EEG Signal Discrimination using Non-linear Dynamics in the EMD Domain

S. M. Shafiul Alam, *Member, IEEE*, M. I. H. Bhuiyan, *Member, IEEE*, Aurangozeb, and Syed Tarek Shahriar

Abstract—An EMD-chaos based approach is proposed to discriminate EEG signals corresponding to healthy persons, and epileptic patients during seizure-free intervals and seizure attacks. An electroencephalogram (EEG) is first empirically decomposed to intrinsic mode functions (IMFs). The nonlinear dynamics of these IMFs are quantified in terms of the largest Lyapunov exponent (LLE) and correlation dimension (CD). This chaotic analysis in EMD domain is applied to a large group of EEG signals corresponding to healthy persons as well as epileptic patients (both with and without seizure attacks). It is shown that the values of the obtained LLE and CD exhibit features by which EEG for seizure attacks can be clearly distinguished from other EEG signals in the EMD domain. Thus, the proposed approach may aid researchers in developing effective techniques to predict seizure activities.

Index Terms—Electroencephalogram (EEG), empirical mode decomposition (EMD), largest Lyapunov exponent (LLE), correlation dimension (CD), epileptic seizures.

I. INTRODUCTION

An electroencephalogram (EEG) can be roughly defined as the mean electrical activity of the brain in different sites of the head. Basically, it is the sum of the extracellular current flows of a large group of neurons. EEG signal can be obtained using sensors on the scalp or by placing special electrodes intracranially. The application of intracranial electrodes involves a major surgery and their use is restricted to very particular cases, for example the study of epileptic patients who are candidates for a surgical resection.

Epilepsy is a group of disorders characterized by recurrent paroxysmal electrical discharges of the cerebral cortex that result in intermittent disturbances of brain function. For the majority of patients, seizures always occur suddenly and unexpectedly without any prior external precipitants. The unpredictable nature of seizures is responsible for the enhanced risk of sudden unexpected injury or death. Various methods are reported in the literature for discriminating EEG signals with or without seizure activity [1]–[7]. In [2], the modified Lempel-Ziv complexity is employed, but the analysis is carried out over only two EEG records, not enough to demonstrate its applicability on a wide scale. In [3], a combination of wavelet transform and chaos is used to classify healthy persons and epileptic patients with and

without seizure attacks. However, the classifications are not consistent for different sub-bands and satisfactory performance is achieved for only one sub-band. EEG signals are classified using a complex time-frequency-based analysis and neural network in [4]. In [5], the intrinsic mode functions (IMFs) of the EEG signals related to seizure free intervals and seizure activities are first obtained. Mean frequencies of the different IMFs are used to discriminate the signals. However, the analysis carried out in [5] does not include the signals corresponding to healthy persons. In [6] and [7], fractal dimension and higher order moments of EEG signals are used as classifying features, respectively. These features are employed in a support vector machine and neural network, respectively, for automatic detection of seizure activities.

In this paper, a nonlinear approach, based on chaos analysis of EEG signals in the empirical mode decomposition (EMD) domain, to classify healthy persons, and epileptic patients with and without seizure attacks, is introduced. The signals are discriminated using the values of the largest Lyapunov exponent (LLE) and correlation dimension (CD) calculated from the corresponding IMFs. A comprehensive analysis is carried out in the EMD domain using a very large set of EEG data and the results show that the values of LLE and CD can be an effective tool for discriminating EEG signals.

II. EMPIRICAL MODE DECOMPOSITION OF EEG SIGNALS

The EMD of a signal is based on the simple assumption that any data consists of different simple intrinsic modes of oscillations derived using the basis obtained from the data [8]. Each of the oscillatory modes is represented by an intrinsic mode function (IMF) that can have an amplitude and frequency as functions of time. For an N -point data, $X\{x_1, x_2, \dots, x_N\}$, the decomposition is carried out as follows:

1. Set input as h_0 . (At the beginning, $h_0 = X$.)
2. Set $h_{old} = h_0$.
3. The local maxima and minima of h_{old} are identified.
4. Envelops of local maxima e_{max} and that of local minima e_{min} are obtained using cubic spline interpolation.
5. The values of the mean of e_{max} and e_{min} are calculated as $m = (e_{max} + e_{min})/2$ and subsequently subtracted from h_{old} as

$$h_{new} = h_{old} - m \quad (1)$$

6. Set $h_{old} = h_{new}$.

The Steps 3-6 are repeated until the following stopping criterion is fulfilled.

