International Journal of Multimedia and Ubiquitous Engineering
Vol.9, No.6 (2014), pp.81-92
http://dx.doi.org/10.14257/ijmue.2014.9.6.09

Ordinal Pattern Analysis Method Applied in a P300-based Brain Computer Interface

Mohammed J. Alhaddad, Mahmoud I Kamel and Dalal M. Bakheet

Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

{Malhaddad, miali, dbakheet} @ kau.edu.sa

Abstract

Ordinal Pattern analysis has been used recently for extracting qualitative information from non-linear time series and it has been applied to usefully track brain dynamics. In this paper, we proposed a novel P300-based BCI system which depends on ordinal time series analysis as a feature extraction method. We have shown that this method can efficiently revel P300 feature, and therefore good classification accuracies and bitrates have been achieved for healthy and disabled subjects.

Keywords: *Brain computer interface; P300; Nonlinear analysis; Ordinal pattern; Pattern recognition; Signal processing*

1. Introduction

Brain computer interface (BCI) is a novel pattern recognition system that translates human thoughts or intentions into a control signal. This can achieve by measuring brain signals, such as electroencephalogram (EEG) signals, and translate them into control commands for a computer or other devices. Such a system allows people with severe motor disabilities to manipulate their environment in an alternative way.

Signal processing operation (which corresponds to the “feature extraction” process in machine learning terms) is the first step toward the translation of brain signals. Processing operations intended to reduce the dimensionality and also the complexity of a raw EEG signal to a few dimensions with the largest informative features. In general, this can be done by transforming the EEG signal from time series representation to another representation which is useful to reveal characteristics or features of the data that are relevant for BCI task.

There are varying numbers of possible transformation methods have been used to handle the EEG time series problems. They range from traditional linear analysis methods such as time domain analysis methods, frequency domain analysis methods, and spatial domain analysis methods [1]; to nonlinear analysis approaches derived from nonlinear dynamical system theory (also called deterministic chaos theory) [2] such as correlation dimension, Lyapunov exponents, and entropy [3].

Recently, Bandt and Pompe [4] developed an ordinal analysis method as a novel nonlinear analysis method. The basic idea behind this method is to convert a given time series into a sequence of ordinal patterns. These patterns describe the order relations between the values of a time series and not the values themselves. The main advantage of this representation is its robustness where the non-stationary process transforms to a (nearly) stationary (ordinal process). Furthermore, the transformation of a time series into a sequence of ordinal pattern

ISSN: 1975-0080 IJMUE

can be done in a computationally extremely fast way, and it is robust in the presence of observational and dynamical noise [4].

These advantages facilitate the use of ordinal time series method for the analysis of long and complex EEG signals. So far the main application to EEG analysis has been the analysis of epileptic activity. Faul *et al*. [5] have discussed the possibility of using permutation entropy (which proposed in [4] to measure the complexity of time series) for the detection of epileptic seizures in brain signals taken from newborn babies. Cao *et al*. [6] have used permutation entropy for describing the intrinsic EEG activity. Keller and Lauffer [7] also used permutation entropy to analyze side differences in EEG measuring and the effect of vagus stimulation.

Apart from the permutation entropy, Ouyang *et al*. [8] have tested the dissimilarity measure for investigating dynamic changes in EEG, in particular in the transition of epileptic seizures. On the other hand, Keller and Wittfeld [9] have discussed correspondence analysis based on symbolic dynamics in order to detect coupling between 19-channel of EEG recordings taken from children with epileptic disorders.

Although ordinal pattern gained significant attention in the literature, the use of ordinal method as a nonlinear feature extraction method of a BCI system have not investigated yet. Hence, we aim to propose a novel P300-based BCI system which depends on the ordinal pattern transformation as a feature extraction method.

The layout of this paper is as follows. In Section 2, a brief introduction of P300-based BCI, is given. Then, Section 3 gives a description of nonlinear time series analysis. In Section 4, the precise definition of ordinal time series analysis is given. Then, In Section 5, the datasets and the methods used for data preprocessing, feature extraction and classification are described. The results of the proposed system are discussed in Section 6, and the conclusion follows in the final section.

