Analysis of Chaos in EEG Signals for Estimation of Drowsiness and Classification of Epilepsy Risk Levels

HARI VIKRAM. T. S
SREENITHI. P
R. HARIKUMAR
PROFESSOR AND HEAD
Department of Electronics and Communication Engineering
Bannari Amman Institute of Technology
Sathyamangalam, Erode District
INDIA
sharivikram@gmail.com, sreenithip@gmail.com

Abstract: - Chaos in nonlinear dynamical systems has become a widely-known phenomenon and its presence has been identified in many different systems in virtually all the fields of science. In medical world, analyzing chaos in the brain and explore its dynamics is a challenging task to every individual. In this paper, an effective and a practical method for exploring such brain activities are studied. This paper relates a method to analyze an Electroencephalogram (EEG) using Correlation Dimension (CD) for drowsiness estimation in sleep onset and epilepsy. Dimension is a critical property since, it indicates how many independent state variables are required to reproduce the system dynamics in state space and this in turn indicates how many state variables should be included in a mathematical model of the system. Aside from this practical issue, the dimension is an indicator of the degree of "complexity" of a system and tracking any changes in dimension due to pathology or other manipulations to the system is a useful diagnostic criterion. For many chaotic systems, accurate calculation of the CD from measured data is difficult because of very slow convergence as the scale size is reduced. This paper proposes a method for collecting data at large scales, creating the time series and determining the possibility of constructing an attractor for establishing the deterministic character of dynamics of the underlying system.

Keywords: - EEG Signals, Chaos, Sleep onset, Epilepsy, Correlation Dimension

1 Introduction
The EEG signals are highly subjective and the information about the various states may appear at random in the time scale [1]. The complexity of estimating the drowsiness and characterizing the EEG signal can be done using neuro-fuzzy models or through some computational techniques [2] [3]. Recent progress in the theory of nonlinear dynamical systems has provided new methods for the study of time series in fields such as hydrodynamics, chemistry, climatic variability, biochemistry and human brain activity [4] [5]. The study of such complex systems may be performed by analyzing experimental data recorded as a series of measurements in time of a pertinent and easily accessible variable of the system. In most cases, such variables describe a global or averaged property of the system. For example, a time series may be obtained by recording at regular time intervals the mean electrical activity of a portion of the mammalian brain. Although it may seem that such data offer only one dimensional view of activity of the brain, this is not the case: it can be shown that a
time series may provide information about a large number of pertinent variables, which may subsequently be used to explore and characterize the system's dynamics [6]. More specifically, by using a time series one can determine the possibility of constructing an attractor and thereby establishing the deterministic character of dynamic underlying system.

Such methods from the non linear dynamical theory can be dragged for better perception of EEG signals. The complexity of drowsiness estimation and characterizing the EEG signals can be brought under some chaotic optimization techniques. On careful application of these techniques may provide excellent results and can clearly exhibit the hidden dynamics inside the brain.

1.1 Chaos & Non-Linear Dynamics
The word chaos is a piece of jargon used particularly to describe a complex type of behavior. It is the term used to describe the apparently complex behavior of what considered being simple, well-behaved systems [7]. Chaotic behavior, when looked at casually, looks erratic and almost random. The behavior of a system strongly influenced by outside, random "noise" or the complicated behavior of a system with many and many degrees of freedom, each "doing its own thing" [8]. Chaotic system are almost free of noise, these systems are essentially deterministic; that is, precise knowledge of the conditions of the system at one time allow us, at least in principle, to predict exactly the future behavior of that system. The problem of understanding chaos is to reconcile these apparently conflicting notions: randomness and determinism.

The study of chaos has provided new conceptual and theoretical tools enabling us to categorize and understand complex behavior that had confounded previous theories. Chaotic behavior seems to be universal; it sums up in mechanical oscillators, electrical circuits, LASERS, nonlinear optical systems, chemical reactions, nerve cells and also in EEG signals [9]. Chaos is really only one type of behavior exhibited by nonlinear systems. The field of study is more properly called nonlinear dynamics, the study of the dynamical behavior of a nonlinear system. A nonlinear system is a system whose time evolution equations are nonlinear; that is, the dynamical variables describing the properties of the system (for example position, velocity, acceleration, pressure, etc.) appear in the equations in a nonlinear form.