Manuscript received April 12, 2012; revised May 15, 2012.

The Authors are with the Electrical and Electronic Engineering Department, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh (e-mail: imamul@eee.buet.ac.bd)

$$SD = \frac{\sum |h_{new} - h_{old}|^2}{\sum h_{old}^2} < \alpha \quad (2)$$

where α is an arbitrary value preset in the range of 0.2~0.3. Upon satisfying eqn. (2) h_{new} becomes an IMF c and a residue value, r is calculated as

$$r = h_0 - c \quad (3)$$

The entire process is then repeated by setting r as h_0 . Thus, the input signal can be decomposed into L IMFs until the residue becomes a monotonic function such that further extraction of an IMF is not possible. The input X can be reconstructed from all the IMFs as

$$X = \sum_{j=1}^L c_j + r_L \quad (4)$$

Fig. 1 shows a sample EEG signal along with its first three IMFs.

Fig. 1. An EEG signal and its first three IMFs.

III. EEG SIGNAL ANALYSIS WITH CHAOS

The first step in the chaos analysis is to obtain a phase space portrait, by using the time-dependent variables of the system (here, the EEG signal and its IMFs) as the components of the vectors, constituting the multidimensional phase space of the system. The sequential plots of the time-dependent vectors represent the evolution of the system's state over time. For an N -point data, $\{x_1, x_2, \dots, x_N\}$, the phase space portrait having M embedding vectors are reconstructed employing the method of delay coordinates [9]. First, a matrix X is formed as

$$X = [X_1 X_2 \dots X_M]^T \quad (5)$$

where $X_n = [x_n x_{n+l} \dots x_{n+(D-1)l}]^T$ represents the n -th embedding vector and $n = 1, 2, \dots, M$. For $n = M$, $N = M + (D-1)l$, with l and D denoting the *lag* or *reconstruction delay* and *embedding dimension*, respectively. The optimum values of the lag and embedding dimension are determined using the algorithm of equidistant histogram box and Cao's method, respectively [3].

The phase space portrait corresponds to the dynamics of an attractor, confined to a sub-region of the system. The

Lyapunov exponents and correlation dimension (CD) are used as a measure of the chaoticity and complexity of the attractors, respectively. The exponents quantify the average exponential separation between the nearby trajectories of a phase space. An exponential divergence coupled with folding of the trajectories represents deterministic randomness and unpredictability. Thus, a positive Lyapunov exponent for almost all initial conditions in a bounded dynamical system is used as an indicator of deterministic chaos.

Various methods are proposed to date for obtaining the Lyapunov exponents from a time series data. In the direct method proposed by Wolf et al. [10], the largest Lyapunov exponent (LLE) of a time-series data is obtained as

$$LLE = \frac{1}{n_N - n_0} \sum_{k=1}^N \log_2 \left(\frac{L'_k}{L_{k-1}} \right) \quad (6)$$

where n_0 , n_N , L_{k-1} , L'_k and N represent initial and final discrete time instant, the distance between two nearest neighbours on separate trajectories at n_{k-1} , distance at time n_k between those two neighbours as evolved from L_{k-1} and number of replacement steps, respectively. However, the LLE can also be estimated from the maximum slope of $p(i)$ vs. i diagram [11], where $p(i)$, the prediction error is given by

$$p(i) = \frac{1}{NT_s} \sum_{k=1}^N \log_2 \left(\frac{L_{k+i}}{L_k} \right) \quad (7)$$

where i and T_s represents the number of time steps and sampling interval respectively. In this paper, the method of [11] is used since it involves less computational complexity as compared to that of [10]. To obtain CD, the correlation sum $C(r)$ is first calculated using the Grassberger-Procaccia algorithm as [12]

$$C(r) = \frac{2}{(N-D+1)(N-D+1-w)} \sum_{j=D}^N \sum_{k=j-w}^j \Theta(r - |X_j - X_k|) \quad (8)$$

where, r , Θ and w represent cell size, Heaviside function and Theiler window, respectively. Next, the optimum value of CD is obtained employing the Takens-Theiler estimator as [13]