2. P300 Brain Computer Interfacing

Brain computer interface (BCI) is a new communication system that does not depend on normal neural or muscular peripheral pathways of the brain. Therefore, it is an excellent choice for people who are suffering from motor disabilities caused by any disease impairs the neural pathways that control muscles, or impairs the muscles themselves [10]. Furthermore, it provides a communication option for nondisabled users with situational disabilities induced by their environment, such as with aircraft pilots and astronauts subjected to extreme forces, soldiers in hostile territory, and even video gamers [11].

The system used in this study is a stimulus-driven BCI system based on Event Related Potentials (ERPs). ERPs are distinct patterns generated by the brain of the subject after or during the presentation of preselected visual and/or audio stimuli. These patterns can be detected by analyzing the recorded EEG signals, and it can be specified which stimulus has drawn the subject's attention [12].

P300 is endogenous Event Related Potentials that has gained significant attention in neuroscience and medical research communities. It is a positive deflection in the human EEG, appearing around 300 ms after the presentation of unusual or unexpected stimuli. To evoke P300, different stimulus modalities and paradigms can be used. Regarding the stimulus modality, often auditory or visual stimuli types can be presented to the subject. Regarding the paradigms, either oddball or three-stimulus paradigms are used. In oddball paradigm, two different stimuli, namely target (or oddball) and non-target stimuli are used [12]. The two stimuli are presented to the subject, but the target stimulus appears occasionally. Subjects should respond to each occurrence of the target stimulus, and should ignore the non-target stimuli.

In three-stimulus paradigm, however, a third distracter stimulus is also used. This stimulus appears in sequence with the same frequency as target stimulus, but the subject is asked to ignore it and does not perform any task when observing this stimulus.

Different types of P300 can be observed in the two paradigms described above [12]. In classical oddball paradigm, the target stimuli evoke the so-called P3b. The P3b has latency of about 300-500 ms and can be observed mainly over centre-parietal brain regions. This component appears only if subjects pay attention to the stimuli. When subjects do not pay attention to the stimuli, the target stimuli in the oddball paradigm evoke the so-called P3a. The P3a has latency of about 200-400 ms and can be observed mainly over front-central brain regions. In three-stimulus paradigm, the target stimuli also evoke P3b. The distracter stimuli, however, evoke P3a.

The basic idea underlying P300 based BCI systems are to use an oddball-like paradigm and to allow the user decide which stimulus is the target stimulus. Figure 1 illustrates a typical P300-based BCI system framework. In general, the sequence of events in a system is as follows. First the user decides on a command he wants to accomplish with the help of the BCI. Then stimuli are presented, and the user concentrates on the stimulus associated to the desired command. After stimulus presentation, the brain activity is recorded using a signal acquisition device, and then preprocessing operations are used to remove unnecessary information such as noises and artifacts. Brain signals are then processed with signal processing and classification algorithms. The outcome of the classification is fed into an application, which usually generates feedback to inform the subject about the result of classification.

3. Nonlinear time Series Analysis

3.1. Nonlinearity of the Medium

Neurons are known to be nonlinear devices because they become activated due to the changes in brain metabolism (as a result of biological and physiological phenomena in the human body) or due to the activity of the brain itself [3] [13]. Hence, an EEG signal can be considered as the output of a dynamical nonlinear system. Although the analysis of such a system is complicated, some measures borrowed, from chaos theory, to accomplish this task. A historical review of nonlinear EEG analysis can be found in [2].

