This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2505238, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Analysis of high-dimensional phase space via Poincaré section for patient-specific seizure detection

Morteza Zabihi*, *Student Member, IEEE*, Serkan Kiranyaz, *Senior Member, IEEE*, Ali Bahrami Rad, *Student Member, IEEE*, Aggelos K. Katsaggelos, *Fellow, IEEE*, Moncef Gabbouj, *Fellow, IEEE*, and Turker Ince, *Member, IEEE*

***Abstract*— In this paper, the performance of the phase space representation in interpreting the underlying dynamics of epileptic seizures is investigated and a novel patient-specific seizure detection approach is proposed based on the dynamics of EEG signals. To accomplish this, the trajectories of seizure and non-seizure segments are reconstructed in a high dimensional space using time-delay embedding method. Afterwards, Principal Component Analysis (PCA) was used in order to reduce the dimension of the reconstructed phase spaces. The geometry of the trajectories in the lower dimensions is then characterized using Poincaré section and seven features were extracted from the obtained intersection sequence. Once the features are formed, they are fed into a two-layer classification scheme, comprising the Linear Discriminant Analysis (LDA) and naïve Bayesian classifiers. The performance of the proposed method is then evaluated over the CHB-MIT benchmark database and the proposed approach achieved an 88.27% sensitivity and 93.21% specificity on average with 25% training data. Finally, we perform comparative performance evaluations against the state-of-the-art methods in this domain which demonstrate the superiority of the proposed method.**

***Index Terms*— Dynamics, EEG, phase Space, Poincaré section, seizure detection, two-layer classifier topology.**

I. Introduction

EPILEPTIC seizures are transient excessive neuronal discharges originated from cortical gray matter and considered as the main definition of epilepsy. Indeed the concept of epilepsy covers a wide range of disorders, which can be classified according to the variety in types of seizures. Therefore, the epileptologist in the first place, should diagnose if an epileptic seizure occurred and then determine the seizure type. For this purpose, electroencephalography (EEG) is commonly used due to its unique properties, such as cost-effectiveness and high temporal resolution, which make it an influential and compulsory tool for exploring the brain functioning of patients with epilepsy.

The estimated number of people suffering from epilepsy in the world is around 50 million [1]. In addition, an increasing need for recording EEG signals in the long term, and the contamination of these signals with physiological and non-physiological artefacts renders their interpretation through visual inspection only a daunting and challenging task. These factors add impetus to the need of an automatic seizure detection system to ease the neurologist's burden of inspecting such long-term EEG data [2]. Thus, in the recent years, several techniques have been developed in order to detect patterns of interest from background patterns, including time [3], frequency [4], time-frequency [5], and nonlinear methods [6] - [8].

Despite conventional time series analysis, nonlinear dynamics addresses nonlinear relationships among the variables of a system by investigating only the variables (i.e., states) in phase space whilst discarding time or spectral components. The main power of this approach is that it provides information regarding the underlying dynamics of the system without knowing all the factors in the system evolution. Hence, nonlinear time series analysis, unlike differential equations, is a top-down approach where information about the states of the system or the relationship among the states is not available. Therefore, the approach is to reconstruct the system dynamics in phase space and then quantify the reconstructed attractors (e.g., [9] and [10]).

Numerous measures originated from nonlinear dynamics have been introduced and used for the analysis of EEG signals. Correlation dimension [11], Lyapunov exponents [12], phase synchronization [13], and mutual dimension [14] can be named among the traditional and novel measures. Nevertheless, most of these measures do not have straightforward interpretations and can only be used as tentative indices. This may instigate a false impression of chaos, hence surrogate data tests are needed in order to check the validity of the analysis. In other words, surrogate data

Manuscript received April 17, 2015; accepted November 24, 2015. Asterisk indicates corresponding author.

M. Zabihi*, and M. Gabbouj are with the Department of Signal Processing, Tampere University of Technology, Tampere, Finland (e-mail: morteza.zabihi@tut.fi, moncef.gabbouj@tut.fi).

S. Kiranyaz is with the Electrical Engineering Department, Qatar University, Doha, Qatar (e-mail: mkiranyaz@qu.edu.qa)

A. B. Rad is with the Department of Electrical Engineering and Computer Science, University of Stavanger, Stavanger, Norway (e-mail: ali.bahrami.rad@uis.no).

A. K. Katsaggelos is with the Electrical Engineering Department, Northwestern University, Evanston, IL, USA (e-mail: aggk@eecs.northwestern.edu).

T. Ince is with the Department of Electrical and Electronics Engineering, Izmir University of Economics, Izmir, Turkey (e-mail: turker.ince@ieu.edu.tr).

analysis is required in this case to reject the least null hypotheses of linear stochastic and white noise [15] [16]. On the contrary, phase space reconstruction can be considered as a tool to demonstrate the evolution of a dynamic system through time, while the dynamic system can fall into several categories such as conservative or dissipative, linear or nonlinear, deterministic or stochastic. Therefore, using geometrical features, based only on the phase space and without asserting any assumption about the type of the underlying dynamic, made the studies independent from chaotic hypothesis and consequently independent from the surrogate data analysis [6].

However, despite the emphasized properties of the phase space, characterizing the reconstructed trajectories of EEG signals, purely based on its geometry, is quite rare and only few works exist up to date. In [17], the phase space of an EEG signal was reconstructed after decomposing the signal using empirical mode decomposition with application to seizure detection. Two measures, namely the 95% confidence ellipse area and the interquartile range of the Euclidian distance, were extracted from two- and three-dimensional phase spaces, respectively. M. Chang et al. compared the efficiency of features obtained by amplitude-frequency analysis [18] and autoregressive (AR) model [10] using phase space and raw data in a Brain Computer Interface (BCI) task. In their study, the phase space of two EEG channels was reconstructed and then the AR parameters, peak and mean values of the absolute value of amplitude samples in two frequency bands 8-13 Hz and 14-25 Hz were calculated as features. It was shown that these features improve the classification results in contrast to the same features extracted from the raw data.

In [19], the wavelet (Daubechies 4) coefficients of an EEG signal at 5 levels were used to plot a two dimensional phase space. Then, the Euclidean distances between the origin and every point in the phase space were calculated. The mean, median, average power and standard deviation of these distances in each sub-band were used as features in order to detect seizure events. The classification results on real EEG data showed the significance of the extracted features. Furthermore, the phase space of EEG signals has been studied in a behavioral neuroscience research. In [20], the slope of a regression line in a two-dimensional phase space was obtained as a function of different time lags and considered as a feature in classification of sleep-wake states. The results showed the features based on phase space achieved higher performance than the power spectral approach.

Accurate reconstruction of the phase space has a great impact on the characterization of its trajectory properties. In the time-delay embedding method this accuracy depends on the proper selection of the time lag and the embedding dimension. According to the embedding theorem, any time lag will be acceptable; however, it should be noted that choosing a too small or a too large time lag value prompts completely dependent and independent coordinates [6]. Thus, the necessary and major task in keeping the physical properties of attractors is to determine a large enough embedding dimension. In the aforementioned studies, the maximum embedding dimension of three (in [17]), or two (in [18], [10], [19], and [20]) were used, this is generally insufficient and hence inadequate for mathematical modeling of such complex signals. Many published works in EEG signal processing (e.g., [21] - [24]) propose high dimensional phase spaces, i.e., $d>3$. However, using a higher dimensional phase space makes the interpretation and visualization of the trajectories a challenging task.

In this study, in order to address the aforementioned deficiencies such as the false impression of chaos and the necessity of applying surrogate data analysis, a novel phase space method is proposed. The proposed method aims to capture the underlying dynamics of the epileptic seizures and hence to discriminate them from the non-seizure segments in an efficient way. Therefore, our primary goal is to create a new set of nonlinear features for seizure detection which decreases the computational complexity while at the same time increases the seizure detection accuracy. This method describes the characterization of the geometry of a high dimensional phase space in such a way that it keeps the reconstructed trajectories unfolded. More specifically, this paper describes a novel feature extraction method based on a high dimensional phase space along with a classification scheme, where the main objective is to maximize the seizure detection accuracy with a minimal feedback from a human expert. Besides the discrimination ability of the extracted features, there are two main factors potentially affecting the performance of the automatic seizure detection methods: variation in seizure types and the brain regions where seizures have originated. Epileptic seizures develop as the results of different disorders, and as such, they cause significant variations of seizure types among patients (inter-patient variability). Therefore, in this work patient-specific setting is employed as a convenient candidate for such classification problems since this approach has more potential to learn the patterns of seizures in each individual specifically. In order to address the second factor, the signals of all channels are utilized. Besides providing the information of seizures in different brain areas, this approach gives the flexibility on the proposed framework to be compatible with various EEG recording montages.