2. Materials and Methods

2.1 EEG Signals and Data Base
The Electroencephalogram (EEG) is a recording of the electrical potentials generated by the brain. Analyzing such complex signals provides a better way for medical diagnostics, drowsiness detection, and schizophrenia [10]. Frequency of EEG signals is 50 Hz and its amplitude varies between 10-100 micro volts. Its maximum amplitude is about 50-60 micro volts. The analog signals from the brain must be translated into digital signals for processing and storing the data.

Typically, sixteen channels of data are recorded by measuring the potential difference between pairs of electrodes placed on the scalp. EEG patterns have shown to be modified by a wide range of variables including biochemical, metabolic, circulatory, hormonal, neuroelectric and behavioral factors. In the past Encephalographer, by visual inspection was able to qualitatively distinguish normal EEG activity from localized or generalized abnormalities contained within relatively long EEG records. The different types of epileptic seizures are characterized by different EEG waveform patterns [11].

Sleep EEG database used in the study is encoded in European data format were obtained from Caucasian males and females (21 - 35 years old) without any medication; they contain FpzCz and PzOz EEG signals, each sampled at 100 Hz. The recordings also contain the sub mental-EMG
envelope, oro-nasal airflow, rectal body temperature and an event marker, all sampled at 1 Hz. The recordings contain sub mental EMG sampled at 100 Hz and an event marker sampled at 1 Hz.

The epilepsy EEG data used in the study were acquired from eight epileptic patients who had been under the evaluation and treatment in the Neurology department of Sri Ramakrishna Hospital, Coimbatore, India. A paper record of 16 channel EEG data is acquired from a clinical EEG monitoring system through 10-20 international electrode placing method. The EEG signal was band pass filtered between 0.5 Hz and 50Hz using five pole analog Butterworth filters to remove the artifacts. With an EEG signal free of artifacts, a reasonably accurate detection of epilepsy is possible; however, difficulties arise with artifacts. This problem increases the number of false detection that commonly plagues all classification systems. With the help of Neurologist artifact free EEG records with distinct features were selected. These records were scanned by Umax 6696 scanner with a resolution of 600dpi.

3. State Space Reconstruction

3.1 State Variables and State Space

State variables are those that change over time and reflect the behavior of a dynamic system. This is a mathematical construction, in which each state variable is plotted along one of the axes no matter what the initial state, some system will always come to rest at (0,0); this point is called an attractor, since it attracts all trajectories. In general an attractor can be much more complex than this simple point and many research papers [12],[13] speak loosely of "state-space attractors" when referring to sets of trajectories whether or not they have been demonstrated to be true attractors. In order to conform that a physical system contains an attractor is to perturb the system and see if the trajectories return to some subset of the state space. Another feature of this simple example is that the trajectory does not cross itself. In fact it is a general feature of the state space for a deterministic system. If the trajectory crosses itself, and if the state of the system at a given moment is at the point of intersection, then it cannot be determined which path the trajectory will follow. This contradicts the concept of determinism. Therefore, if trajectories appear to cross, the system is either random, or the dimension of the state space is not high enough to depict the trajectories accurately [12].

3.2 Time Delay Reconstruction

Time-delay reconstruction is almost very simple and yet extremely powerful. Instead of using the actual state variables, such as x(t) and its derivatives, the successively delayed values of x(t) can be used. The plotted trajectories spiral in to a fixed point at the origin, and they do so without any self-crossings. The exact shapes of the plots are slightly different, depending on the value chosen for the time delay parameter in the reconstruction [13].