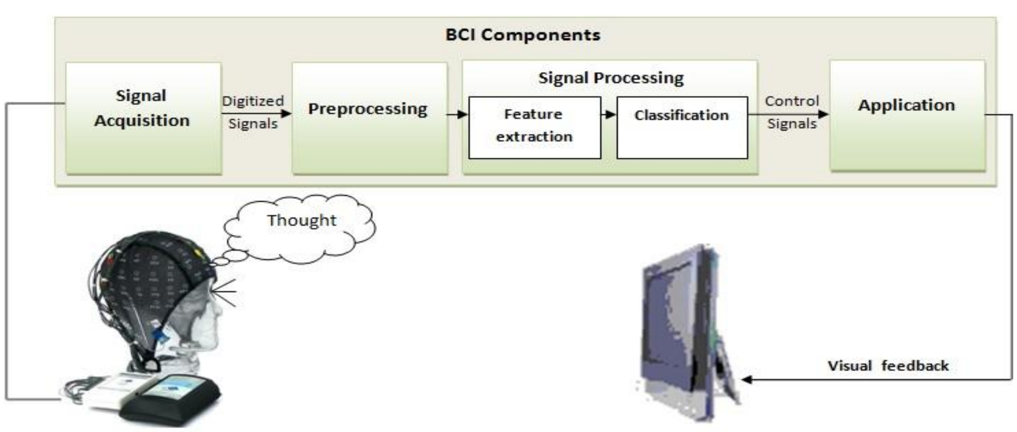

Figure 1. Basic Block Diagram of BCI Systems

A dynamical system is defined as a system that changes its state over time, sometimes in rather complex ways. Understanding, processing, and classifying such changes is of greatest importance for the analysis of EEG signals. Formally, a dynamical system is given by (1) a phase space, (2) a continuous or discrete time and (3) a time-evolution law. The elements or points that represent possible states of the system are called state variable, and the space made up of the state variables is called phase space or state space [14]. Let us assume that the state of such a system at a fixed time t can be specified by m components. These parameters can be considered to constitute a vector

$$\vec{x}(t) = (x_1(t), x_2(t), \dots\dots\dots\dots\dots\dots\dots, x_m(t))^T, \tag{1}$$

in the m-dimensional phase space of the system. The time-evolution law is a rule that allows calculating all future states given a state at any moment. The vector $\vec{x}(t)$ defines a trajectory in phase space which is a path followed by a dynamical system as time progresses [14].

A dynamical system may be a linear system if all the equations describing its dynamics are linear; otherwise it is nonlinear. On the other hand, a dynamical system can be deterministic if the equations of motion (which every subsequent state of the system must follow) are predefined, and stochastic otherwise. However, the neural networks of the brain, which concern us in the present context, are likely to be a chaotic system. The important features of such a system are its nonlinearity and deterministic. Although chaotic systems are deterministic, their behavior shows sustained irregularity [3].

An important property of the chaotic systems is that, after a long time observation, the trajectory will converge to a subspace of the whole phase space. This subspace is called the attractor of the system since it 'attracts' trajectories from all possible initial conditions [2]. The attractor, in chaotic systems, is an important object since it gives us an image of the complexity of the systems dynamics.

3.2. Reconstruction of Dynamics from Observation

From the previous section, we can conclude that nonlinear EEG analysis approach aims to reveal, from observations, the nature of the underlying dynamics, as well as its complexity. This means that the nonlinear analysis works in reverse; starting with the output of the system (EEG signals), and working back to the state space, attractors and their properties. Thus, the first step toward nonlinear EEG analysis is to reconstruct, from one or a few time series of observations, an attractor in the phase space of the underlying system (the brain). In other words, EEG signals need to be transformed from time series representation into the attractor representation which reflects the dynamics of the brain. However, due to unknown mixing of the underlying system variables, the observations usually do not have a one to one correspondence with these variables. This problem can be solved using the embedding procedure. Embedding allows us to reconstruct an equivalent attractor of the underlying system dynamics. In other words, even if we do not have direct access to all system variables, embedding procedure allows us to obtain informative information about the dynamics of the system.

A frequently used method for embedding procedure is the time delay method [15].

$$X_t = \left[X_t, X_{t+\tau}, X_{t+2\tau}, \dots\dots\dots\dots\dots\dots, X_{t+(m-1)\tau}\right]. \tag{2}$$

In this method, we start with a single time series of observation X_t. By using time delay τ we reconstruct the m-dimensional vectors by taking m consecutive values of the time series as the values for the m coordinates of the vector. By repeating this procedure for the next m values of the time series, we obtain the series of vectors in the state space of the system. The connection between successive vectors defines the trajectory of the system. In practice, we do not use values of the time series of consecutive digitizing steps, but we use values separated by a small lag. Thus, time-delay embedding is characterized by two parameters: time lag τ and the embedding dimension m. The proper choice of these parameters is necessary but difficult task in nonlinear time series analysis.

As found by researchers, too small delays result in near linear reconstructions with high correlations between consecutive state space variables and too large delays may ignore any deterministic structure of the sequence [3]. The auto-correlation function and mutual information are two early approaches that are used to estimate time delay parameter [16] [17]. On the other hand, embedding dimension m also should be estimated properly. In general, time-delay theorem requires that the proper value of m should be at least 2d+1 (where d is the actual dimension of the attractor) in order to guarantee reconstruction of the phase space. The false nearest neighbors [18] is the most known technique to estimate the optimum dimension m. However, in data with significant amounts of noise, such as EEG signals, it is common to see extremely high embedding dimensions [19].

