u[1], X^u[2],..., X^u[N-1]]^T$. Inverse Fast Fourier Transform of U independent sequences $\{X^u[v]\}$ are considered here to produce the sequences $x^u = [x^u[0], x^u[1], ... x^u[N-1]]^T$, among which the one $\tilde{x} = x^u$ and the lowest value of PAPR is accepted for transmission purposes as shown in the following equation

$$\tilde{u} = \arg\min_{u=1,2,...,U}\left(\max_{n=0,1,...,N-1}\left|x^u[n]\right|\right)$$

For the successful implementation of SLM technique it requires U IFFT operations to be performed. For the receiver to recover the original data block, the pertinent information regarding the selected phase sequence B^u should be transmitted as side information.

Mahalanobis Based Sparse Representation Classifier (M-SRC)

Here in this method, as a way of true and original representation of a low dimensional sample, say, $b \in R^c$ from a training dictionary consisting of q samples, $\phi \in R^{c \times q}$ can be obtained by solving $\hat{b} = \phi e$, where $a \in R^q$ is the corresponding weight of each and every training exemplar in the entire dictionary ϕ, and $\hat{\bullet}$ denotes the estimate value. In practical conditions, in most of the cases, the system has either multiple solutions of no solutions. The main aim of sparse representations is to trace the smallest number of non zero coefficients e, such that $\hat{b} = \phi e$. The convex relaxation approaches[20] shows that under certain vital conditions, considering the sparsity of the representation, the minimal solution obtained in equivalent to the solution of the regression problem in statistics:

$$\hat{e} = \arg\min\|e\|_1 \text{ s.t. } \hat{b} = \phi e$$

where $\|e\|_1 = \Sigma|e|$. The main advantage of implementing l^1 minimization is that the problem can be solved easily with the help of convex optimization algorithms. Orthogonal Matching Pursuit[21] is a technique widely used to solve the l^1 minimization. Given the sparse representation coefficients $\hat{e}$ of a particular test signal using the dictionary ϕ, the class p^* of a query sample b is estimated using reconstruction error method. Given p classes, the reconstructed sample using sparse coefficients $\hat{e}$ from all the classes is compared to that of the reconstructed sample with the help of coefficients $\hat{e}_i$ from each and every respective class as

$$p^* = \min_{i=1:p}\left\|b - \phi\hat{e}^i\right\|_2$$

With the help of above equation, the test sample classification is assigned to the class with high similarity of the fully reconstructed signal. This necessitates the choice for choosing the degrees of freedom easily, thereby choosing the sparsest solution and producing graceful coefficients throughout the test samples. Once the sparsest solution is chosen, then the similarity measures is found out using the Mahalonobis distance[22]. For measurement of the similarity between the original data and its respective reconstructed data, Mahalonobis distance is employed instead of Euclidean distance. With the introduction of Mahalonobis distance, a generalized distance measure for signal classification is found out which embodies the different weights on various components of the feature vector. Mahalonobis distance has always been proved as a versatile similarity measure and so it is implemented with SRC to find out the similarity measures. For the given data points $v_1, v_2 \in R^q$, the Mahalonobis distance is given as $d_M(v_1, v_2) = \sqrt{(v_1 - v_2)^T M (v_1 - v_2)}$, where $M \in R^q$ is defined as a positive definite matrix. The distance between the original data and reconstructed data can be thus easily computed.

RESULTS AND DISCUSSION

The PAPR Results, BER Results and the Classification Results are discussed in this section.

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 Figure 3 and BER Analysis is shown in Table 2.

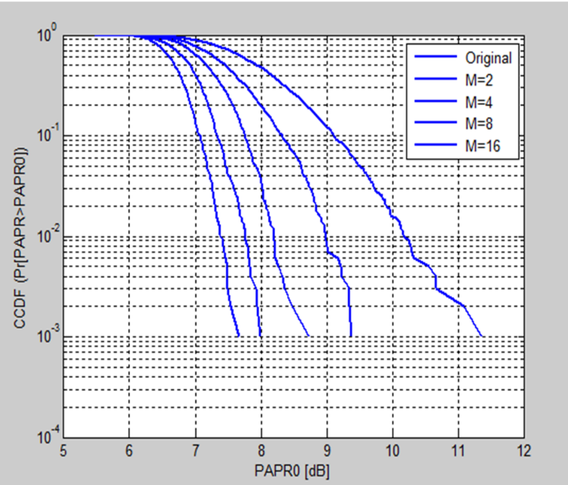

Figure 3: PAPR Reduction Analysis Using SLM Techniques

Classification Results

For VBMF as dimensionality reduction techniques Mahalonobis Based SRC as a Post Classifier at the receiver side, based on the Quality values, Time Delay and Accuracy the results are computed in Table 3 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 denotes the scaling constant,

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

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

P_{dct} gives the percentage of perfect classification

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

The time delay is expressed as follows

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

Table 2: BER Analysis

SNR	BER
0	0.1349
1	0.1109
2	0.0781
3	0.0588
4	0.0359
5	0.0212
6	0.0127
7	0.0077
8	0.0037
9	0.0017
10	0.0003
11	0.0002
12	0.0001
13	0
14	0
15	0
16	0
17	0
18	0

Table 3: Consolidated Average Values

Parameters	Epoch 1	Epoch 2	Epoch 3	Average
PC	95.41	95.62	95.625	95.55
MC	4.58	4.375	4.375	4.44
FA	0	0	0	0
PI	95.17	95.41	95.41	95.33
Sensitivity	100	100	100	100
Specificity	95.41	95.625	95.625	95.55
Time Delay	2.18	2.17	2.17	2.17
Quality Value	22.91	22.99	22.99	22.96
Accuracy	97.70	97.81	97.81	97.77

CONCLUSION

Thus epilepsy being one of the most prevalent neurological disorders, affects the daily life and duties of the patient to a great extent. Thus, in this paper, the dimensions of the EEG data were reduced initially with the help of Variational Bayesian Matrix Factorization concept.