To accomplish these objectives, first we reconstruct the trajectories of each 1-second EEG segment with fixed values of time lag and embedding dimension (Section II.B). Then, a Poincaré map of the reconstructed trajectories is obtained using Poincaré section, which is mainly chosen according to the first and second Principal Components (PCs) of the phase space coordinates (Section II.C). Afterwards, 7 discriminative features (Section II.D) are extracted from the obtained Poincaré map and fed into a classifier topology with two-layer architecture. The first layer consists of 23 Linear Discriminant Analysis (LDA) classifiers and a Naïve Bayes classifier formed in the second layer (Section II.E) in order to fuse the decisions of the first layer classifiers and hence perform the final classification. Finally, the proposed approach is tested on a benchmark dataset with EEG recordings of pediatric patients with intractable seizure and compared with the state-of-the-art

methods in this domain. The paper is concluded in Section IV.

II. EEG Data Processing

A. EEG Dataset

The EEG recordings were collected from 23 pediatric patients (males, ages 3-22; and females, ages 1.5-19) at the Children's Hospital, Boston, to assess their candidacy for surgical operation [25] [26]. Nine to twenty four EEG recordings were recorded for each individual. All the recordings were labeled as seizure or non-seizure with one second resolution. The sampling frequency was 256 Hz with 16-bit resolution. There are 23 common channels (FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8, P8-O2, FZ-CZ, CZ-PZ, P7-T7, T7-FT9, FT9-FT10, FT10-T8, and T8-P8) for each recording, which are based on the international 10-20 system of EEG electrode positions (see Fig. 1). In this study, only those records which contain at least one seizure event were used. In TABLE I, the lengths of each EEG recording (used in this study) are shown. Recordings 21 and 1 were obtained from the same female patient with 1.5 year apart, which were considered as two extra patients in this study. In addition, patient 15 was excluded from our analysis because we failed to read the EEG data of this patient.

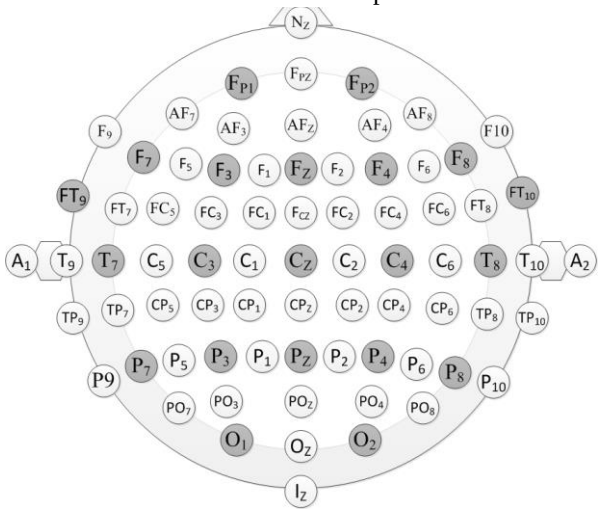

Fig. 1. Location of electrodes in international 10-20 system; the used channels are shown with gray color.

B. Phase Space Reconstruction

The phase space of a system represents how the states' dynamics evolve over time. The aim of phase space reconstruction is to obtain state vectors using the sequence of observed measurements. In this study, time-delay embedding method [27] is used for reconstruction of the EEG phase space. Takens' theorem expresses that the topological features of any higher dimensional system with coupled variables are reconstructable from a single time series of observations [28]. This theorem proves the independence of our study from the surrogate data analysis, which is mentioned in Section I.

The main idea is to create a series of time-shifted samples in d dimensions so that d coordinates would be provided using the map:

$$X[n] \rightarrow Y[n] = (X[n], X[n+T], \dots, X[n+(d-1)T]) \quad (1)$$

where T is the time lag. In order to determine the convenient dimension and time lag the two commonly known methods of correlation dimension and the mutual information are employed [6]. The embedding dimension 5 and time lag 6 (about 23 milliseconds at a sampling frequency of 256 Hz) were achieved and used for constructing an EEG attractor. The achieved values were validated empirically, where different time lags and embedding dimensions were used and their classification accuracy compared. In Fig. 2, the reconstructed phase spaces of sample seizure and non-seizure segments from the 1st, 17th and 21st patients are shown.

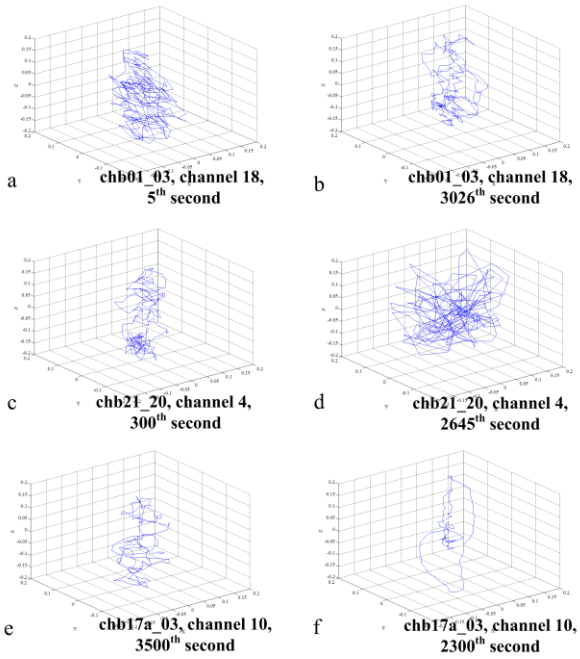

Fig. 2. The phase space plots of 1-s non-seizure segments in the left column (a, c, and e), and seizure segments in the right column (b, d, and f) reconstructed from the 1st, 17th and 21st patients (each row) in the CHB-MIT benchmark database. 3-D phase spaces are plotted for visualization.

C. Poincaré Section Delineation

Poincaré section, which was named in Henry Poincaré honor, is a well-known method for analyzing the type of attractors. In this method, a line (or plane) cut the attractor and then the intersection points are investigated. In fact, Poincaré section provides a geometric view of a trajectory's behavior through those intersection points. One application of Poincaré section, for instance, is to study the stability of limit cycles. This method also can be considered as a sampling method, which converts the continuous trajectories of a phase space to a discrete sequence of intersection points. The investigation of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2505238, IEEE Transactions on Neural Systems and Rehabilitation Engineering

TABLE I. CHB-MIT benchmark. The patients with longest and shortest duration of recordings are shown in bold.

Patient	Gender	Age	Number of seizure events (Tmax-Tmin in seconds)	Total duration of seizures (sec)	Total duration of non-seizures (sec)	Total duration (sec)
1	F	11	7 (28-102)	449	23476	23925
2	**M**	**11**	**3 (10-83)**	**175**	**7984**	**8159**
3	F	14	7 (48-70)	409	24791	25200
4	M	22	4 (50-117)	382	37977	38359
5	F	7	5 (97-121)	563	17437	18000
6	**F**	**1.5**	**10 (13-21)**	**163**	**93053**	**93216**
7	F	14.5	3 (87-144)	328	32209	32537
8	M	3.5	5 (135-265)	924	17076	18000
9	F	10	4 (63-80)	280	34219	34499
10	M	3	7 (36-90)	454	50010	50464
11	F	12	3 (23-753)	809	9250	10059
12	F	2	27 (14-98)	1016	33844	34860
13	F	3	10 (18-71)	450	24750	25200
14	F	9	8 (15-42)	177	25023	25200
16	F	7	8 (7-15)	77	17923	18000
17	F	12	3 (89-116)	296	10528	10824
18	F	18	6 (31-69)	323	19951	20274
19	F	19	3 (78-82)	239	10307	10546
20	F	6	8 (30-50)	302	19734	20036
21	F	13	4 (13-82)	203	13587	13790
22	F	9	3 (59-75)	207	10593	10800
23	F	6	7 (21-114)	431	31823	32254
24	Unknown	Unknown	16 (17-71)	539	42661	43200

how the trajectories pass through the Poincaré section reveals information about the dynamics of a system that is not obtainable otherwise [29].