In addition, it is important to know that time lag τ and dimension m are closely related. Based on this viewpoint, new algorithms to estimate m and τ in the same time have been suggested in the category of the chaotic time series analysis. These new approaches, as well as the traditional ones, have been reviewed in [20].

Unfortunately, the embedding estimation approaches may sometimes give an improper embedding, which, from its side, can cause a considerable amount of spurious correlations between the distances of embedded vectors in phase space [21]. Although this problem has gained a lot of emphasis in the literature, with today's computer power, good embeddings may be found by trial and error. Therefore, we can say that the embedding parameters are optimal if it delivers the best possible estimates for the topological properties of the attractor, or if it provides the most accurate prediction of the time series [22].

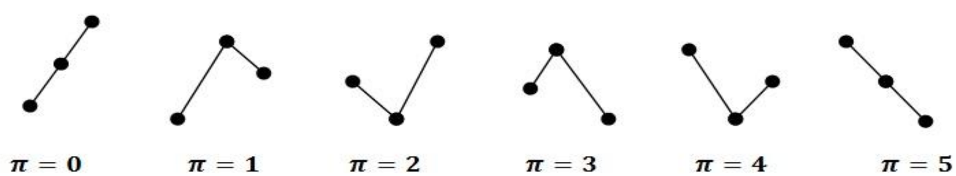

Figure 2. Order Patterns at Embedding Dimension m=3

Once, the attractor has been reconstructed, the next step toward nonlinear analysis is to characterize it in a quantitative way. Many measures are introduced (in the physics literature) to do this; such as the dimension, the Lyapunov exponents, and entropy.

Bandt and Pompe [4] developed ordinal time series analysis method as a new method of nonlinear analysis. The precise definition of ordinal pattern analysis is given in the next section.

4. Ordinal Pattern Analysis

Ordinal time series analysis method is based on transforming a reconstructed vector into a series of ordinal pattern describing the order structure of time-dependent. Consider a given reconstructed vector X_t as shown in Equation 2. This vector can be rearranged in an increasing order and encoded by an ordinal pattern $\pi(t)$ as follows:

$$[X_t \leq X_{t+\tau} \leq X_{t+2\tau} \ldots\ldots\ldots\ldots\ldots \leq X_{t+(m-1)\tau}] \rightarrow \pi(t), \tag{3}$$

where $\pi(t)$ could be an integer number $0 \leq \pi(t) \leq m!-1$. To obtain a unique result, we set $x_{t+(u-1)\tau} < x_{t+(u)\tau}$ in the case of $x_{t+(u-1)\tau} = x_{t+(u)\tau}$. It is clear that there are at most m!

different ordinal patterns π possible, which are also called permutations. For example, at dimension m=3 we get six order patterns as shown in Figure 2, and we interested in the following six relations:

$$\begin{array}{lcl}
x_t < x_{t+\tau} < x_{t+2\tau} & \rightarrow & \pi(t) = 0\,(123) \\
x_t < x_{t+2\tau} < x_{t+\tau} & \rightarrow & \pi(t) = 1(132) \\
x_{t+\tau} < x_t < x_{t+2\tau} & \rightarrow & \pi(t) = 2(213) \\
x_{t+\tau} < x_{t+2\tau} < x_t & \rightarrow & \pi(t) = 3(231) \\
x_{t+2\tau} < x_t < x_{t+\tau} & \rightarrow & \pi(t) = 4(312) \\
x_{t+2\tau} < x_{t+\tau} < x_t & \rightarrow & \pi(t) = 5(321)
\end{array}$$

Having the transformed data on hand, we can use it as nonlinear features to train and test a BCI classifier.

Figure 3. The Paradigm Used for Evoking P300 [23]

5. Materials and Methods

The main aim in this study is to evaluate the performance of a p300-BCI system using ordinal pattern method as a feature extraction method. To this end, we have used the datasets and some of the preprocessing algorithms which are made available on the website of EPFL BCI group (http://bci.epfl.ch/p300). A brief description of these datasets and algorithms is given in the following sections. Please refer to the work of [23] for more details of these materials.