If the original data points in the time series are represented by x(i), then the reconstructed attractor consists of the M-dimensional points y(i), generated from the time series as follows:

\[
y(1) = [x(1), x(1+L), \ldots, x(1+(M-1)L)] \quad (1) \\
y(2) = [x(1+J), x(1+J+L), \ldots, x(1+(M-1)L)+L] \quad (2) \\
y(N) = [x(1+(N-1)L), x(1+(N-1)L)+L, \ldots, x(1+(N-1)L)+(M-1)L)] \quad (3)
\]

Here N is the number of points on the reconstructed trajectory or attractor. M is the embedding dimension, each trajectory point y(i) is composed of M values from the time series x(i), separated by time delay L. L is the interval between first elements of successive attractor points, and is usually set to 1.

Carrying out a proper time-delay embedding or reconstruction, with real data can be thorny. The key parameters in the process are the dimension of the embedding space (the embedding dimension M) and the
time delay \( L \). In other words, how many state variables are there, and how far apart in time should be the delayed elements of each point in the state space. The key idea in the choice of time delay \( L \) is that the elements that make up an attractor point \( y(i) \) has to be close enough in time that they loosely approximate a derivative and are dynamically related, yet far enough apart in time that they are not repetitive. Each point \( y(i) \) should capture some dynamic information about the system, and if the elements \( x(i) \) of that point are too close together, the information they provide will be redundant \([14]\).

### 3.3 Appropriate Selection of Time Delay

One of the simplest and yet more effective methods for choosing \( L \) is that it should be a small multiple (2 or 3) of the correlation time of the signal \( x(t) \). The correlation time is the time shift at which the autocorrelation function \( R_{xx}(\tau) \) of the time series \( x(t) \) has decayed to \( 1/e \) of its peak value. This is one way to quantify the notion that the consecutive \( x(i) \) values should be far enough apart in time to be somewhat but not completely independent (uncorrelated). This simple rule of thumb is a good starting point for the selection of \( L \).

### 3.4 Selection of Embedding Dimension

Selection of embedding dimension \( M \) is the other major issue. The embedding dimension \( M \) should be large enough that the attractor is properly embedded in the topological sense. In particular, there should be no trajectory crossings if the system is truly deterministic (although noise of various types can introduce apparent intersections which can often be safely ignored). A promising approach to both of these questions, which enjoys widespread use, is that of False Nearest Neighbors (FNN)\([14]\). If an attractor is reconstructed in an embedding space with too small a dimension \( M \), then points on the attractor that are actually far apart in space will appear artificially close together - the trajectories are compressed because the embedding space is not big enough for them to fully expand. These points that appear close together in \( M \) dimensions but are actually far apart in a higher-dimensional space are false neighbors and FNN quantifies this concept \([13]\). The distance between two points \( y(i) \) and \( y(j) \) in \( M \)-dimensional can be defined as \([15]\)

\[
D_M(i,j) = \sqrt{\sum_{k=1}^{M} [y_k(i) - y_k(j)]^2}
\]

Here, \( D_M \) denotes the distance as measured in \( M \) dimensions, that is, with \( M \) delayed elements in each point \( y(i) \) and \( y(j) \). The subscript \( k \) to the right indicates that corresponding delayed elements are subtracted from each other in the distance calculation. This is nothing more than the well-known Euclidean distance measure, extended to \( M \) dimensions. A point \( y(j) \) is a false nearest neighbor of \( y(i) \) if the distance between the two in \( M+1 \) dimensions is much greater than the distance in \( M \) dimensions \([15]\)

\[
\frac{D_{M-1}(i,j)}{D_M(i,j)} > R_{thr}
\]

Here, \( R_{thr} \) is a distance-ratio threshold. If the distance increases by more than this factor, then these points are false neighbors. A value of approximately 10 for \( R_{thr} \) is suitable in many cases \([15]\). In operation, an initial value for embedding dimension \( M \) is set. Then, each point on the attractor is taken in turn as a reference point. The nearest neighbor to each reference point is found by computing the distance in \( M \)-dimensional space between the reference and every other point and identifying the minimum distance. Then, the distance between these same two points found in \( M+1 \) dimensions. If the ratio of these two distances is greater than \( R_{thr} \), the points are false nearest neighbors. Across all reference points, the proportion of nearest neighbors that are false nearest neighbors is found, for
the given dimension M. Then M is increased by one, and the process repeated.