$$CD = \frac{C(r)}{\int_0^r \frac{C(r')}{r'} dr'} \quad (9)$$

IV. RESULTS AND DISCUSSIONS

The analysis of the EEG signals is carried out on a large set of data. Five data sets, each containing 100 single-channel EEG segments of 23.6-sec. duration, are used in the analysis. The data sets are obtained from <http://www.epileptologie-bonn.de/>. Set A and Set B consist of segments obtained from surface EEG recordings carried out on five healthy volunteers, in an awake and relaxed state with eyes open and closed, respectively. Segments in Set C are recorded from the hippocampal formation of the opposite hemisphere, whereas those of Set D from within the

epileptogenic zone of the brain, of epileptic patients during seizure free intervals. Set E contains signals corresponding to seizure attacks. The signals are recorded in digital format using a sampling rate of 173.61Hz omitting the artifacts caused either by muscle activity or strong eye movements [14]. Hence, the sample length of each segment from each set is $173.61 \times 23.6 \approx 4097$, and the corresponding maximum frequency content is 86.81 Hz. Since the frequency response of an EEG signal spans over 0~60 Hz [3], frequencies greater than 60 Hz are considered as noise and discarded by passing each signal through a 6th order Butterworth filter having a cut-off frequency of 60 Hz. Fig. 2 shows some band-limited EEG segments of five datasets. After band-limiting, EMD operation is carried out on the EEG signals of all sets. For Set C, each EEG signal is decomposed to nine IMFs and for others, to ten IMFs.

Using (7), the values of LLE are calculated for the band-limited EEG signals as well as for the corresponding IMFs. For each dataset, the mean and standard deviation (SD) of the values of LLE are then obtained and listed in Table I for different datasets. The values of CD are obtained using (8) and (9) and the corresponding mean and SD values are provided in Table II.

Fig. 2. Band-limited EEG segments from each data set.

TABLE I: MEAN LLE (SD SHOWN IN PARENTHESIS)

	Set A	Set B	Set C	Set D	Set E
Band-Limited	18.16 (2.098)	22.16 (1.93)	11.72 (2.643)	12.89 (3.762)	20.49 (3.634)
IMF 1	24.63 (0.763)	24.09 (0.792)	23.99 (0.898)	23.39 (1.028)	22.14 (2.244)
IMF 2	26.4 (1.253)	26.06 (1.315)	26.68 (1.407)	26.15 (1.914)	25.58 (2.47)
IMF 3	27.69 (1.337)	26.8 (1.422)	26.46 (1.406)	25.88 (2.057)	25.58 (2.863)
IMF 4	21.52 (2.065)	20.75 (2.137)	19.35 (1.903)	19.19 (2.29)	19.49 (3.461)
IMF 5	12.68 (1.901)	11.97 (2.41)	11.39 (2.204)	11.96 (2.487)	11.91 (3.448)
IMF 6	5.665 (2.253)	4.711 (1.999)	4.295 (1.957)	4.757 (2.051)	4.957 (2.551)
IMF 7	1.479 (1.152)	1.681 (1.569)	1.327 (1.176)	1.358 (1.444)	1.396 (1.399)
IMF 8	0.697 (0.667)	0.833 (0.894)	0.583 (0.809)	0.644 (0.74)	0.551 (0.858)
IMF 9	0.142 (0.385)	0.179 (0.495)	0.029 (0.118)	0.093 (0.248)	0.141 (0.32)
Residue	0.006 (0.019)	0.013 (0.031)	--	0.009 (0.031)	0.009 (0.025)

It is observed from Table I that the mean value of LLE shows no specific trend for the EEG signals from Set A to Set

E and thus, no decision can be taken from it. However, it decreases as one move from Set A to Set E for IMFs 1 and 3. A similar trend is found for the IMFs 2 and 4, with an exception for Sets C and D, and Set E, respectively. On the other hand, no specific trend is seen for the IMFs 5 to 9. But, for all IMFs it is seen that the mean value of LLE for Set A is always greater than that for Set E (seizure activity). From Table II it is observed that the mean value of CD decreases from Set A to Set E for the IMFs 1, 2, 3, 4 and 7, but no general trend for the EEG signals. As for other IMFs, no decision can be taken based on the mean value of CD since it does not show any specific trend.