5.1. Datasets

The datasets were collected while subjects were facing a laptop screen on which a paradigm consists of six images (see Figure 3) was displayed. The images showed a television, a telephone, a lamp, a door, a window, and a radio. The images were flashed in random sequence, one image at a time. In Hoffman *et al*. [23] study, each subject had to complete four recording sessions. Each session consisted of six runs, one run for each of the six images. The sequence of images to be flashed in each run was block- randomized which means after each block (6 flashes), each image was flashed one time, and after two blocks (12 flashes) each image was flashed twice and so forth. The number of the blocks inside each run was selected randomly between 20 and 25. Thus, on average 22.5 blocks of six flashes were displayed in each run. This means that one run consisted on average of 22.5 target trials (*i.e.*, contained P300 signal) and 22.5× 5=112.5 non-target trials (*i.e.*, did not contain P300 signal).

Each flash of an image lasted for 100 ms, and during the following 300 ms, none of the image was flashed, *i.e.*, the ISI was 400 ms.

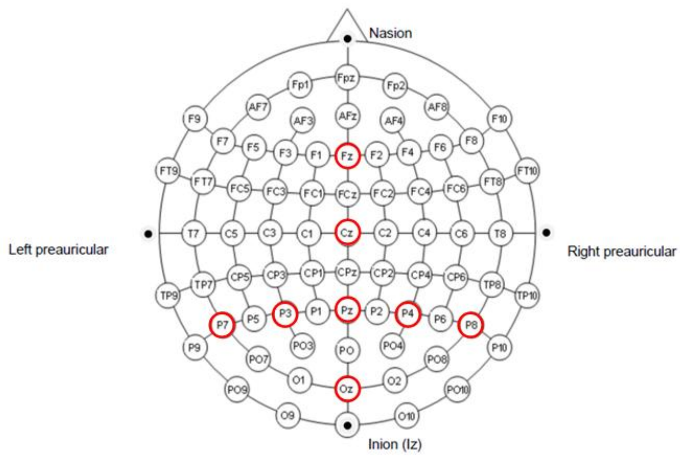

Figure 4. The Eight Electrode Configuration used in the Experiments from the Standard Positions of 10-20 International System

The EEG was recorded at 2048 Hz sampling rate from 32 electrodes placed at the standard positions of 10-20 international system, and a Biosemi Active Two amplifier was used for amplification and analog to digital conversion of the EEG signals.

5.2. Offline Analysis

The impact of using nonlinear ordinal pattern analysis on classification accuracy was tested in an offline procedure. In this analysis, the datasets of eight subjects (four disabled and four healthy subjects) were tested. A breakdown of the pathological condition of the disabled subjects can be found in [23]. For each subject, fourfold cross-validation was used to estimate average classification accuracy. More specifically, data from three recording sessions were used to train a classifier and data from the left-out session were used for testing. This procedure was repeated four times, so each session served once for testing.

5.3. Preprocessing Operations

During the signal acquisition, two electrodes were placed on the mastoids of the subjects. The average signal of these two electrodes was used for referencing. Then, a sixth order forward-backward Butterworth band-pass filter was used to filter the data. Then, we down-sampled the data from 2048 Hz to 256 Hz by selecting each 8th sample from the band-pass filtered data.

To extract the single trials from the whole EEG data gathered during each run, windows of duration 1000 ms were extracted from the data. The extracted trials need then to be purified from artifacts, such as eye blinks, eye movement and muscle activity or subject movement. This was done by computing, for the samples from each electrode, the 10th percentile and the 90th percentile. Then, amplitude values lying below or above this percentile were replaced by the 10th percentile or the 90th percentile respectively. This procedure is called windsorizing. Finally, the preprocessed samples were normalized.

At the end of preprocessing operations, each trial has been represented by a two dimensional matrix $\mathrm{N_e} \times \mathrm{N_t}$, where $\mathrm{N_e}$ denotes the number of electrodes, and $\mathrm{N_t}$ denotes the number of temporal samples. Due to the trial duration of 1s (1000 ms) and the down-sampling to 256 Hz, $\mathrm{N_t}$ equaled 256. In this research, the eight electrode configuration was used as shown in Figure 4. Hence, $\mathrm{N_e}$ equaled 8.

