Phase space analysis of Event Related Potential during episodic memory retrieval

M. Ghoshuni\textsuperscript{a,b}, A.M Nasrabadi\textsuperscript{b,c}, M.A Khalilzadeh\textsuperscript{a}, and M.R. Hashemi\textsuperscript{b,d}

\textsuperscript{a}Department of Biomedical Engineering, Islamic Azad University, Mashhad branch, Mashhad, Iran
\textsuperscript{b}Department of Biomedical Engineering, Islamic Azad University, Science and Research branch, Tehran, Iran
\textsuperscript{c}Department of Biomedical Engineering, Shahed University, Tehran, Iran
\textsuperscript{d}Department of Biomedical Engineering, AmirKabir University of Technology, Tehran, Iran

Abstract— In this study we have used recurrence analysis (RA) to classify ERPs that appear from episodic memory retrieval in old/new recognition task. Since RA is based on embedding phase space, we have used correlation dimension and autocorrelation function in order to estimate embedding dimension and the lag time between successive components of each of embedding space vectors alternatively.

According to RA the rates of classification have been improved in comparison to previous study. We have obtained 98.9\% accuracy for train data and 97.7\% for test data. Furthermore we could classify ERPs with line interference noise using RA.

I. INTRODUCTION

An event-related potential (ERP) is a change in the electrical state of the brain that occurs in response to a discrete sensory or cognitive event \cite{1}. Episodic memory is one of the cognitive events that has been investigated in this study. By definition, episodic memory is the explicit recollection of incidents that occurred at a particular time and place in one’s personal past, and is the subsystem of the explicit (declarative) memory \cite{2}.

Many researches have been done on ERP studies of episodic memory \cite{3}-\cite{7}. A common finding of many ERP studies of episodic memory is that presentation of old/repeated items elicits more positive-going ERPs than does presentation of new/unrepeated items \cite{5}-\cite{7}. Such ERP ‘old/new effects’ typically onset approximately 300–400 ms post-stimulus, last 300–600 ms and, are generally of greatest magnitude at left parietal and adjacent centroparietal electrodes \cite{8}. But most of these studies have been focused on intuitive or statistical analysis of ERPs that appear in episodic memory retrieval and it has not been accomplished any analysis from engineering aspects.

Furthermore an ERP can be extracted by averaging away the portion of the signal that is not time-locked to the stimulus onset \cite{8}. It has long been recognized, however, that the variability rejected by the averaging process might itself be physiologically significant \cite{9}. The presence of nonlinear determinism in the electroencephalogram (EEG) has been studied using Lyapunov exponents and fractal dimension \cite{10}, but neither method has been shown to be useful for analyzing ERPs, perhaps because these methods require a stationary signal, which is a condition often not realized in practice.

In the absence of a priori knowledge regarding how an ERP is generated, the optimal detection procedure is one that makes minimal assumptions regarding the dynamical nature or statistical properties of the recorded signal, but yet affords a requisite sensitivity\cite{11}. We have used recurrence analysis (RA), as a method that meets these conditions and appears to be particularly useful for analyzing nonlinear ERPs. Furthermore we have presented methods to finding some parameters of RA such as embedding dimension and lag time, but other parameters of RA have been chosen empirically. Whereas in the same pervious study\cite{11} all parameters have been chosen empirically. Likewise we have shown that using RA, we can classify ERPs with line interference noise.

II. MATERIAL AND METHODS

A. EEG Measurement

EEGs were recorded from Fz, Cz, and Pz referenced to linked ears (International 10–20 system) using Ag/AgCl electrodes attached to the scalp with conductive paste; The signals were amplified and filtered to pass 0.13–42 Hz, sampled at 256Hz and stored on a computer hard-drive. Electrooculogram (EOG) was recorded from sub- and supraorbital electrodes (above and below the right eye). Digitized data were subsequently analyzed offline using MATLAB software. Prior to data analysis, all data were digitally filtered in the 0.5–35 Hz range and trials were discarded from analyses if they contained eye movements.
(EOG over 100 µV), or one of channels were bad (transit amplitude over 50 µV).