The most common approach in this context is applying the Poincaré section on two- or three-dimensional attractors. The reason is that Poincaré section aims to generate a plot stroboscopically so that it samples the motion of the observed trajectories. In order to construct such a view, the Poincaré section must have a dimension less than the corresponding trajectories. Therefore, for instance, if the embedding dimension of a phase space is three, then a plane is used as a Poincaré section. In our case, where the number of embedding dimension is five, the Poincaré section will be a surface with four dimensions. In order to avoid using such complex geometric shapes and at the same time keep the trajectories unfolded without ambiguities, the following solution is proposed: first, Principal Component Analysis (PCA) is employed in order to convert the 5-dimensional embedding coordinates into a set of values of linearly uncorrelated principle components (PCs). Then, a polynomial with the degree of one (i.e., Poincaré section) is fitted to the space formed by the 1st and 2nd PCs. The intersection points of the fitted line and trajectories in the 2-dimensional space are then obtained. In Fig. 3, the whole process of Poincaré mapping is shown. The reason for deploying PCA is to project the entire phase space onto a different space (the space originated from 1st and 2nd PCs) where the reconstructed trajectories are more spread. In this way, the chosen space contains more information about the dynamics of the states. It is worth mentioning that the first 2 PCs are just linear transformations of the original variables and do not necessarily contain more information than the other PCs. Therefore, we empirically checked all the possible combinations of PCs i.e., 1st and 3rd PCs, 2nd and 3rd PCs, etc., and the best results was achieved by the first and seconds PCs for feature extraction. In order to find the intersection points, the trajectories and the Poincaré section lines were considered as polylines and then the Bézier clipping method [30] was applied.

D. Feature Extraction

In the next step, seven features in total were extracted from the first PC of the intersection sequence obtained in Section II.C. These features are as follows:

Range:

$$\text{Range} = \max(\mathrm{X}) - \min(\mathrm{X}), \tag{2}$$

Quantile and interquantile range:

The 0.13 quantile and the interquartile range (the difference between the first -0.25- and the third -0.75- quartile).

Shannon entropy:

$$\mathrm{H_s(X)} = -\sum \mathrm{P(X)}\log_2\big(\mathrm{P(X)}\big), \tag{3}$$

Root Mean Squared Amplitude (RMS Amp):

$$\mathrm{RMS_{Amp}(X)} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{k=1}}^{\mathrm{N}} \mathrm{X^2(k)}}, \tag{4}$$

Coefficient of Variation:

$$\mathrm{COV(X)} = \frac{\sqrt{\frac{\sum(\mathrm{X}-\bar{\mathrm{X}})}{\mathrm{N}}}}{\bar{\mathrm{X}}}, \tag{5}$$

and energy:

$$\mathrm{En(X)} = \sum_{\mathrm{k=1}}^{\mathrm{N}} |\mathrm{X(k)}|^2 \tag{6}$$

where X is the sequence of the intersection points, and $\bar{\mathrm{X}}$ is its mean value. N is the number of intersection points and P(X) is the probability distribution function.

E. Classification and Post Processing

Once the feature vectors of each patient were formed, they are fed into a two-layer classifier network. In the first layer, a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2505238, IEEE Transactions on Neural Systems and Rehabilitation Engineering

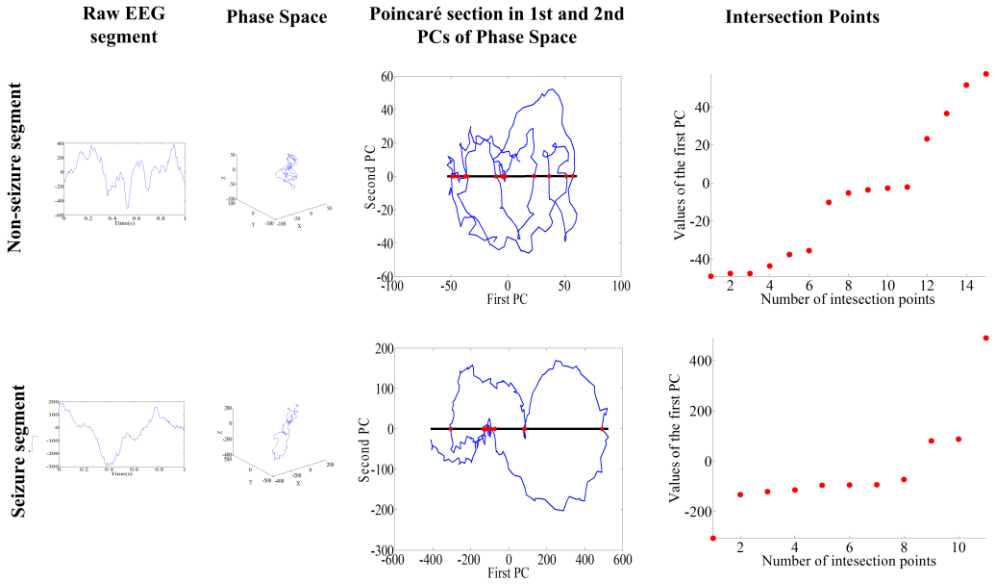

Fig. 3. Poincaré mapping procedure. The phase space (column 2) is obtained from the raw signal (column 1) in a non-seizure (top row) and a seizure segment (bottom row). In order to draw the Poincaré section, a polynomial curve with degree of one is fitted to the 1st and 2nd PCs of the phase space coordinates (column 3). Once the intersection points were determined, their values on the first PC were used for feature extraction. For visualization purpose, only the first three coordinates of phase space are shown.

single LDA classifier trained over the features of each 1-second EEG segment of each EEG channel. Once the first layer classifiers are trained, then their outputs (class vectors) are fed into a naïve Bayes (NB) classifier as a feature matrix, which makes the final classification for the input 1-second EEG segment. In the naïve Bayes classifier the "*multivariate multinomial distribution"* is used to model the outputs of the LDA classifiers because the input features of the second layer are discrete (i.e., binary). The schematic diagram of the proposed classification framework is shown in Fig. 4.

The main scenario is that once the system is trained for a specific patient it can then be used over and over for the same patient. In this case, after the system is trained using the labeled data, the system can be used and help the neurologist for the same patient over and over. Therefore, we divided the benchmark database is divided into a training and test datasets, both of which contain seizure and non-seizure frames, where the training set contains seizure and non-seizure segments which occurred earlier in time and the remaining segments constitute the test set. The classifier network was trained over the EEG recording of each patient's training set that is formed using two different training sizes: 25% and 50% of the available data.

In the post processing step, the fuzzy rule-based morphological filter proposed in our previous work [31] was applied to the outputs of each classifier in both layers (i.e., 23 LDA and 1 NB). The principal aim of the morphological filter is to filter out the classification outliers based on some global properties such as continuity and neighborhood similarity.