5.4. Ordinal pattern calculations

After the preprocessing operations, ordinal patterns for each trial (from each channel separately) were calculated. To this end, first the attractor of a chosen delay τ and dimension m, was reconstructed using the embedding method of delays [15] discussed in (Section 3.2). Then, permutation vectors of the reconstructed attractor were calculated and rearranged. At the end, a list of pattern corresponding with the underlying trial was found.

To choose the optimal values of delay τ and dimension m, we found that using traditional methods, such as auto-correlation function (for choosing delay τ) and false nearest neighbors (for choosing dimension m) is not useful in our case. These traditional methods were used frequently in the literature where EEG activity, under study, represented by a single time series of measurements. Thus, auto-correlation function, or false nearest neighbors give single optimal results. However, BCI systems, in general, deal with multiple time series of independent measurements resulted from different channels, runs, and sessions. Hence, the traditional methods for choosing optimal delay and dimension fail to give single optimal results.

Therefore, the trial and error approach has been used in this study to choose optimal values of delay τ and dimension m. This approach was inspired from Kantz and Schreiber [22], where it was mentioned that the optimal embedding parameters are which provide the most accurate prediction of the time series. On other words and from pattern recognition point of view, we can say that the embedding parameters are optimal if they provide satisfactory classification accuracy.

Hence, for each subject, the procedure of the whole offline analysis was repeated a number of time (using random values of delay τ and dimension m) in order to get the embedding dimensions values which give best classification accuracy in a shortest possible time.

Finally, the calculated patterns were normalized and concatenated into feature vectors. The dimensionality of the feature vectors was $\mathrm{N_e} \times \mathrm{N_l}$, where $\mathrm{N_l}$ denotes the number of samples in one ordinal pattern series, and it equals to $\mathrm{N_t} - (\mathrm{m} - 1)\tau$.

5.5. Machine Learning and Classification

The percentile values used for windsorizing and the classifier were trained on data from three sessions and tested on the left-out fourth session. On average, training data sets contained 405 target trials and 2025 non-target trials, and testing data sets contained 135 target and 675 non-target trials. In this study, we used Bayesian Linear Discriminant Analysis (BLDA) algorithm, which developed by Hoffman [24], to learn the classifier. The algorithm of BLDA is also made available on the website of EPFL BCI group (http://bci.epfl.ch/p300). After we trained the classifier, we applied it to the testing data as follows: For each run in the testing session, we extracted (using preprocessing operations) the single trials corresponding to the first twenty blocks of flashes. Then, for each trial, we calculated its ordinal patterns. The resulted ordinal patterns were then concatenated to form the feature vectors. After that, the single trials were classified. Twenty blocks of classifier outputs were resulted from this process. Each block consisted of six outputs (one output for each image on the paradigm). In order to decide which image the subject was concentrating on, the summing scores aggregation method was used. This method was first described by Farwell and Donchine [25], and is often used by many researchers to aggregate information from several single trials. Hence, the classifier outputs were summed over blocks for each image. Then, the image with the maximum summed classifier output was selected. Different tradeoffs between the classification accuracy and the time needed to take a decision were simulated by varying the number of blocks (*i.e.*, the number of summed classifier outputs).

All preprocessing, ordinal pattern calculation and classification algorithms were implemented with MATLAB.

6. Results

The resulted classification accuracy of the system using ordinal pattern features were averaged over sessions. Then, the averaged accuracies were plotted against the time needed to take a decision as shown in Figure 5. All subjects (except subject 6 and 9) reached an average of 100% classification accuracy after 5 blocks of stimulus presentations. The total numbers of blocks needed to achieve 100% classification accuracy for each subject are shown in Table 1. The lower performance of subject 6 and 9 is due to wrong concentration on some sessions as explained in [23].

In addition, the bitrates of the system, were computed by applying the Wolpaw *et al.* method [10] to the average accuracy curves. The resulted bitrates for each subject are listed in Table 2.

However, these good results can only be achieved using specific values of embedding dimension τ and m. Trial and error method was used, in this study, to get these values.