As the work is implemented for telemedicine applications, the dimensionally reduced data set is given to the DSTBC MIMO-OFDM system. The PAPR and BER is reduced with the help of Selective Mapping Scheme. At the receiver, Mahalonobis based SRC is employed to classify the epilepsy from EEG signals. The results show that an average accuracy of about 97.77% is obtained; an average Perfect Classification of about 95.55% is obtained. A high quality value of about 22.96 is also obtained in this work. Thus this work can be widely used in telemedicine applications. Future works may incorporate many other machine learning techniques to classify the epilepsy from EEG signals.

REFERENCES

1. Subasi, A. (2005b). Epileptic seizure detection using dynamic wavelet network. Expert Systems with Applications, 29, 2005, 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, 2009, 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), 1997, 61–72.
4. N. Acir, I. Oztura, M. Kuntalp, B. Baklan, C. Guzelis, Automatic detection of epileptic form events in EEG by a three-stage procedure based on artificial neural networks, IEEE Trans. Biomed. Eng. 52(1), 2005, 30–40.
5. A. Petrosian, Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns, in: Proceedings of the IEEE Symposium on Computer-Based Medical Systems, 1995, 212–217.
6. T.P. Exarchos, A.T. Tzallas, D.I. Fotiadis, S. Konitsiotis, S.Giannopoulos, A data mining based approach for the EEG transient event detection and classification, in: Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems, 2005, 35–40.
7. L.D. Iasemidis, P.M. Pardalos, D.S. Shiau, W. Chaovalitwongse, K. Narayanan, A. Prasad, K. Tsakalis, R. Carney, J.C. Sackellares, Long term prospective on-line real-time seizure prediction, Clin. Neurophysiol. 116(3), 2005, 532–544.
8. J.J. Niederhauser, R. Esteller, J. Echauz, G. Vachtsevanos, B. Litt, Detection of seizure precursors from depth-EEG using a sign periodogram transform, IEEE Trans. Biomed. Eng. 51(4), 2003, 449–458.
9. Klaus Lehnertz, Florian Mormann, Thomas Kreuz, Ralph G. Andrzejak, Christoph Rieke, Peter David, And Christian E. Elger, "Seizure Prediction By Nonlinear EEG Analysis" in proceeding of *IEEE engineering in medicine and biology magazine*, vol.22, no. 1, January/February 2003, 57-63.
10. Chen H., Zhong S., and Yao D. "Detection singularity value of character wave in epileptic EEG by wavelet", in proceedings of *IEEE International Conference on Communications, Circuits and Systems and West Sino Expositions* 2002, vol. 2, 29 June-1 July 2002, 1094-1097.
11. N.F. Gulera, E.D. Ubeylib, I. Guler, "Recurrent neural networks employing Lyapunov exponents for EEG signals classification" in proceedings of *Journal on Expert Systems with Applications* vol. 29, no. 3, 2005, 506–514.
12. D. Najumnissa and S. Shenbaga Devi, "Intelligent identification and classification of epileptic seizures using wavelet transform", *International Journal of Biomedical Engineering and Technology*, Vol. 1, No. 3, 2008, 293-314.
13. M. Seeger, G. Bouchard, "Fast Variational Bayesian Inference for Non-conjugate Matrix Factorization Models", Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS) 2012, PG: 1012-1016.
14. Lozano A (2005), "High-SNR Power Offset in Multi-antenna Communication. *IEEE Tr. Comm*", 4134-4151.
15. Xiaodong Zhul (2012), "Low-BER Clipping Scheme for PAPR Reduction in STBC MIMO-OFDM System" *Wireless Pers. Commun*, 65, 2012, 335-346.
16. Tarokh V (2000) "Space-time block codes form orthogonal designs*", IEEE Trans. J.Sel.Areas Commun*., 18, 7, 2000, 1169-1174.
17. Ganesan G and Stoica P (2002), "Differential modulation using space-time block codes", *IEEE Signal Process. Lett*., 9, 2, 2002, 57-60.
18. Tao Jiang, Chunxing Ni, and Lili Guan, "A Novel Phase Offset SLM Scheme for PAPR Reduction in Alamouti MIMO-OFDM Systems Without Side Information, "IEEE Signal Processing Letters", Vol. 20, No 4, April 2013.
19. Seyran Khademi, "Constant Modulus Algorithm for Peak-to-Average Power Ratio (PAPR) Reduction in MIMO OFDM", IEEE Signal Processing Letters, Vol.20, No.5, May 2013.
20. J. Wright, A. Ganesh, S. Rao, Y. Peng and Y. Ma, "Robust Principal Component Analysis: Exact Recovery of corrupted low-rank matrices via convex optimization", in Advances in Neural Information Processing Systems, NewYork, USA: MIT Press, Dec 2009.
21. D. Cai, X. He, J. Han and H. J. Zhan. "Orthogonal Laplacian faces for face recognition", IEEE Trans. Image Process., vol,15, no.11, pg:3608-3614, Nov 2006.
22. S. Pourmohammed, R. Soosahabi, A.S. Maida, "An efficient character recognition scheme based on k-means clustering", 5th International Conference on Modelling, Simulation and Applied Optimization (ICMSAO), 2013, pg:1-6, Hammamet.

Source of Support: Nil, **Conflict of Interest:** None.