III. Experimental Results

In this section, first the overall results of the proposed patient-specific approach are presented. For comparative

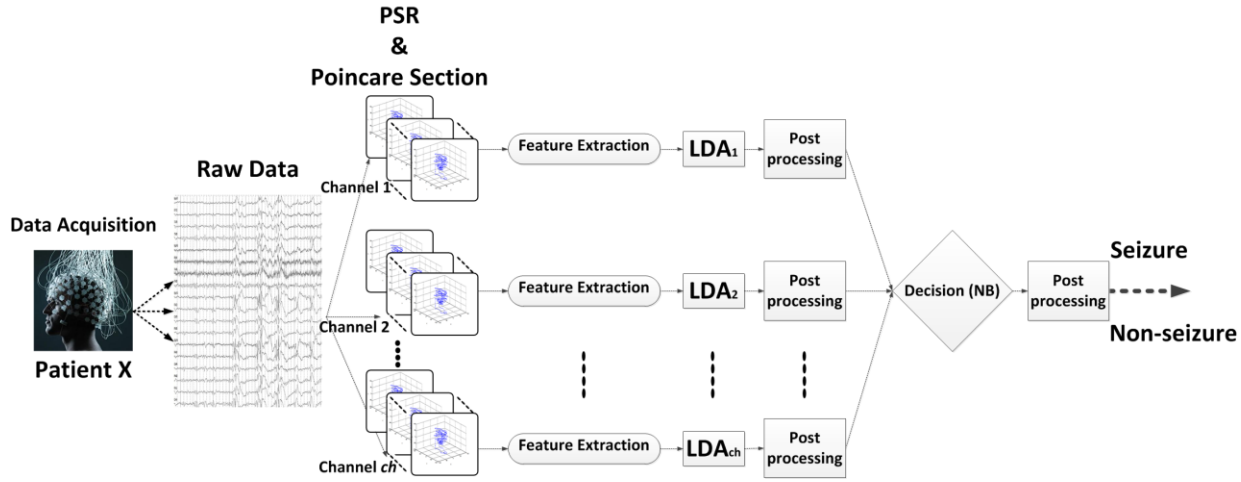

Fig. 4. The proposed classification framework (PSR is the phase-space reconstruction, LDA and NB are the linear discriminant analysis and naïve Bayes classifier, respectively).

evaluations, in section III.B, the seizure detection systems using the CHB-MIT benchmark are briefly presented. In section III.C, the proposed feature extraction approach is compared against the three state-of-the-art methods. In section III.D, four different classifiers within the proposed classification topology are evaluated against the proposed classifier. Finally, the computational complexity analysis is presented in section 0.

A. Classification Performance Evaluations

In this work, the standard performance measures of sensitivity (*Sen*), specificity (*Spe*) and accuracy (*Acc*) are used. They are defined as follows,

$$Sen = \frac{TP}{TP+FN}, \quad (7)$$

$$Spe = \frac{TN}{TN+FP}, \quad (8)$$

$$Acc = \frac{TP+TN}{TP+TN+FP+FN}, \quad (9)$$

where TP (True Positive) is the number of segments correctly detected as seizure, FN (False Negative) is the number of segments incorrectly detected as non-seizure, TN (True Negative) is the number of segments correctly detected as non-seizure, and FP (False Positive) is the number of segments incorrectly detected as seizure. The confusion matrix and the overall classification performance measures are shown in TABLE II and TABLE III. In addition, the Region of Convergence (ROC) plots are presented in Fig. 5 and Fig. 6 for better visualization of the performance of the proposed method.

As TABLE III shows, the best classification performance is achieved using a training rate of 50% with an average sensitivity and specificity of 89.10% and 94.80%, respectively. However, even with 25% training rate only an insignificant performance loss is encountered, i.e., 88.27% and 93.21% are the average sensitivity and specificity rate, respectively. This demonstrates a delicate generalization capability of the proposed approach and effectiveness of the proposed feature extraction on the discrimination of the seizures segments.

However, the results given in TABLE III indicate that a relatively low classification accuracy is obtained on a few patients i.e., 6, 12 and 24. The reason is that in the recordings of these patients, there are many similarities between seizure and non-seizure segments as well as the high variations within each type. This is visible in Fig. 7 where few segments of non-seizure and seizure recordings from patient 6 are displayed. As can be seen in the figure, there is a high variability between patterns of non-seizure segments c, *e* and *g*. In addition, this difference is evident between seizure segments *d* and *h*. Furthermore, both segments in *a* and *b* have high frequency and low amplitude signals while the former is a non-seizure and the latter is a seizure segment.

Furthermore we calculated the Average Detection Sensitivity Rate (ADSR), Average False Alarm (AFA) per hour and Average Alarm Delay (AAD) as expressed below in order to evaluate the "seizure event detection" performance of the proposed method. In order to calculate these metrics, we defined a seizure event if at least 7 consecutive seizure segments (with resolution of 1 second) are detected.

$$\text{ADSR} = \frac{\sum_{S=1}^{NS} number\ of\ detected\ seizure\ events\ in\ subject\ "S"}{Total\ number\ of\ seizure\ events\ in\ all\ subjects}, \quad (10)$$

TABLE II. True Positive (TP), False Negative (FN), True Negative (TN), and False Positive (FP) achieved for test set using 50% and 25% training rate.

	50% Training rate				25% Training rate			
patients	Seizure detected as seizure (TP)	Seizure detected as non-seizure (FN)	Non-seizure detected as non-seizure (TN)	Non-seizure detected as seizure (FP)	Seizure detected as seizure (TP)	Seizure detected as non-seizure (FN)	Non-seizure detected as non-seizure (TN)	Non-seizure detected as seizure (FP)
1	206	18	11693	45	327	9	17371	236
2	80	7	3955	37	131	0	5645	343
3	203	1	12087	308	306	0	17084	1509
4	169	22	18409	579	276	10	23663	4819
5	219	62	8677	41	350	72	13011	66
6	58	23	41672	4854	99	23	47232	22557
7	124	40	16069	35	211	35	23988	168
8	373	89	8522	16	574	119	12404	403
9	133	7	17084	25	206	4	24844	820
10	212	15	24750	255	293	47	37129	378
11	383	21	4521	104	597	9	6556	381
12	407	101	11344	5578	560	202	20491	4892
13	219	6	11751	624	325	12	17455	1107
14	83	5	11814	697	129	3	17258	1509
16	33	5	6681	2280	42	15	11859	1583
17	148	0	4862	402	191	31	7618	278
18	161	0	8460	1515	242	0	11953	3010
19	101	18	5048	105	135	44	7647	83
20	142	9	9790	77	213	13	14525	275
21	101	0	6727	66	151	1	10040	150
22	102	1	5278	18	151	4	7836	108
23	183	32	15757	154	193	130	23716	151
24	179	90	21142	188	272	132	31195	800

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2505238, IEEE Transactions on Neural Systems and Rehabilitation Engineering

TABLE III. The classification results using the proposed method. Patients with measures less than 70% are highlighted. Sen, Spe, and Acc are the sensitivity, specificity and accuracy obtained over the test data.

	50% Training (%)			25% Training (%)		
Patient	**Sen.**	**Spe.**	**Acc.**	**Sen.**	**Spe.**	**Acc.**
1	91.96	99.62	99.47	97.32	98.66	98.63
2	91.95	99.07	98.92	100	94.27	94.39
3	99.51	97.52	97.55	100	91.88	92.02
4	88.48	96.95	96.87	96.50	83.08	83.21
5	77.94	99.53	98.86	82.94	99.50	98.98
6	71.60	89.57	89.54	81.15	67.68	67.70
7	75.61	99.78	99.54	85.77	99.30	99.17
8	80.74	99.81	98.83	82.83	96.85	96.13
9	95.00	99.85	99.81	98.10	96.80	96.82
10	93.39	98.98	98.93	86.18	98.99	98.88
11	94.80	97.75	97.51	98.51	94.51	94.83
12	80.12	67.04	67.42	73.49	80.73	80.52
13	97.33	94.96	95.00	96.44	94.04	94.08
14	94.32	94.43	94.43	97.73	91.96	92.00
16	86.84	74.56	74.61	73.68	88.22	88.16
17	100	92.36	92.57	86.04	96.48	96.19
18	100	84.81	85.05	100	79.88	80.20
19	84.87	97.96	97.67	75.42	98.93	98.39
20	94.04	99.22	99.14	94.25	98.14	98.08
21	100	99.03	99.04	99.34	98.53	98.54
22	99.03	99.66	99.65	97.42	98.64	98.62
23	85.12	99.03	98.85	59.75	99.37	98.84
24	66.54	99.12	98.71	67.33	97.50	97.12
Average	**89.10**	**94.80**	**94.69**	**88.27**	**93.21**	**93.11**