### 3.5 Dimension and Box Counting Dimension

Dimension is a critical property because it indicates how many independent state variables are required to reproduce the system dynamics in state space, and this in turn indicates how many state variables should be included in a mathematical model of the system. Since brain signals are purely chaotic, those chaotic dynamics are termed as strange attractors. They occupy a well-defined and bounded region of the state space. Yet the system behavior is aperiodic, so no matter how much data acquire the attractor trajectory will never return to the same location in state space. One way to accomplish this is if the attractor forms a fractal, such that there is finer and finer detail as look at it more and more closely; in this sense, no matter how "dense" the trajectory in any given area of the state space, there is always room to squeeze in another trajectory passage. Scaling process provides us a purposeful way to determine the dimension of an attractor in state space, where dimension is defined as the exponent. The box-counting dimension implements the idea of power-law scaling in a more general form. An object in an M-dimensional space, count the number N of M-dimensional boxes, each with side of length \( \varepsilon \), that are needed to cover the object. If N increases as a power law function of \( \varepsilon \), then we can define the dimension D as [16]

\[
D = \lim_{\varepsilon \to 0} \frac{\log \left( \frac{N}{\varepsilon^D} \right)}{\log \left( \frac{1}{\varepsilon} \right)} \tag{9}
\]

In particular, as the box size gets smaller, fewer points are enclosed in each box, on average, yet each box is included in N no matter how many points it contains. Thus information regarding the probability of the attractor visiting certain boxes is lost. For these conceptual and practical reasons, box-counting dimension has been almost completely surpassed in most applications by the correlation dimension. Box-counting dimension is one of a series of dimensions based on a more general form. The basic form of Renyi dimensions is [6]

\[
D_q = \frac{1}{q(q-1)} \lim_{\varepsilon \to 0} \frac{\log I(q, \varepsilon)}{\log \varepsilon} \tag{10}
\]

Here, q indicates which in the series of dimensions is being considered (q=0 is the box-counting dimension), \( \varepsilon \) is the size of a box as before, and

\[
I(q, \varepsilon) = \sum_{i=1}^{M} \left[ \mu(C_i) \right]^q \tag{11}
\]

\[
\mu(C_i) = \lim_{t \to \infty} \frac{\eta(C_i, T)}{T} \tag{12}
\]

The quantity \( \eta(C_i, T) \) is the amount of time that a trajectory spends in box \( C_i \) in the time span from 0 to \( T \). Hence, \( \mu(C_i) \) is the proportion of time that the trajectory spends in box \( C_i \), (in the long run, as \( T \) increases), and this is essentially the probability that the attractor trajectory passes through box \( C_i \). If \( q=0 \), \( \mu(C_i) \) is raised to the power zero in the equation for \( I(q, \varepsilon) \), so that \( \mu(C_i)^q \) is zero if \( C_i \) is not visited at all, and one if it is visited by the trajectory no matter how briefly. In other words, I(q, \( \varepsilon \)) is a count of the number of boxes \( C_i \) visited by the trajectory. Incrementing q to 1, the next in the series of dimensions, D1, is known as the information dimension and it is not used in process. The next dimension, when q=2, is of by far the most interest called the
correlation dimension and the equation 10 reduces to

$$D_2 = \lim_{\xi \to 0} \frac{\log \left[ \sum_{i} \mu^2(C_i) \right]}{\log (\xi)} \quad (13)$$

3.6 Correlation Integral

Grassberger and Procaccia [16] showed that the summation in equation (13) could be approximated by a correlation integral which is much easier to compute from experimental data:

$$\sum_{i,j} \mu^2(C_i) \approx \frac{1}{N^2} \sum_{i} \sum_{j \neq i} U(\xi, |y_i - y_j|) \quad (14)$$

$$U(\xi, |y_i - y_j|) = \begin{cases} 
1 & \text{if } |y_i - y_j| < \xi \\
0 & \text{otherwise} 
\end{cases} \quad (15)$$

Although expressed as a discrete summation, the quantity on the right in the upper equation is known as a correlation integral. The operator $U(.)$ is a step function; as expressed here, it is one if the distance between the attractor points $y(i)$ and $y(j)$ are within distance $\varepsilon$ of each other, and zero otherwise. Thus, the correlation integral counts the number of pairs of points on the entire attractor that are within distance $\varepsilon$ of each other, and divides this by $N^2$, the total number of pairs of points.