Twenty-seven subjects (27 male) participated in the study. They were generally undergraduate or postgraduate students and all had normal or corrected vision.

Visual stimuli were presented to each subject. All stimuli were complex random shapes. They were designed to be neither figurative nor verbalizable. Another details and the procedure of experiment has been described in pervious study [12].

B. Recurrence Analysis

Recurrence analysis was developed by Webber and Zbilut to detect deterministic behavior in time series data, such as the EEG [13]. The deterministic behavior may be linear or nonlinear; RA imposes no constraints on the stationarity or statistical characteristics of the time series [11].

Use of RA to detect ERPs involves phase space embedding of successive intervals of the EEG signal, calculation of the corresponding recurrence plots, and quantification of the plots using an appropriate nonlinear quantifier [11].

The mathematical details of RA have been described elsewhere [13], [14], [15]. After choosing an embedding dimension \( d \) and a lag time \( t_l \), the brain’s electrical activity is represented by a series of \( d \)-dimensional vectors, the sequence of which corresponds to a trajectory in the phase space [11]. The trajectory is represented in two dimensions by a recurrence plot [14], which can be quantified using any of nonlinear quantifiers [13]; the quantifier used here is percent recurrence \( %R \), defined as the ratio of the number of recurrent points to the total number of points in the recurrence matrix [14]. Points in phase space are said to be recurrent if the distance between them in phase space is less than an adjustable parameter [11] (here, chosen to be 15% of the maximum distance). For calculating the distances, we used the Euclidean norm [16].

A phase space can be constructed for an entire epoch of the EEG, leading to a single value of \( %R \) [11]. However, to detect the transient changes in the EEG produced by the ERPs, it was necessary to iterate the calculation, using a sliding window of points in \( V(t) \) to produce a corresponding time series, \( %R(t) \); this process captured the dynamic activity (both linear and nonlinear) in the EEG occurring over small time intervals [11].

The changes induced in the EEG by the stimulus are more easily detected by analyzing \( %R(t) \), a smoothed version of the \( %R \) produced by use of a sliding averaging window. To synchronize \( V(t) \) and \( %R(t) \), we adopted the convention that each point in \( %R(t) \) was plotted in the middle of the time interval from which it was computed. The same rule was followed to synchronize \( %R(t) \) and \( %R(t) \) [11].

The presence of transient deterministic changes in the EEG (an ERP) was assessed by time averaging \( V(t) \) and by evaluating \( %R(t) \), either by time averaging an appropriate number of independent epochs[11].

Recurrence analysis of experimental signals involves the choice of specific values for important parameters including embedding dimension \( d \), lag time \( t_l \), scale for calculation of \( %R \), EEG window for RA calculation, and averaging window for calculation of \( %R(t) \).

In next two sections, we present methods for estimating embedding dimension \( d \) and time lag \( t_l \) of recurrence analysis parameters, but as mentioned previously other parameters must be determined empirically by systematically varying their values and assessing the effect on the ability to analysis an ERP.

C. Embedding Dimension Estimation

For embedding dimension estimation, we must first illustrate data organization.

We write out a set of \( d \)-dimensional vectors, which are called \( \tilde{V} \). The lag time \( t_l \) is given as a multiple of the time interval between successive samples: \( t_l = L t_o \). The jump time \( t_j \) between successive vectors is also given as a multiple of \( t_o \). Both \( L \) and \( J \) are taken to be positive integers. Assume \( x_n \) is the variable value recorded at time \( t=nt_o \). The set of vectors is:

\[
\tilde{V}_0 = (x_0, x_L, \ldots, x_{(d-1)L})
\]

\[
\tilde{V}_1 = (x_{J}, x_{J+L}, \ldots, x_{J+(d-1)L})
\]

\[
\vdots
\]

\[
\tilde{V}_J = (x_{JL}, x_{JL+L}, \ldots, x_{JL+(d-1)L})
\]