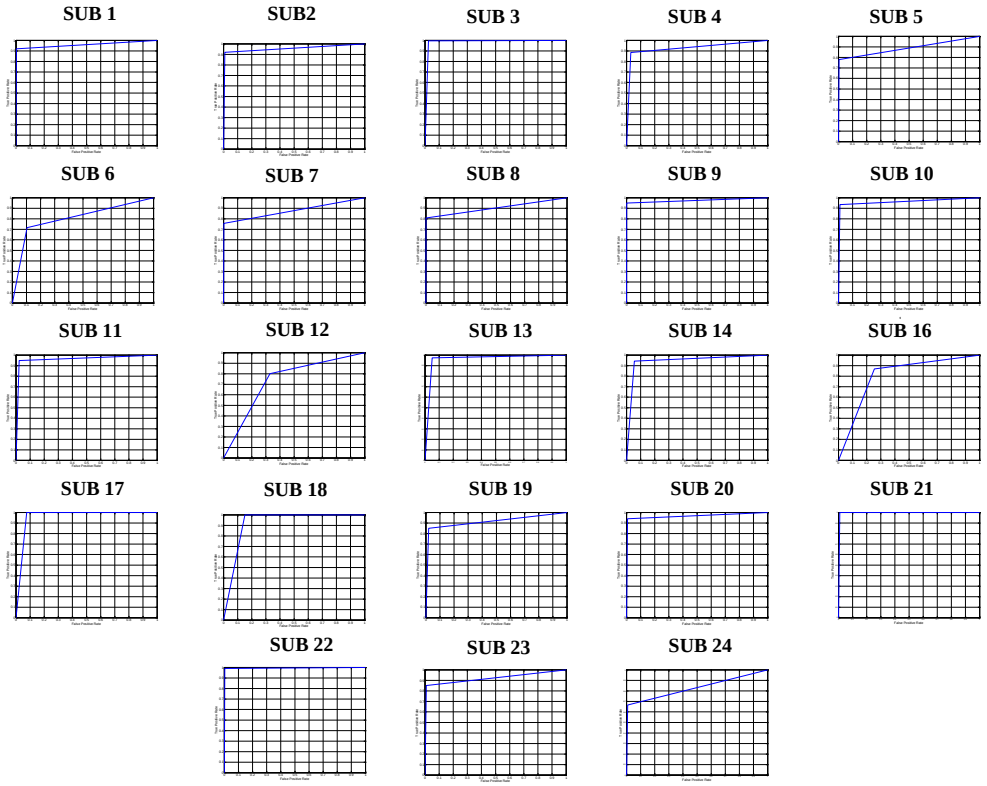

Fig. 5. ROC plots for 50% training rate per patient. The x- and y axis represent the false positive rate and true positive rate, respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2505238, IEEE Transactions on Neural Systems and Rehabilitation Engineering

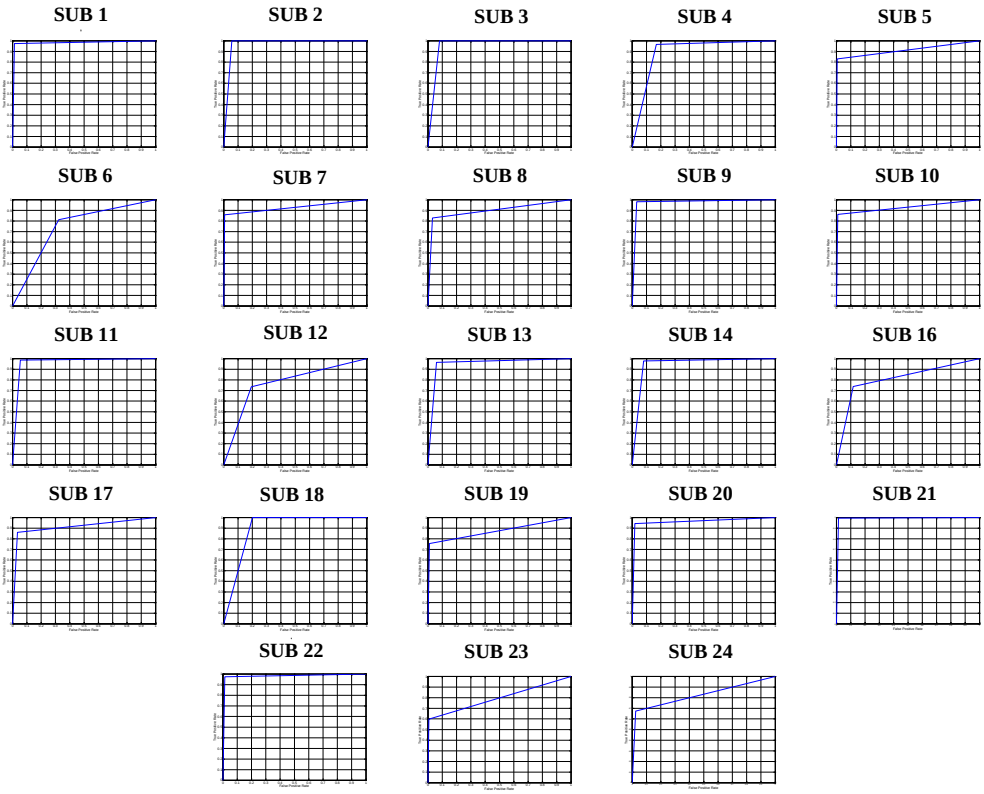

Fig. 6. ROC plots for 25% training rate per patient. The x- and y axis represent the false positive rate and true positive rate, respectively.

$$\text{AFA per hour} = \frac{\sum_{S=1}^{NS} number\ of\ false\ alarms\ in\ subject\ "S"}{Total\ duration\ of\ EEG\ recordings\ (in\ hours)\ in\ all\ subjects}, \quad (11)$$

$$\text{AAD} = \frac{\sum_{S=1}^{NS} length\ of\ delays\ in\ subject\ "S"}{Total\ number\ of\ detected\ seizure\ events\ in\ all\ subjects}, \quad (12)$$

where NS = 23 is the total number of patients. The ADSR, AFA, and AAD metrics are reported for the test set in TABLE IV.

TABLE IV. Average Detection Sensitivity Rate (ADSR), Average False Alarm (AFA) and Average Alarm Delay (AAD) achieved on test set for seizure event detection

	Using 50% training rate	Using 25% training rate
ADSR (%)	96.29	91.34
AFA per hour	3.04	4.86
AAD (second)	4.65	5.03

As can be seen in TABLE IV, the best results have been achieved using 50% training rate as expected. Because of the existence of noise and artifacts in the EEG records of the patients 6 and 12, the average of false alarms increased significantly. To be more specific, using 50% training rate and not taking into account patients 6 and 12, we achieved 97.14%, 1.85, and 5.63 for ADSR, AFA and AAD, respectively. However, the delay is increased by about 1 second on average while the number of false alarms per hour is approximately reduced by half. Similarly, with 25% training rate, 94.93%, 2.75 and 5.65 were obtained for ADSR, AFA and AAD, respectively.

B. Comparative Evaluations of the Classification Performance

Few recent studies have used this benchmark for evaluation. There are three main reasons for this: 1) high seizure variations both within- and among patients, 2) only bipolar longitudinal montage information provided (lack of full montage information), and 3) long-term EEG recordings in this dataset contain other patterns such as sleep and physiological artefacts which reduce the performance of seizure detection (in particular in patients 6 and 12).

In TABLE V, we summarized the seizure detection methods (i.e., methods which detect the entire duration of seizure events and not only the onset of seizure) applied on CHB-MIT dataset. In order to perform a fair comparative evaluation we compared our approach only with those studies which used a training rate higher than or equal to ours and also used complete data from the benchmark database for evaluation.