The demonstration of this equality can be found in Grassberger and Procaccia [16], but an intuitive argument can be made to justify it. If, at a given box size $\varepsilon$, the attractor visits box $C_i$ for $P$ points out of a total number $N$ of points on the attractor, then $\mu^2(C_i) = (P/N)^2$. On the other hand, in box $C_i$, since there are $P$ points there will be approximately $P^2$ pairs of points - that is, box $C_i$ will contain $P^2$ pairs of points within distance $\varepsilon$ of each other. By the definition of the function $U$, this means that $C_i$ will contribute an amount $P^2$ to the correlation integral. Since this is divided by the total number of point pairs $N^2$, this contribution ($(P/N)^2$) is identical to that of the contribution of $C_i$ to the summation of $\mu^2(C_i)$ and the two quantities are equal. The equality is an approximation, largely due to the fact that the correlation integral is expressed in terms of inter-point distances.

The correlation dimension has become a standard measure of the fractal dimension of attractors that have been reconstructed in the state space. It approximates and is a lower bound for, the box-counting dimension (i.e., it is less than or equal to the box counting dimension, with equality in the case when all the boxes $C_i$ are occupied equally). Its use is simple in principle, but nontrivial in practice.

3.7 Correlation Dimension

The correlation integral obtained in the equation (14) can be used to approximate the correlation dimension.

$$D_2 = \lim_{r \to 0} \frac{\log \left[ C(r) \right]}{\log (r)} \quad (16)$$

$$C(r) = \frac{1}{N(N-1)} \sum_{i} \sum_{j \neq i} U(r, |y_i - y_j|) \quad (i \neq j) \quad (17)$$

The notation has been changed to use $r$ (radius) rather than $\varepsilon$ to designate the criterion distance; when two points $y_i$ and $y_j$ are closer together than $r$, they are "spatially correlated" and contribute to the correlation integral $C(r)$. The divisor has also been changed to reflect the fact that, since the case $i=j$ is always skipped in the summation (since the distance between $y_i$ and $y_j$ is zero when $i=j$ and counting this does not accurately reflect how close different points are to each other), the total number of inter-point pairs being compared is $N(N-1)$ rather than $N^2$. If $C(r)$ increases as a power-law function of $r$, then $C(r)$ versus $r$ on a log-log plot should be a straight line, and the slope will be the correlation dimension $D_2$. 
4 Estimation of Attractor Dimension from Sleep Subjects and Epileptic Patients

As EEG signals are Chaotic, on reconstruction the time series will create an attractor, estimating the attractor dimension directly exhibits the number of state variables controlling the system. The usefulness of this parameter can be extended to classify the risk level of epileptic patients. Patients epochs were clubbed together and the resulted time series were involved in the classification process. The figure 1 represents the autocorrelation plot of patient 1, through the autocorrelation plot the appropriate time delay can be estimated and the value of embedding dimension can be found through false nearest neighborhood method. From the above two parameters an attractor can be designed through the time delay reconstruction process. The figure 2 represents such an attractor constructed from the patient -1 EEG signal. The next step is to obtain the correlation integral value for each value of radius which is to be set manually. Grassberger Procaccia algorithm can be utilized for estimating the suitable dimension value for each value of radius. The figure 3 depicts the logCr vs logr plot for the patient 1 EEG signal. The slope of the logCr vs logr plot will directly provides the dimension value of the reconstructed attractor. Since the curve obtained through the logCr vs logr plot is not linear, approximation can be done through linear regression algorithm. The figure 4 represents the linearly approximated plot calculated through linear regression algorithm. Subjects having different level of risk can be classified through the estimated dimension value. Similar approach can be applied to the normal EEG signals acquired during sleep; the figure shows the logCr vs. logr plot and the fitted curve to easily estimate the correlation dimension value from the slope of the fitted straight line.
Figure 4: Linear Regression for Correlation Dimension

The slope of the logCr vs logr plot will directly provide the dimension value of the reconstructed attractor. Since the curve obtained through the logCr vs logr plot is not linear, approximation can be done through linear regression algorithm. The figure 4 represents the linearly approximated plot calculated through linear regression algorithm. Lot of subjects having different levels of risk can be classified through the estimated dimension value. Similar approach can be applied to the normal EEG signals acquired during sleep; the figure shows the logCr vs. logr plot and the fitted curve to easily estimate the correlation dimension value from the slope of the fitted straight line.