Also we shall use correlation sum to illustrate how to implement the embedding technique. The correlation sum as defined in the following form [17]:

\[
C^{(d)}(R) = \frac{1}{N(N-1)} \sum_{i,j=1, i \neq j}^{N} \theta[R - |\tilde{x}_i - \tilde{x}_j|] \tag{2}
\]

The subscript of \( C \) in equation (2) indicates that the correlation sum may depend on \( d \), the number of embedding dimension, furthermore \( N \) indicates number of vectors in embedding space, \( R \) is the radius in \( d \)-dimensional space and \( \theta(x) \) is the Heaviside step function which can be written in the following form:

\[
\theta(x) = \begin{cases} 
0 & \text{if } x < 0 \\
1 & \text{if } x \geq 0 
\end{cases} \tag{3}
\]

Also \( \tilde{x}_i \) is \( d \)-dimensional vector that is the collection of \( d \) components \( \tilde{x}_i = (x_i, x_{i+L}, x_{i+2L}, \ldots, x_{i+(d-1)L}) \).

The length of difference between two vectors is usually taken to be the “Euclidean norm”[16].

Now we are ready to find the correlation dimension. We define \( D_c(d) \) to be the number that satisfies:

\[
C^{(d)}(R) = kR^{-D_c(d)} \tag{4}
\]

for some range of \( R \) values, which we call the scaling region [17].

What we do in practice is computing of \( D_c(d) \) for
and plot the values of \(D_c\) as a function of \(d\). We expect \(D_c\) to vary with \(d\) until \(d\) is equal to or becomes greater than about twice the dimension of the state space. For \(d=d_{sat}\) (a saturation value), \(D_c\) becomes independent of the embedding dimension \(d\) [17].

D. Lag Time Estimation

In this section, we first introduce some definitions. The first definition is the autocorrelation time, which can be defined in terms of the autocorrelation function:

\[
g(n) = \frac{\sum_{k} x_k x_{k+n}}{\sum_{k} |x_k|^2} \tag{5}
\]

\(g(n)\) compares a data point in the series with a data point located \(n\) units of time away. If, on the average, they are uncorrelated, then we have \(g(n)=0\). If they are nearly the same, then we find \(g(n)=1\) [17]. For data sets from stochastic or chaotic systems, the autocorrelation function is expected to fall of exponentially with time:

\[
g(n) = ae^{-n\tau} \tag{6}
\]

where \(\tau\) is called the autocorrelation time [17].

In practice, we find the autocorrelation function of EEG signal from equation (6) and compute the sample \(n_0\) that \(g(n_0)=ae^{-1}\), then we find autocorrelation time as follows:

\[
\tau = \frac{1}{n_0} \tag{7}
\]

Now we can estimate the lag time \((\tau_L)\) by assuming that \(\tau \approx 3(d-1)\tau_s\) [17].

III. RESULTS

In this section the results of embedding dimension estimation, lag time estimation and recurrence analysis are presented.

A. Results of Embedding Dimension Estimation

We have used four epochs of ERP signals that were related with episodic memory process for reconstruct 1024 points of EEG signal. We computed correlation sum for reconstructed EEG signal with equation (2) and plotted \(logCR\) versus \(logR\), then we found the scaling region and fitted a line to points in scaling region. Finally we found the slope of line and assigned it to correlation dimension.

Then we computed correlation dimension \((D_c(d))\) for \(d=1,2,3,\ldots\) and plotted the values of \(D_c\) as a function of \(d\) in figure 1. It is shown that \(D_c\) change with \(d\) until \(d\) is equal to or becomes greater than \(d_{sat}\).