Note that no comparisons can be drawn between our results and the ones from [32], [33] and [34], since the patients used in these studies were not specified. In [35], only accuracy is reported which is not a proper metric for such highly unbalanced dataset. Still in the proposed approach with only 25% training rate we achieved around 13% higher accuracy level than the method in [35] although they used 80% training rate.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2505238, IEEE Transactions on Neural Systems and Rehabilitation Engineering

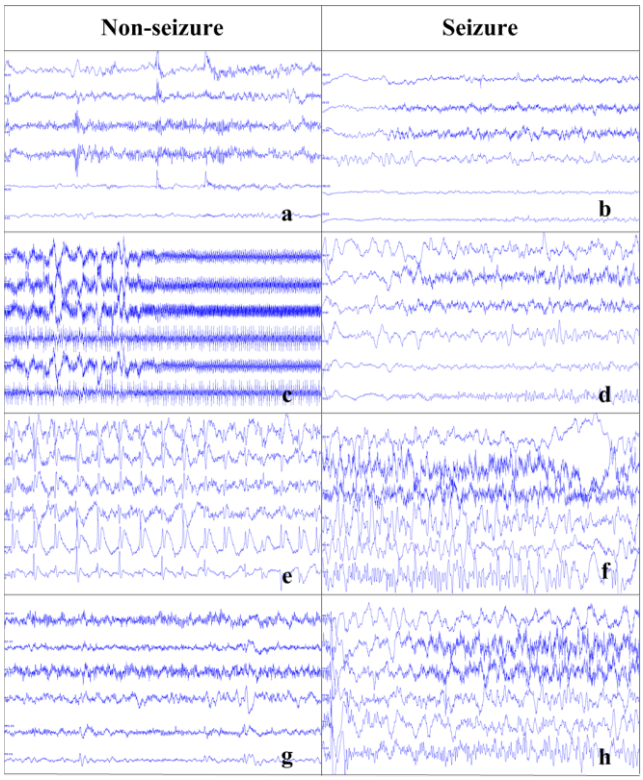

Fig. 7. The non-seizure (first column) and seizure (second column) segments of patient 6 illustrate a high variability within the same class and similarity between distinct classes (length of each segment=10 seconds, channels: FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, and F3-C3).

Direct comparison with [36] is not feasible either because there is an ambiguity in how the training and test sets are constructed. Regardless of this ambiguity, the best achieved detection sensitivity rate of 83% was reported in the fifth experiment which corresponds to the 80% training rate. Regardless, we still achieve a significantly higher sensitivity rate with only 25% training rate.

Additionally, in this work an average sensitivity and specificity rate of 89.31% and 95.03% were achieved using only the *early* 25% of each EEG record as training (excluding the patients 6 and 12), which further shows approximately 1% improvement in contrast to our previous work [31] with the same training rate but using a large set of features in a significantly sophisticated classifier network topology. Therefore, our previous method needs a cloud computing implementation with massive parallelization for any on-line processing because we used a large network of classifiers in each Network of Binary Classifier (NBC) [37]. While in the proposed method, only 24 and significantly faster classifiers (23 LDAs and 1 Naïve Bayes) are used. In addition, the proposed method needs only about 3 ms for feature extraction from a 1-s EEG segment while in our previous method it took 280 ms. As a result, the proposed method can easily be ported on a tablet or pocket size computer which makes the proposed EEG classification approach feasible on a mobile application.

C. Comparative Evaluations of the Feature Extraction Approach

In order to perform a fair comparison between the discriminative powers of the proposed feature extraction technique against the other state-of-the-art methods, three different feature sets are extracted from the CHB-MIT benchmark and classified using the same classification scheme explained in section II.E. The first feature set consists of six features used in [38], which are the energy of details and approximation coefficients (d_1, d_2, d_3, d_4, d_5, and a_5). The second feature set was proposed by Kumar et al. [39], where approximate entropy (ApEn) of details and approximation coefficients (d_1-d_5, and a_1-a_5) were proposed for epileptic seizure detection. In these two feature sets, wavelet Daubechies 4 were used in order to decompose the EEG segments into five levels. The third feature set is a combination of eight nonlinear features, including ApEn [40], correlation dimension [41], and recurrence quantification analysis (RQA). The RQA based features are recurrence rate, determinism, averaged diagonal line length, entropy, laminarity, and trapping time [28]. TABLE VI presents the obtained results against the three competitor feature sets over the test data.

The results clearly indicate that the proposed feature extraction approach yields the highest average sensitivity and specificity rates. Based on TABLE VI, the proposed features obtained relatively low results for only two patients, i.e., less than 70% in either sensitivity or specificity while the competing feature sets totally fail in some patients; for instance, the first, second and third feature sets obtained less than 50% specificity for patients 4, 13, and 18. This makes them entirely unreliable in practice for a medical diagnosis application using the proposed classification scheme.

D. Comparative Evaluations of the Classifiers

To evaluate the selection of the classifiers used in the first layer of the proposed approach (i.e., LDA), the performances of different classifiers are also presented while the NB classifier in the second layer is used. In TABLE VII, we compared the sensitivity and specificity of 4 state-of-the-art classifiers including Quadratic Discriminant Analysis (QDA), Mahalanobis Discriminant Analysis (MDA), Naïve Bayes (NB), and SVM (with a linear kernel). The NB classifier used in the first layer employed *kernel smoothing density estimation* in order to model the input features. It is clear that the proposed LDA classifiers can reach much higher sensitivity compared to QDA and NB classifiers. In addition, they can achieve approximately 2% higher average sensitivity and specificity compared to MDA classifier. Even though SVM classifier with linear kernel achieved 1% higher sensitivity than LDA, it fails to classify the EEG data of patients 4 and 6. Moreover, LDA achieved 0.4% higher specificity on average, which is a noteworthy improvement in such unbalanced data, i.e., recall that the average duration of non-seizure segments over 23 patients is around 26444 seconds, 0.4% of which are about 106 seconds. So on average, using SVM would lead to an extra of 106 false alarms. Furthermore, the computational complexity of the SVM classifier is much higher than that of LDA, which may cause problems in practice for real-time processing. This is further investigated in detail in the next sub-section.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2505238, IEEE Transactions on Neural Systems and Rehabilitation Engineering

TABLE V. The proposed seizure detection methods using CHB-MIT dataset

Authors	Features	Patients	Channels	Training Rate (%)	Av. Sen. (%)	Av. Spe. (%)	Av. Acc. (%)
Rafiuddin et. al. [35]	Interquartile range, median absolute deviation of raw data, energy and coefficient of variation extracted from the Daubiches (db4) wavelet coefficients	23	23	80	-	-	80.16
Uzzaman Khan et. al. [32]	Relative values of normalized coefficient of variation (NCOV) based measure	5 (not specified)	Not specified	80	83.6	100	91.8
Hunyadi et. al. [36]	16 features extracted from time and frequency domain	23	Not specified	≈ 80	83	-	-
Supratak et. al. [33]	Uses stacked autoencoders as unsupervised feature learner	6 (randomly selected)	Not specified-channels were selected manually	Totally 30 epochs used	100	-	-
Fürbass et al. [34]	EpiScan (automatic seizure detection method)	23	-	-	67	-	-
Kiranyaz et. al. [31]	342 features including time, frequency, and time-frequency features	21(excluding patients 6, 12, and 15)	18	25	89.01	94.71	-
Proposed Method	7 features extracted from intersection points of Poincaré section and phase space	23 (excluding patient 15)	23	25	88.27	93.21	93.11
				50	89.10	94.80	94.69

TABLE VI. The classification results of three different feature sets using the proposed classification scheme. Performance metrics less than 70% are highlighted.