Furthermore, a one-way analysis of variance (ANOVA) is performed at 99% confidence level, where no significant difference among the means of the five data sets represents a null hypothesis. The null hypothesis will be rejected if the corresponding *p* value is less than 0.001. Table III shows the results obtained from the ANOVA.

TABLE II: MEAN CD (SD SHOWN IN PARENTHESIS)

	Set A	Set B	Set C	Set D	Set E
Band-Limited	13.8 (26.6)	20.47 (101.9)	6.906 (1.572)	6.289 (1.475)	5.252 (1.793)
IMF 1	6.627 (1.356)	6.165 (0.833)	5.769 (0.798)	4.891 (1.264)	4.346 (1.219)
IMF 2	4.682 (0.407)	4.189 (0.478)	4.049 (0.395)	3.858 (0.611)	3.739 (0.829)
IMF 3	3.073 (0.221)	2.907 (0.317)	2.729 (0.231)	2.71 (0.319)	2.702 (0.488)
IMF 4	2.16 (0.122)	2.111 (0.14)	2.079 (0.112)	2.075 (0.126)	2.074 (0.244)
IMF 5	1.874 (0.146)	1.807 (0.206)	1.845 (0.142)	1.815 (0.196)	1.771 (0.239)
IMF 6	1.551 (0.221)	1.483 (0.241)	1.505 (0.222)	1.495 (0.248)	1.451 (0.262)
IMF 7	1.268 (0.182)	1.264 (0.166)	1.252 (0.148)	1.25 (0.159)	1.23 (0.165)
IMF 8	1.182 (0.085)	1.171 (0.088)	1.174 (0.085)	1.168 (0.094)	1.143 (0.148)
IMF 9	1.095 (0.136)	1.098 (0.07)	1.003 (0.282)	1.084 (0.13)	1.072 (0.204)
Residue	0.587 (0.524)	0.839 (0.422)	--	0.558 (0.528)	0.579 (0.526)

TABLE III: P-VALUES OBTAINED FROM ONE WAY ANOVA

Signals	<i>p</i> -value for LLE	<i>p</i> -value for CD
Band-Limited	0	0.1095
IMF 1	0	0
IMF 2	0.0002	0
IMF 3	2.3908e-14	0
IMF 4	1.59983e-13	0.0002
IMF 5	0.0113	0.0021
IMF 6	0.0003	0.0579
IMF 7	0.3593	0.5238
IMF 8	0.104	0.0812
IMF 9	0.02	0.0017

It is seen from Table III that the null hypothesis is rejected for both the LLE and CD in the case of the IMFs 1, 2, 3 and 4. Thus, the LLE and CD values of the five sets are significantly different from each other if their first four IMFs are considered. This is further illustrated in Fig. 3 that shows the confidence interval plots for the EEG signals and the corresponding first two IMFs. Fig. 3(a) shows plots for original band-limited EEG signals whereas Figs. 3(b) and 3(c) the plots for first and second IMFs, respectively. Notice that, the LLE values calculated directly from the signal shows discriminating feature but in an inconsistent manner, while

the values of CD are all similar, thus useless for discrimination. On the other hand, the values of CD and LLE obtained from the IMFs clearly show distinguishing feature-enabling one to easily discriminate signals belonging to various categories including seizure-free intervals and seizure attacks. Overall, it is clear that the values of LLE and CD can be quite effectively employed to identify seizure activities from the various IMFs of the EEG signals. In addition, one can take decisions from multiple IMFs, which is more convincing.

Fig. 3. Confidence interval (99% level of confidence) plots of LLE and CD values for (a) original EEG signals, and the corresponding (b) IMF1s and (c) IMF2s.

V. CONCLUSIONS

In this paper, EEG signals have been analysed using chaotic parameters in the EMD domain to differentiate the signals of healthy and epileptic patients. The largest Lyapunov exponent (LLE) and correlation dimension (CD) have been employed as the measure of chaoticity and complexity of a signal, respectively, and calculated for various IMFs of a large group of EEG signals of healthy and epileptic persons (with and without seizure attacks). It has been observed that both the value of LLE and CD decreases as the non-linearity as well as randomness of IMF decreases from healthy to signals corresponding to seizure activity. Thus, the parameters have demonstrated distinguishing features enabling one to discriminate the EEGs of seizure attacks from the others.