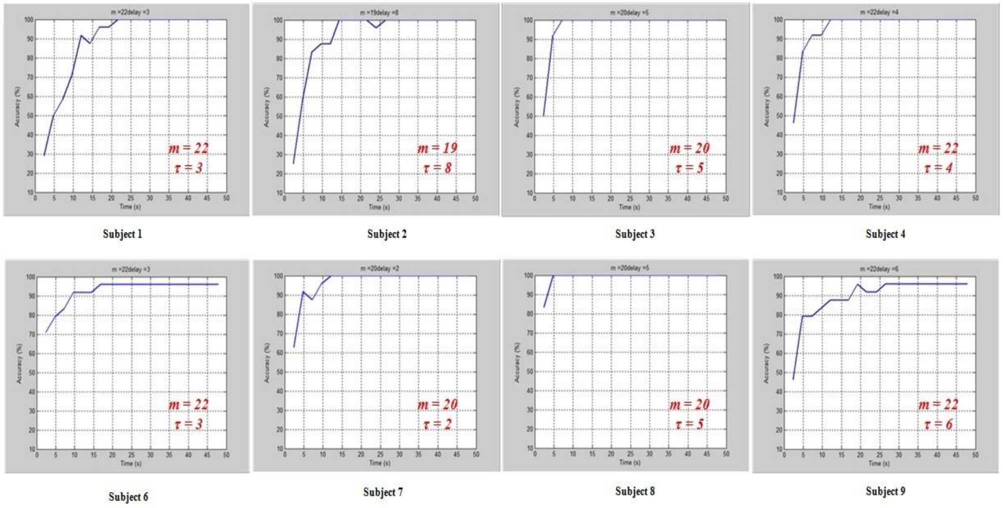

Figure 5. Classification accuracy plotted vs. time. The panels show the classification accuracy obtained using ordinal features (through the shown values of m and τ) for disabled subjects (S1-S4) and able-bodied subjects (S6-S9)

The optimal embedding values for each subject are shown in Figure 5.

Differences between disabled and healthy subjects can be strongly observed using the bitrate measure. In general, healthy subjects achieved higher maximum bitrates than disabled subjects.

Table 1. Total Numbers of Blocks Needed to Achieve 100% Accuracy are shown for all subjects except S6 and S9. Mean blocks numbers were computed for all subjects

Subject	S1	S2	S3	S4	S7	S8	Average (all)
No. of blocks	9 blocks	10 blocks	3 blocks	5 blocks	5 blocks	2 blocks	5 blocks

Table 2. Maximum average bitrates per minute. Bitrates were computed from average accuracy curves and are shown for all subjects. Mean bitrates were computed for disabled subjects (S1-S4), healthy subjects (S6-S9), and all subjects

Subject	S1	S2	S3	S4	S6	S7	S8	S9	Average (S1-S4)	Average (S6-S9)	Average (all)
Bitrate per minute	9.888	12.900	24.721	19.349	25.922	24.721	38.699	17.037	16.715	26.595	21.655

7. Conclusion and Discussion

In this work we presented a novel P300 based BCI system that depends on ordinal time series analysis as a feature extraction method. We have shown that the ordinal method can efficiently revel P300 features, and hence achieve good classification accuracy for healthy and disabled subjects. However, the success of this method depends strongly on the optimal choice of embedding dimensions. The trial and error approach, which inspired from [22], have been used in this study in order to choose the optimal values of delay τ and dimension m. However, it doesn't practical to use trial and error approach in the real (online) sessions of

BCI systems where time is crucial. This approach will be more appropriate if we can consider the values of the delay τ and the dimension m as subject specific parameters. Therefore, we can choose the optimum dimensions in an offline sessions (as we done in this work), and then train a BCI system with it prior to any real sessions. To this end, we need to test the proposed system using larger numbers of sessions in order to verify that the delay τ and the dimension m can be considered as subject specific parameters.

Acknowledgements

Many thanks go to Hoffmann *et al.* [23]. Their code helped us in developing many preprocessing algorithms. We would like also to thank our team for their efforts in the BCI project. This research was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, under grant, No. (16-15-1432 HiCi). The authors, therefore, acknowledge with thanks DSR technical and financial support.

