Epileptic Seizure Prediction Using Lyapunov Exponents and Support Vector Machine

Conference Paper · July 2007

DOI: 10.1007/978-3-540-71629-7_42

4 authors:


Bartosz Swiderski
Warsaw University of Technology


Stanislaw Osowski
Warsaw University of Technology, Military University of…


Andrzej Cichocki
Skolkovo Institute of Science and Technology


Andrzej Rysz
Medical University of Warsaw

Epileptic Seizure Prediction Using Lyapunov Exponents and Support Vector Machine

Bartosz Świderski[1], Stanisław Osowski[1,2], Andrzej Cichocki[1,3], and Andrzej Rysz[4]

[1] Warsaw University of Technology, Warsaw, Poland
Koszykowa 75, 00-662 Warsaw, Poland
{sto, markiewt}@iem.pw.edu.pl
[2] Military University of Technology, Warsaw, Poland
[3] Brain Science Institute, RIKEN, Japan
[4] Banach Hospital, Warsaw, Poland

Abstract. The paper presents the method of predicting the epileptic seizure on the basis of EEG waveform analysis. The Support Vector Machine and the largest Lyapunov exponent characterization of EEG segments are employed to predict the incoming seizure. The results of numerical experiments will be presented and discussed.

1 Introduction

Epilepsy is a group of brain disorders characterized by the recurrent paroxysmal electrical discharges of the cerebral cortex, that result in irregular disturbances of the brain functions [3]. The spatio-temporal dynamical changes in EEG, beginning several minutes before and ending several minutes after the seizure, evolve in a characteristic pattern, culminating in a seizure.

It has been shown [2] that EEGs do not merely reflect stochastic processes, and instead that they manifest deterministic chaos. Using the EEG recording of a human epileptic seizure the papers [1], [5], [6] have shown the existence of a chaotic attractor, being the consequence of the deterministic nature of the brain activity.

In this paper we will investigate the phenomena of the EEG internal dynamics, characteristic for the epileptic activity, through the use of the short-term largest Lyapunov exponent (STLmax) and Support Vector Machine (SVM). The Lyapunov exponents and the other measures based on them will serve as the features used by the SVM to predict the incoming seizure. This approach is motivated by the special properties observed in EEG waveforms in the baseline (interictal) and epileptic states.

The typical EEG waveforms corresponding to three stages of an epileptic seizure: interictal, preictal and ictal are shown in Fig. 1. The ictal stage corresponds to the seizure period. The preictal stage is the time period preceding directly the seizure onset. No strict borders of the preictal period are defined. The postictal stage follows the seizure end. The interictal stage is the period between the postictal stage of one seizure and the preictal stage of the next seizure. In this state the EEG signal is chaotic of very unpredictable nature and relatively small magnitude. Observing the

B. Beliczynski et al. (Eds.): ICANNGA 2007, Part II, LNCS 4432, pp. 373 – 381, 2007.

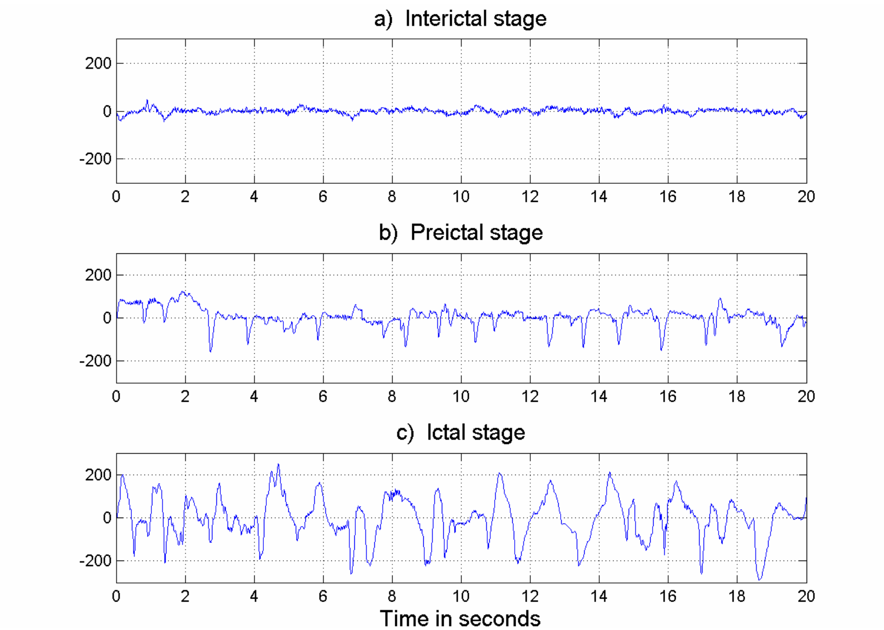

Fig. 1. The typical EEG waveforms corresponding to epilepsy: a) interictal, b) preictal and c) ictal stage (*horizontal axis* - seconds, *vertical axis* – microvolts)

EEG signals in all stages we can understand the difficulties in recognizing the seizure onset. The differences between the preictal and ictal stages are hardly noticeable for non-specialist.