	Length of One EEG Segment = 1 second, Training Rate = 50%							
	Proposed feature set		**1st Feature Set Energy of DWT Coeff. [38]**		**2nd Feature Set DWT-based ApEn. [39]**		**3rd Feature Set Nonlinear [40] [41] [28]**	
Patient	**Sen.**	**Spe.**	**Sen.**	**Spe.**	**Sen.**	**Spe.**	**Sen.**	**Spe.**
1	91.96	99.62	94.64	98.11	94.64	99.48	93.75	90.45
2	91.95	99.07	98.85	16.51	96.55	36.62	78.16	95.57
3	99.51	97.52	96.08	97.01	95.59	97.52	94.61	97.04
4	88.48	96.95	24.61	26.01	21.99	34.78	67.02	83.96
5	77.94	99.53	79.36	98.65	78.29	99.60	98.93	85.66
6	71.60	89.57	65.43	67.99	64.20	73.44	64.20	70.28
7	75.61	99.78	73.17	99.79	69.51	99.96	50.00	93.45
8	80.74	99.81	74.89	99.23	74.46	99.40	77.71	96.78
9	95.00	99.85	82.14	100	87.14	100	68.57	66.46
10	93.39	98.98	84.58	99.86	83.26	99.64	94.27	91.89
11	94.80	97.75	95.30	97.69	94.31	94.14	99.01	74.68
12	80.12	67.04	76.97	79.66	87.80	87.24	77.76	65.88
13	97.33	94.96	100	28.75	98.22	32.21	99.11	42.97
14	94.32	94.43	64.77	98.63	76.14	98.63	85.23	99.22
16	86.84	74.56	81.58	91.75	84.21	97.51	97.37	76.97
17	100	92.36	92.57	93.43	97.30	93.79	91.22	74.58
18	100	84.81	97.52	78.42	96.27	82.52	100	43.19
19	84.87	97.96	84.87	95.85	85.71	99.28	77.31	91.97
20	94.04	99.22	92.05	97.67	94.04	97.14	94.04	95.46
21	100	99.03	90.10	98.45	90.10	98.45	90.10	84.31
22	99.03	99.66	96.12	99.30	97.09	98.51	83.50	91.77
23	85.12	99.03	93.02	99.39	88.84	99.08	90.23	92.76
24	66.54	99.12	65.80	97.05	63.94	99.99	76.95	93.84
Average	**89.10**	**94.80**	**82.80**	**85.18**	**83.46**	**87.78**	**84.74**	**82.57**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2015.2505238, IEEE Transactions on Neural Systems and Rehabilitation Engineering

TABLE VII. The classification results of the proposed features using four different classifiers. Patients with performance metrics of less than 70% are highlighted.

	Length of One EEG Segment = 1 second, Training Rate = 50%									
	LDA		**QDA**		**MDA**		**NB**		**SVM**	
Patient	**Sen.**	**Spe.**	**Sen.**	**Spe.**	**Sen.**	**Spe.**	**Sen.**	**Spe.**	**Sen.**	**Spe.**
1	91.96	99.62	95.98	99.45	100	91.00	95.87	98.93	95.98	99.80
2	91.95	99.07	98.85	13.45	90.80	99.97	93.98	98.76	97.70	98.95
3	99.51	97.52	99.51	95.35	100	97.05	100	96.51	99.51	97.10
4	88.48	96.95	23.04	40.85	18.32	97.70	11.11	97.52	32.46	96.00
5	77.94	99.53	86.83	93.90	100	67.80	96.59	93.62	99.64	94.23
6	71.60	89.57	58.02	29.72	18.75	98.10	16.05	77.51	55.56	91.24
7	75.61	99.78	82.32	96.35	91.46	94.98	87.90	96.41	89.02	97.44
8	80.74	99.81	88.74	99.33	93.51	97.21	83.63	99.30	89.83	99.06
9	95.00	99.85	97.14	96.05	97.14	95.69	97.84	76.14	96.43	97.21
10	93.39	98.98	96.92	96.70	96.04	94.55	95.96	98.20	95.59	99.08
11	94.80	97.75	96.53	95.35	97.03	87.87	96.47	98.67	96.04	99.11
12	80.12	67.04	59.76	84.85	86.22	68.38	53.66	77.17	87.40	74.74
13	97.33	94.96	56.89	94.56	96.41	95.31	72.85	97.50	96.44	95.01
14	94.32	94.43	88.64	91.99	88.64	96.92	91.67	95.38	94.32	95.72
16	86.84	74.56	84.21	41.95	76.32	97.07	63.64	79.22	92.11	72.05
17	100	92.36	100	92.00	100	87.08	100	92.97	100	90.81
18	100	84.81	100	77.30	100	75.26	100	85.81	100	82.09
19	84.87	97.96	88.14	93.13	97.48	87.70	82.18	94.83	95.80	96.35
20	94.04	99.22	90.73	99.04	93.38	96.73	87.50	99.13	94.04	98.62
21	100	99.03	85.15	99.85	86.14	99.87	87.75	99.63	100	99.44
22	99.03	99.66	100	97.90	100	96.85	100	98.33	100	98.92
23	85.12	99.03	97.67	99.13	98.60	97.94	98.07	99.06	98.14	99.06
24	66.54	99.12	72.12	98.48	80.67	95.49	68.83	99.53	66.54	99.50
Average	**89.10**	**94.80**	**84.66**	**83.77**	**87.26**	**92.02**	**81.81**	**93.48**	**90.11**	**94.41**

E. Computational Complexity Analysis

In this work, we implemented the proposed method using MATLAB version R2014b and the computations were performed on a standard desktop computer with a 3.4 GHz processor and 16 GB ram. Two metrics are used for evaluating the computational complexity: The first metric is the time taken to extract each feature set from every 1-s long segment of one EEG channel. This value was repeatedly calculated across one hour (3600 segments) to ensure the repeatability of the measured time and is reported in TABLE VIII for the proposed and competing feature extraction methods.

TABLE VIII. Run time (in millisecond) using three different feature extraction methods in the proposed classification scheme

	Proposed Feature Set	**1st Feature Set**	**2nd Feature Set**	**3rd Feature Set**
Average elapsed time for feature extraction of 1s-long EEG segment	2.6	2.8	5.3	146.2

The second metric is the elapsed times for classification (including training and test) over the proposed features for one hour recording with different classifiers, which are shown in TABLE IX. The results clearly demonstrate the superiority of the proposed approach with LDA classifiers over the proposed features in terms of computational complexity.

Clearly, the proposed method has the advantages of having relatively high sensitivity and specificity, and low computational complexity. The proposed method on the other hand lacks preprocessing steps for removing noise and artifacts. Although, this decreases the computational complexity, it can increase the false alarms due to the contaminated EEG signals especially in patients 6 and 12. Additionally, the proposed method is not able to detect the pre-ictal states.

TABLE IX. Classification run times (in seconds) using four different classifiers over the proposed features

	LDA (Proposed)	**QDA**	**MDA**	**NB**	**SVM**
Average elapsed time for classification of 1h EEG recording	0.25	0.27	0.26	58.29	24.28

IV. CONCLUSIONS

In this study, a new multi-channel EEG seizure detection method is presented based on the dynamics of the trajectories in phase space. The proposed Poincaré mapping procedure enables us to study the difference between the dynamics of the seizure and non-seizure segments in a high dimensional phase space. The proposed approach keeps the reconstructed

trajectories unfolded and the computational complexity low.

The proposed seizure detection approach was performed over CHB-MIT database in order to achieve the aforementioned objectives. The results indicate an improved classification performance over competing techniques without any pre-processing. The proposed approach achieves the highest accuracy and minimum false alarm rate among the three state-of-the-art feature extraction methods and four different classifiers, and offers the best trade-off between the anomaly detection accuracy and computational burden. Furthermore, the achieved run time shows the potential application of the proposed approach in Epilepsy Monitoring Units (EMUs).

Extracting better features directly from the phase space rather than from its principal components is the subject of our future study. To accomplish this, noise and artifacts in EEG signals will need to be significantly suppressed prior to subsequent analysis.

ACKNOWLEDGMENT

The authors gratefully acknowledge Professor Tarmo Lipping from Tampere University of Technology in Finland for his fruitful comments on the proposed method. We would also like to thank Professor Ville Jäntti from the Department of Clinical Neurophysiology, Seinäjoki Central Hospital in Finland for his helpful assistance in the interpretation of EEG signals.

REFERENCES

[1] W. H. Organization, "Fact sheet N°999," October 2012. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs999/en/.

[2] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG analysis of epilepsy: A review," *Knowledge-Based Systems*, vol. 45, pp. 147-165, 2013.

[3] J. Zhang, J. Zou, M. Wang, L. Chen, C. Wang, and G. Wang, "Automatic detection of interictal epileptiform discharges based on time-series sequence merging method," *Neurocomputing*, vol. 110, pp. 35-43, 2013.