References

[1] M. W. Tangermann, "Feature Selection for Brain-Computer Interfaces", Citesee, Berlin, (**2007**).
[2] C. Stam, "Nonlinear dynamical analysis of EEG and MEG:Review of an emerging field", Clinical Neurophysiology, vol. 116, (**2005**), pp. 2266–2301.
[3] S. Sanei and J. A. Chambers, “EEG SIGNAL PROCESSING”, UK: John Wiley and Sons, Ltd, **(2007)**.
[4] C. Bandt and B. Pompe, "Permutation Entropy: A Natural Complexity Measure for Time Series", Physical Review Letters, **(2002)**.
[5] S. Faul, G. Boylan, S. Connolly, W. Marnane and G. Lightbody, "Chaos Theory Analysis Of The Newborn EEG – Is It Worth The Wait?", in IEEE International Workshop on Intelligent Signal Processing, Faro (Portugal), **(2005)**.
[6] Y. Cao, W.-w. Tung, J. B. Gao, V. A. Protopopescu and L. M. Hively, "Detecting dynamical changes in time series using the permutation entropy", PHYSICAL REVIEW E, vol. 70, **(2004),** pp. 046217.
[7] K. Keller and H. Lauffer, "Symbolic analysis of high-dimensional time series," Int. J. Bifurcation Chaos, vol. 13, no. 9, **(2003),** pp. 2657-2668.
[8] G. Ouyang, C. Dang, D. A. Richards and X. Li, "Ordinal pattern based similarity analysis for EEG recordings", Clinical Neurophysiology, **(2009)** December 22, pp. 694–703.
[9] K. Keller and K. Wittfeld, "Distances of Time Series Components by Means of Symbolic Dynamics", Int. J. Bifurcation Chaos, vol. 14, no. 2, **(2004),** pp. 693-704.
[10] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, "Brain–computer interfaces for communication and control", Clinical Neurophysiology, **(2002),** pp. 767-791.
[11] A. B. Randolph and M. M. Moore Jackson, "Assessing Fit of Nontraditional Assistive Technologies", ACM Transactions on Accessible Computing, **(2010)**.
[12] A. Yazdani, J.-S. Lee and T. Ebrahimi, "Implicit Emotional Tagging of Multimedia Using EEG Signals and Brain Computer Interface", in WSM '09 Proceedings of the first SIGMM workshop on Social media, New York, **(2009)**.
[13] N. Marwan and A. Meinke, "Extended Recurrence Plot Analysis and its Application to ERP Data", International Journal of Bifurcation and Chaos, Special Theme Issue "Cognition and Complex Brain Dynamics", **(2002)** December 20**,** pp. 761-771.
[14] T. Munakata, "Chaos", in Fundamentals of the New Artifical Intelligence, London, Springer, **(2008)**, pp. 206-245.
[15] F. Takens, "Detecting strange attractors in turbulence", Dynamical Systems and Turbulence, Lecture Notes in Mathematics, Springer, **(1981),** pp. 366–381.
[16] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees and P. E. Rapp, "Singular-Value Decomposition and the Grassberger-Procaccia Algorithm", Physical Review A, vol. 38, **(1988)**.
[17] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information", Phys. Rev. A, vol. 33, no. 2, **(1986),** pp. 1134–1140.
[18] H. Abarbanel, "Analysis of Observed Chaotic Data", Springer, **(1996)**.
[19] N. Rohrbacker, "Analysis of Electroencephologram Data Using Time-Delay Embeddings to Reconstruct Phase Space", Dynamics at the Horsetooth , vol. 1, **(2009)**.
[20] M. Hong-guang and H. Chong-zhao, "Selection of Embedding Dimension and Delay Time in Phase Space Reconstruction", Frontiers of Electrical and Electronic Engineering in China, 12, **(2005),** pp. 111-114.

[21] M. THIEL, M. C. ROMANO and J. R. KURTHS, "Spurious Structures in Recurrence Plots Induced by Embedding", Nonlinear Dynamics, **(2006),** pp. 299–305.
[22] H. Kantz and T. Schreiber, "Nonlinear Time Series Analysis", Cambridge University Press, **(2004)**.
[23] U. Hoffmann, J. M. Vesin, T. Ebrahimi and K. Diserens, "An efficient P300-based brain–computer interface for disabled subjects", Journal of Neuroscience Methods, vol. 167, 15 January **(2008),** pp. 115-125**.**
[24] U. Hoffmann, "Bayesian Machine Learning Applied in a Brain-Computer Interface for Disabled Users", EPFL, LAUSANNE, **(2007)**.
[25] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials", Electroencephalography and Clinical Neurophysiology, **(1988),** pp. 510-523.