As we approach the epileptic seizure the signals are less chaotic and take more regular shape. Moreover observing many channels at the same time we can see that the signals of all channels are becoming coupled and to some degree locked. The simultaneous synchronization of signals of many channels proves the decreasing chaoticity of signals. Thus the chaoticity measures of the actual EEG signal may provide good features for the prediction of the incoming seizure.

2 Largest Lyapunov Exponents for Characterization of EEG

Each EEG signal recorded from any site of the brain may be associated with the so called embedding phase space. If we have a data segment x(t) of duration T we can define the vector $\mathbf{x}_i$ in the phase space as following [5]

$$\mathbf{x}_i = \left[x(t_i),\ x(t_i+\tau),\ ...,\ x(t_i+(p-1)\tau)) \right] \tag{1}$$

with τ - the selected time lag between the components of each vector, p - the selected dimension of the embedding space and t_i - the time instant within the considered period [T-(p-1)τ]. Each vector $\mathbf{x}_i$ in the phase space represents an instantaneous state of the system.

The Lyapunov exponents have been proven to be useful dynamic diagnostic measure for EEG waveforms [1], [2], [4], [10]. Lyapunov exponents define the average exponential rate of the divergence or convergence of the nearby orbits in the phase space. It may be described in the form $d(t) = d_0 e^{Lt}$, where L means the Lyapunov exponent and d_0 – the initial distance of two nearby orbits. The magnitude of the exponent reflects the time scale on which the system dynamics become unpredictable [1],[5]. Any system containing at least one positive Lyapunov exponent is defined to be chaotic and the nearby points, no matter how close, will diverge to any arbitrary separation.

The estimation L of the $STL_{\max}$ exponent may be presented as following [1], [4]

$$L = \frac{1}{N\Delta t}\sum_{i=1}^{N} \log_2 \frac{\left|\Delta \mathbf{x}_{ij}(\Delta t)\right|}{\left|\Delta \mathbf{x}_{ij}(0)\right|}, \tag{2}$$

where $\Delta\mathbf{x}_{ij}(0) = \mathbf{x}(t_i)\text{-}\mathbf{x}(t_j)$ is the displacement vector at the time point t_i, that is the perturbation of the fiducial orbit observed at t_j with respect to t_i, while $\Delta\mathbf{x}_{ij}(\Delta t) = \mathbf{x}(t_i + \Delta t) - \mathbf{x}(t_j\text{-}\Delta t)$ is the same vector after time Δt. The vector $\mathbf{x}(t_i)$ is the point in the fiducial trajectory for $t = t_i$ and $\mathbf{x}(t_j)$ is a properly chosen vector adjacent to $\mathbf{x}(t_i)$ in the phase space. The time increase Δt is the evolution time, that is the time which is allowed for $\Delta\mathbf{x}_{ij}$ to evolve in the phase space. For the time given in sec the value of L is in bits/sec. N is the number of local $STL_{\max}$'s that will be estimated within period T of the data segment, where $T = N\Delta t+(p\text{-}1)\tau$. For the exponent L to be a good estimate of $STL_{\max}$ the candidate vector $\mathbf{x}(t_j)$ should be chosen in such a way that the previously evolved displacement vector $\Delta\mathbf{x}_{i\text{-}1,j}(\Delta t)$ is almost parallel to the candidate displacement vector $\Delta\mathbf{x}_{ij}(0)$. Moreover $\Delta\mathbf{x}_{ij}(0)$ should be small in magnitude to avoid computer overflow in the evolution within chaotic region.

In this work we have applied special method of choosing two time segments of EEG taking part in estimation of the largest exponent [12]. It is based on the statistical hypothesis that two chosen time segments are alike to each other. Each time segment is characterized by the vector $\mathbf{x}$ described by (1) of the length equal p=8. The length of the vector was chosen in such a way that the time series within this time span may be regarded as a stationary and at the same time the series is long enough to provide stabilization of the estimate to obtain the credible results of the performed calculations. We have applied the Kolmogorov-Smirnov test [7]. Let us denote the set of points belonging to the observed fiducial EEG trajectory corresponding to time t_i by $\mathbf{x}_i$ as shown by equation (1). By $\mathbf{x}_j$ we denote the vector generated from the same EEG waveform placed at t_j not far from the time point t_i. Similarity of these two vectors in the Kolmogorov-Smirnov test is characterized by the minimal significance level h needed to reject the null-hypothesis H0 that these vectors are drawn from the same underlying continuous population, that is belong to the same distribution (at the assumed dimension p of the vectors). Let H1 be the alternative hypothesis. On the basis of the observed actual statistics of the Kolmogorov-Smirnov test we determine the minimal significance level h needed to reject the null hypothesis. High value of the actual significance level h means high similarity of the processes, characterized by the vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ (small distance between both vectors). For the highest possible value of h they contribute to the summation terms in the equation (2).