[4] O. Blanke, G. Lantz, M. Seeck, L. Spinelli, R. Grave de Peralta, G. Thut, T. Landis, and C.M. Michel, "Temporal and spatial determination of EEG-seizure onset in the frequency domain," *Clinical Neurophysiology*, vol. 111, no. 5, pp. 763-772, 2000.

[5] B. Boashash, and GH. Azemi , "A review of time–frequency matched filter design with application to seizure detection in multichannel newborn EEG," *Digital Signal Processing*, vol. 28, pp. 28-38, 2014.

[6] C. Stam, "Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field," *Clinical Neurophysiology*, vol. 116, no. 10, p. 2266–2301, 2005.

[7] V. Bajaj and R. B. Pachori, "Epileptic seizure detection based on the instantaneous area of analytic intrinsic mode functions of EEG signals," *Biomedical Engineering Letters*, vol. 3, no. 1, pp. 17-21, 2013.

[8] R. Pachori, "Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition," *Research Letters in Signal Processing*, vol. 2008, 2008.

[9] P. Gifani, H.R. Rabiee, M.R. Hashemi, and M. Ghanbari,, "Dimensional Characterization of anesthesia dynamic in reconstructed embedding space," in *Engineering in Medicine and Biology Society (EMBS 2007).*, Lyon, 2007.

[10] Y. Fang, M. Chenb, and X. Zhengc, "Extracting features from phase space of EEG signals in brain-computer interfaces," *Neurocomputing*, p. In Press, 2014.

[11] N. Sriraam, "Correlation dimension based lossless compression of EEG signals," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 379-388, 2012.

[12] *F. Shayegh, S. Sadri, R. Amirfattahi, and K. Ansari-Asl, "A model-based method for computation of correlation dimension, Lyapunov exponents and synchronization from depth-EEG signals," Computer Methods and Programs in Biomedicine, vol. 113, no. 1, pp. 323-337, 2014.*

[13] *C. Allefeld, and J. Kurths, "An approach to multivariate phase synchronization analysis and its application to event-related potentials," International Journal of Bifurcation Chaos, vol. 14, pp. 417-426, 2004.*

[14] *X. Meng, J. Xu, and F. Gu, "Generalized dimension of the intersection between EEGs," Biological Cybernetics, vol. 85, no. 4, p. 313, 2001.*

[15] C. J. Stam, Nonlinear Brain Dynamics, Nova Publishers, 2006.

[16] D. Kugiumtzis, "Surrogate Data Test on Time Series," in *Modelling and Forecasting Financial Data*, Springer Science & Business Media, 2002, pp. 267-282.

[17] R. Sharma , and R. B. Pachori, "Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1106-1117, 2015.

[18] M. Chen, Y. Fang, and X. Zheng, "Phase space reconstruction for improving the classification of single trial EEG," *Biomedical Signal Processing and Control*, vol. 11, pp. 10-16, 2014.

[19] S. Lee, J. S. Lim, J. Kim, J. Yang, and Y. Lee, "Classification of normal and epileptic seizure EEG signals using wavelet transform, phase-space reconstruction, and Euclidean distance," *Computer Methods and Programs in Biomedicine*, vol. 116, no. 1, pp. 10-25, 2014.

[20] A. Brignol, T. Al-ani b, and X. Drouot, "Phase space and power spectral approaches for EEG-based automatic sleep–wake classification in humans: A comparative study using short and standard epoch lengths," *Computer Methods and Programs in Biomedicine*, vol. 109, no. 3, pp. 227-238, 2013.

[21] J. Jeong, D. Kim, J. Chae, S. Y. Kim, H. Ko, and I. Paik, "Nonlinear analysis of the EEG of schizophrenics with optimal embedding dimension," *Medical Engineering & Physics*, vol. 20, no. 9, pp. 669-679, 1998.

[22] N. Kannathal, U. R. Acharya, C.M. Lim, and P.K. Sadasivan, "Characterization of EEG—A comparative study," *Computer Methods and Programs in Biomedicine*, vol. 80, no. 1, pp. 17-23, 2005.

[23] Y. Yuan,Y. Li, and DP. Mandic, "A comparison analysis of embedding dimensions between normal and epileptic EEG time series.," *Journal of Physiological Sciences*, vol. 58, no. 4, pp. 239-247, 2008.

[24] C. Anderson, E. Forney, D. Hains, and A. Natarajan, "Reliable identification of mental tasks using time-embedded EEG and sequential evidence accumulation," *Journal of Neural Engineering*, vol. 8, no. 2, 2011.

[25] *A. L. Goldberger, L. Amaral , L. Glass, J. M. Hausdorff , P. C. Ivanov, R. G. Mark, J. E. Mietus, G.B. Moody, C.K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," Circulation, vol. 101, no. 23, p. e215–e220, 2000.*

[26] A. Shoeb, "Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment," Ph.D. dissertation, Massachusetts Institute of Technology, September 2009.

[27] F. Takens, "Detecting strange attractors in turbulence," *Lecture Notes in Mathematics*, vol. 898, pp. 366-381, 1981.

[28] C. L. Webber, J. P. Zbilut, "Recurrence Quantification Analysis of Nonlinear Dynamical Systems," in *Tutorials in Contemporary Nonlinear Methods for Behavioral Sciences*, Retrieved March 6th from http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp, 2005, pp. 27-94.

[29] A. Lyon, and M. Colyvan , "The Explanatory Power of Phase Spaces," *Philosophia Mathematica*, vol. 2, no. 16, pp. 227-243, 2007.

[30] T.W. Sederberg, and T. Nishita, "Curve intersection using Bézier clipping," *Computer-Aided Design*, vol. 22, no. 9, pp. 538-549, 1990.

[31] S. Kiranyaz, T. Ince, M. Zabihi, and D. Ince, "Automated patient-specific classification of long-term Electroencephalography," *Journal of Biomedical Informatics*, vol. 49, pp. 16-31, 2014.

[32] Y. Uzzaman Khan, N. Rafiuddin, and O. Farooq, "Automated seizure detection in scalp EEG using multiple wavelet scales," in *Signal Processing, Computing and Control (ISPCC), 2012 IEEE International Conference on*, Waknaghat Solan, 2012.

[33] A. Supratak, L. Li, and Y. Guo, "Feature Extraction with Stacked Autoencoders for Epileptic Seizure Detection," in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Chicago, IL, 2014 .

[34] F. Fürbass, P. Ossenblok, M. Hartmann, H. Perko, A.M. Skupch, G. Lindinger, L. Elezi, E. Pataraia, A.J. Colon, C. Baumgartner, and T. Kluge, "Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units," *Clinical Neurophysiology*, p. In Press, 2014.

[35] N. Rafiuddin, Y. Uzzaman Khan, and O. Farooq, "Feature Extraction and Classification of EEG for Automatic Seizure Detection," in *International Conference on Multimedia, Signal Processing and Communication Technologies*, Aligarh, 2011.

[36] B. Hunyadi, M. Signoretto, W. Van Paesschen, J. A.K. Suykens, S. Van Huffel, and M. De Vos, "Incorporating structural information from the multichannel EEG improves patient-specific seizure detection," *Clinical Neurophysiology*, vol. 123, no. 12, p. 2352–2361, 2012.

[37] S. Kiranyaz, T. Ince, and M. Gabbouj, "The Classifier Framework: Collective Network of Binary Classifiers," in Multidimensional Particle Swarm Optimization for Machine Learning and Pattern Recognition, Springer Science & Business Media, 2013, pp. 265 - 270.

[38] N. Sadati, H. R. Mohseni, and A. Maghsoudi , "Epileptic Seizure Detection Using Neural Fuzzy Networks," in *IEEE International Conference on Fuzzy Systems*, Vancouver, 2006.

[39] *Y. Kumar, M. L. Dewal, and R. S. Anand, "Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network," Signal, Image and Video Processing, vol. 8, no. 7, pp. 1323-1334, 2014.*

[40] N. Kannathal, M. L. Choo, U. R. Acharya, and P.K. Sadasivan, "Entropies for detection of epilepsy in EEG," *Computer Methods and Programs in Biomedicine*, vol. 80, no. 3, pp. 187-194, 2005.

[41] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D: Nonlinear Phenomena*, vol. 9, no. 1-2, pp. 189-208, 1983.