Here, we achieved \(d_{sat}=5\). Therefore embedding dimension \((d)\) is equal to 5 and actual dimension of EEG signal in state space is equal to 2. Because embedding dimension is twice the actual dimension of the state space plus one [17].

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{A schematic graph of the correlation dimension plotted as a function of embedding dimension. The embedding dimension of the EEG signal in this case is 5.}
\end{figure}

B. Results of Lag Time Estimation

We considered one epoch of ERP signals with length 256 samples, then computed autocorrelation function with equation (5) and plotted it versus samples in figure 2. As we see in figure 2, \(g(n)\) starts from one value and descent exponentially to zero. We can find autocorrelation time \((\tau)\) through equation (7), then by assuming \(\tau \approx 3(d-1)\tau_s\), we can estimate the lag time \((\tau_L)\). From this we obtained \(\tau_L = \frac{3}{48}\).

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig2.png}
\caption{A schematic graph of the autocorrelation function plotted as a function of samples.}
\end{figure}

From equation (1) and considered lag time as a multiple of the time interval between successive samples: \(\tau_L=\tau_s\), we can find \(L\) as follows:

\[
L = \frac{\tau_L}{\tau_s} = \frac{3/48}{1/256} = 16 \Rightarrow \frac{L}{\tau_s} = \frac{3}{48} \tag{8}
\]

C. Results of Recurrence Analysis

After finding embedding dimension and lag time, now we can use RA to analysis ERP signals. But other parameters have not specified yet. We consider scale 15% Euclidean norm for calculation of %R, a window of 70 EEG points for
RA calculation and 30 points averaging window for calculation of \( \%R(t) \) empirically.

In the first step, we found \( \%R(t) \) for each epoch of ERPs, and then averaged \( \%R(t) \) signals according to old/new stimulus for each subject. In the next step, we extracted time domain features from each averaged \( \%R(t) \) same as previous study[12] and then classified averaged \( \%R(t) \) according to old/new stimulus, using linear discriminant analysis (LDA) and stepwise method via SPSS software. Likewise we used leave one out (LOO) method to validation data. Table1 shows the classification results.

**TABLE 1**

<table>
<thead>
<tr>
<th></th>
<th>OLD</th>
<th>NEW</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAIN</td>
<td>98.5%</td>
<td>100%</td>
<td>98.9%</td>
</tr>
<tr>
<td>TEST</td>
<td>97.0%</td>
<td>100%</td>
<td>97.7%</td>
</tr>
</tbody>
</table>

Here, we have improved classification results of about 10\% in comparison with previous study[12].

In the next level, we used ERPs without digitally filtering (data with line interference noise), and ran RA and all other processing to classifying them. The results are shown in table2.

As we can see, the classification results in this level are reliable yet, which express RA can classify noisy data too.

**TABLE 2**

<table>
<thead>
<tr>
<th></th>
<th>OLD</th>
<th>NEW</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAIN</td>
<td>78.8%</td>
<td>81.8%</td>
<td>79.5%</td>
</tr>
<tr>
<td>TEST</td>
<td>74.2%</td>
<td>81.8%</td>
<td>76.1%</td>
</tr>
</tbody>
</table>

IV. CONCLUSION

In this study we presented a method based on RA for analyzing ERPs which appear in episodic memory retrieval. Because RA consider nonlinear nature of ERPs, it can classify ERPs with better results than previous study[12]. Furthermore RA can classify ERPs with line interference noise.

Since RA based on relativity and phase space, then it does not have any sensitivity to amplitude and scaling of data, and it does not need to data normalization.

Likewise in this study, we presented methods to finding some parameters of RA. We could estimate embedding dimension and lag time, but other parameters have been chosen empirically yet. Whereas in the same previous study[11] all parameters have been chosen empirically.

At last, it seems that finding a method for choosing all parameters of RA through special criterion based on nonlinear and phase space analysis will improve the classification results.

REFERENCES