For each data segment of the length T we choose the vectors $\mathbf{x}_i$ (i=1, 2, 3...) and each of these vectors is compared with many vectors $\mathbf{x}_j$ formed from the data belonging to the same segment. The vectors $\mathbf{x}_j$ are generated at the initial points t_j placed at different distances, starting from the distance larger than three times the length of the vector. For each $\mathbf{x}_i$ many pairs ($\mathbf{x}_i$, $\mathbf{x}_j$) are compared and the significance levels h calculated. The vector $\mathbf{x}_j$ corresponding to the highest value of h is selected and used together with $\mathbf{x}_i$ in the process of the determination of largest Lyapunov exponent according to equation (2).

The value h of the significance level needed to reject the null hypothesis has been also used by us for the determination of the distance $|\Delta\mathbf{x}_{ij}|$ between two vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ that contribute to the Lyapunov exponent. We have applied here the measure of distance defined in the form

$$\left|\Delta\mathbf{x}_{ij}\right| = 1 - h \, . \tag{3}$$

This measure was used in the relation (2) for the determination of $|\Delta\mathbf{x}_{ij}(0)|$ and $|\Delta\mathbf{x}_{ij}(\Delta t)|$ in all our experiments.

The important observation from many experiments is that in the direct preictal stage we can observe the phenomena of progressive locking of the EEG waveforms observed at many sites of the brain. To quantify many observed sites simultaneously we have introduced some measure of similarity of profiles. We have employed the T_{index} from the well known t-test statistics as a measure of distance between the mean values of pairs of $STL_{\max}$ profiles over time [7]. The value of T_{index} at the time t between the electrode i and j is defined as

$$T_{ij}(t) = \frac{E\left\{ \left|STL_{\max,i}(t) - STL_{\max,j}(t)\right|\right\}}{\sigma_{i,j}(t)/\sqrt{N}} \, , \tag{4}$$

where E{} means the average of all absolute differences $|STL_{max,i}(t) - STL_{max,j}(t)|$ within a moving window w_t, where w_t is equal 1 in the interval $[t\text{-}N\text{-}1,t]$ and zero elsewhere, while N is the length of the moving window. The variable $\sigma_{ij}(t)$ is the sample standard deviation of $STL_{\max}$ differences between electrode sites i and j within the moving window w_t. To observe the general trend for all electrodes at the same time we define the global T_{index} value for the combination of all pairs of electrodes

$$T_{index} = E\{T_{ij}\} \, . \tag{5}$$

The effect of gradual locking of EEG registrations at different channels is manifested in the form of the decreasing trend of T_{index} versus time. In the vicinity of the onset point the T_{index} assumes the minimum value.

Fig. 2 presents the typical EEG waveform (lower subplot) registered within the period of 40 min with the seizure occurring at 31min. The seizure point indicated by the neurologist is denoted by the vertical line. The upper and middle subplots present respectively, the changes of the largest Lyapunov exponents of the EEG of 8 channels and the T_{index} measure corresponding to these channels. The time axis is