5 Results And Discussion

Three subjects were involved in the sleep stage classification process, each window consists of hundred dataset and nine thousand data were used for the reconstruction process. The table 1 shows the correlation dimension value for several windows. In the epilepsy risk level detection process eight patients having different risk levels were classified. The table 2 shows the correlation dimension value of eight different patients and also their mean and standard deviation. The table 3 makes the comparison between the normal subjects during sleep and the epileptic patients; it is evident that the epileptic patients have low dimension value when compared to the normal subjects during sleep.

<table>
<thead>
<tr>
<th>Window Size</th>
<th>SUBJECT 1</th>
<th>SUBJECT 2</th>
<th>SUBJECT 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-90</td>
<td>6.3074</td>
<td>6.3806</td>
<td>6.2451</td>
</tr>
<tr>
<td>100-190</td>
<td>5.9902</td>
<td>6.7092</td>
<td>6.5435</td>
</tr>
<tr>
<td>700-790</td>
<td>6.5141</td>
<td>6.8207</td>
<td>6.0563</td>
</tr>
<tr>
<td>1200-1290</td>
<td>6.0981</td>
<td>6.6723</td>
<td>5.8714</td>
</tr>
<tr>
<td>7000-7090</td>
<td>6.6138</td>
<td>6.5634</td>
<td>6.2673</td>
</tr>
<tr>
<td>12000-12090</td>
<td>6.054</td>
<td>6.1254</td>
<td>6.3983</td>
</tr>
<tr>
<td>Mean</td>
<td>6.2629</td>
<td>6.5452</td>
<td>6.2303</td>
</tr>
</tbody>
</table>

Table 1: Dimension Values for Sleep Subjects

<table>
<thead>
<tr>
<th>PATIENT</th>
<th>CORRELATION DIMENSION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPOCH-1</td>
</tr>
<tr>
<td>Patient-1</td>
<td>2.0506</td>
</tr>
<tr>
<td>Patient-2</td>
<td>4.7539</td>
</tr>
<tr>
<td>Patient-3</td>
<td>2.4017</td>
</tr>
<tr>
<td>Patient-4</td>
<td>3.5604</td>
</tr>
<tr>
<td>Patient-5</td>
<td>3.0379</td>
</tr>
<tr>
<td>Patient-7</td>
<td>5.2193</td>
</tr>
<tr>
<td>Patient-8</td>
<td>5.9284</td>
</tr>
</tbody>
</table>

Table 2: Representation of Risk level Classifications


6 Conclusion

Since non linear dynamics have the ability for the brilliant classification of variety of EEG signals, we have utilized the chaotic optimization technique to the full extent for better characterization of EEG signals. Many researches on the characterization of brain dynamics rely only on the Machine learning or Neural Networks or Neuro-fuzzy models which cannot not be stretched to the desired extent. Using only the correlation dimension algorithm, the dynamics of brain at different states are calculated. This closely relates how many independent variables needed to control a system at a time.

Low value of dimension indicates that the system is somewhat affected from its normal working phenomenon. Normal subjects while sleeping provides a dimension value of approximately 6 indicate normal working activity of brain. We have utilized this concept for the classification of risk levels of Epilepsy patients which closely coincides with the results estimated by other methods. Finally the drowsiness of the patient can be clearly distinguished and correlation dimension method is found to be better when compared to the other methods.

### References:


Computer Technique”, Nonlinear Dynamics, Psychology, and Life Sciences, Volume 03, Number 1, January 1999 , pp. 49-63(15).