ACKNOWLEDGMENT

The authors are highly grateful to the department of Electrical & Electronic Engineering (EEE) of Bangladesh University of Engineering and Technology (BUET) for providing necessary laboratory facilities for carrying out the research.

REFERENCES

- [1] L. D. Iasemidis, J. C. Principe, and J. C. Sackellares, "Measurement and Quantification of Spatio-Temporal Dynamics of human Epileptic Seizure," *Nonlinear Biomedical Signal Processing*, IEEE Press, Piscataway, New Jersey, vol. 2, pp. 294-318, 2000.
- [2] W. Jia, N. Kong, F. Li, X. Gao, S. Gao, G. Zhang, Y. Wang, and F. Yang, "An epileptic seizure prediction algorithm based on second-order complexity measure," *Physiol. Meas.*, vol. 26, pp. 609-625, 2005.
- [3] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A Wavelet-Chaos Methodology for Analysis of EEGs and EEG Sub-bands to Detect Seizure and Epilepsy," *IEEE Trans. on Biomedical Engineering*, vol. 54, no. 2, pp. 205-211, Feb. 2007.
- [4] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic Seizure Detection in EEGs Using Time-Frequency Analysis," *IEEE Trans. on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 703-710, 2009.
- [5] R. B. Pachori, "Discrimination between Ictal and Seizure-Free EEG Signals Using Empirical Mode Decomposition," *Research Letters in Signal Processing*, Hindawi Publishing Corporation, vol. 2008.
- [6] M. Schneider, P. N. Mustaro, and C. A. M. Lima, "Automatic Recognition of Epileptic Seizure in EEG via Support Vector Machine and dimension fractal," in *Proc. of Intl. Joint Conf. on Neural Networks*, pp. 2841-2845, 2009.
- [7] V. M. Bedeuzzaman, O. Farooq, and Y. U. Khan, "Automatic Seizure Detection Using Higher Order Moments," in *Proc. of Intl. Conf. on Recent Trends in Information, Telecomm and Computing*, pp. 159-163, 2010.
- [8] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. -C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proc. of R. Soc. London, Ser. A*, vol. 454, pp. 903-995, 1998.
- [9] F. Takens, "Detecting strange attractors in turbulence", *Dynamical Systems and Turbulence, Lecture notes in Math.*, Springer-Verlag, Heidelberg, pp. 366-381, 1981.
- [10] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov Exponents from a time series," *Physical D*, vol. 16, pp. 285-317, 1985.
- [11] U. Parlitz, *Nonlinear Modelling-Advanced Black-box Techniques*, Kluwer Academic Publishers, Boston, pp. 209-239, 1998.
- [12] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Phys. Rev. Lett*, no. 50, pp. 346-349, 1983.
- [13] F. Takens, "Numerical determination of the dimension of an attractor", *Dynamical Systems and bifurcations, Lecture notes in Math.* Springer, Berlin, vol. 1125, pp. 99-106, 1985.

- [14] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E*, vol. 64, no. 6, pp. 1-8, 2001.

S. M. Shafiul Alam obtained his M. Sc. And B. Sc. Degrees in Electrical and Electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh in 2009 and 2011, respectively respectively. His major field of study is signal and image processing, biomedical signal processing and communication. Since 2009, he had been a Lecturer in the Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology. In 2011, he has joined in the same department as Assistant Professor. Cuurently, he is onleave pursuing Ph. D. in the University of Kansas, USA. Mr. Alam is a member of IEEE. He has published a number of papers in the area of biomedical signal processing.

M. I. H. Bhuiyan obtained his M. Sc. And B. Sc. Degrees in Electrical and Electronic Engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh in 1998 and 2001, respectively. He received his Ph. D. degree from the Department of Electrical and Computer Engineering, Concordia University, Canada in 2007. His major field of study is signal, image and video processing, biomedical signal processing and communication.

Currently, he is an Associate Professor in the Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology. Dr. Bhuiyan is Member of IEEE. He has received the best Paper Award from the Research Institute of Signal processing, Japan in 2003. he has publishes a number of papers on image denoising and compression and biomedical signal and image processing.