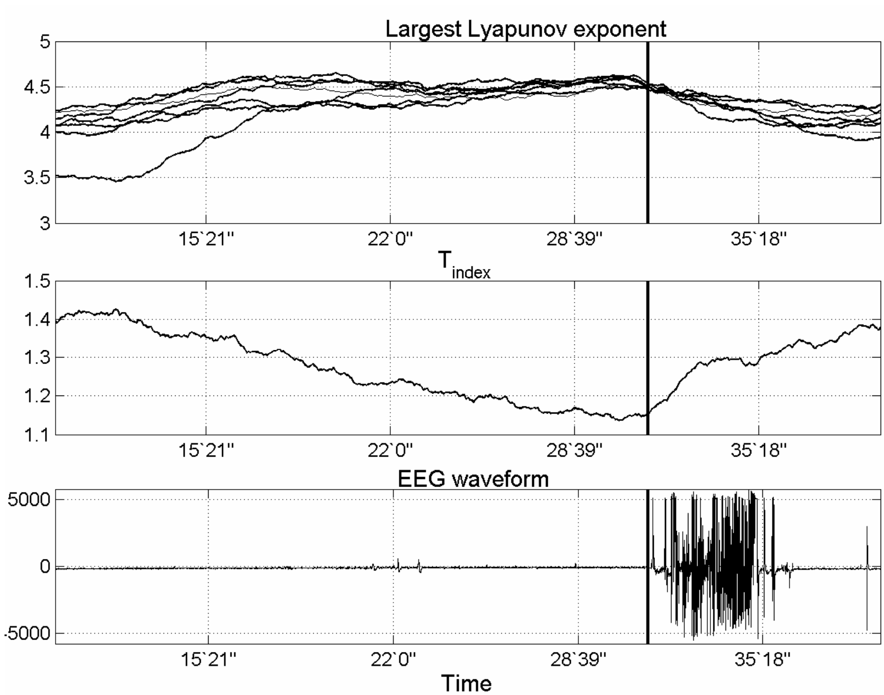

Fig. 2. The typical change of the Lyapunov exponents of the epileptic EEG waveform

given in minutes and seconds. The seizure onset represents a temporal transition of the system from a chaotic state to a less chaotic one, corresponding to a synchronized rhythmic firing pattern of neurons participating in a seizure discharge. So it may be associated with the drop in the Lyapunov exponent values (the decreasing trend) determined for all electrode sites participating in the seizure. The synchronization of the channels is manifested by the decreasing trend of T_{index} measure.

3 Prediction of the Seizure Using Support Vector Machine

The main goal of the research is elaboration of the method for prediction of the incoming epileptic seizure on the basis of the registered EEG waveform. To solve the problem we have applied the Support Vector Machine working in the classification mode and EEG characterization using largest Lyapunov exponents. The SVM network is trained to recognize the time span corresponding to the preictal stage, starting 10 minutes before the potential seizure and lasting up to the seizure moment (destination d=1). The time span outside this period (in the interictal stage) is associated with the other destination (d=-1). So the SVM network works in a classification mode.

3.1 Feature Generation

The most important point in training the SVM classifier is the generation of the diagnostic features forming the vector **x** applied to its inputs. We have made use of the observation that the seizure is strictly associated with some pattern of changes of the Lyapunov exponent preceding it. To take into account the change of the Lyapunov exponents within time we consider the EEG waveform split into 10 second segments. Fig. 3 shows the way the EEG waveform is split into 100 segments, each of approximately 10 seconds duration. For each segment we determine the Lyapunov exponent L and T_{index} associated with the channels under observation. On the basis of these values we generate the first set of diagnostic features used by the SVM network in the prediction process.

Fig. 3. The division of the EEG waveform into 100 segments

This set of features is obtained from the polynomial approximation of the series of 100 Lyapunov exponent L and T_{index} values. The values of the polynomial coefficients are generated for the actual segment and 99 past segments, directly preceding it. We apply the polynomial of the second order $P(x)=ax^2+bx+c$, built separately for Lyapunov exponent ($x=L$) and for T_{index} ($x=T_{index}$). The coefficients a, b and c for both cases create the features (6 features together).

The next features are equal to the standard deviations of these 100 Lyapunov exponent and T_{index} values (2 features). The last set of features is generated from the ARMA model [7] of the EEG waveform of the actually considered segment (segment No 1). We have applied (1,3) ARMA model of four parameters forming the features. In this way the input vector **x** to the SVM is composed of 12 components (6 features following from polynomial approximation, 2 standard deviation values and 4 coefficients of ARMA model). Observe that one input vector **x** needs the analysis of the EEG waveform of duration of approximately 1000 seconds. The input data set has been generated for the segments moving each time by 10 seconds. In this way the EEG recording of few hours may result into few hundreds of data points. Some of them have been used in learning and the rest in the testing mode.

3.2 Support Vector Machine

The prediction task is performed in our system by the Support Vector Machine network (SVM) known as the excellent tool of good generalization ability [11]. The Gaussian kernel SVM based classifiers have been found the best. The principle of operation of such classifier and the review of the learning algorithms can be found for example in the book [11].

The important point in designing SVM classifier is the choice of the parameter σ of the Gaussian kernel function and the regularization parameter C. The parameter C controls the tradeoff between the complexity of the machine and the number of non-separable data points used in learning. The small value of C results in the acceptation of more not separated learning points. At higher value of C we get the lower number of classification errors of the learning data points, but more complex network structure. The optimal values of C and σ were determined after additional series of learning experiments through the use of the validation test sets. Many different values of C and σ combined together in the learning process have been used in the learning process and their optimal values are those for which the classification error on the validation data set was the smallest one.

The whole set of data has been split into the learning set (60%) and testing one (40%). The testing set is used only for testing the trained system. The number of representatives of the seizure and no seizure warning periods have been balanced as much as possible in both sets. The SVM network was trained using the learning data and the trained network tested on the testing data not used in learning. The hyperparameters C and σ have been adjusted by applying the validation data (20% of the learning data). The final values of these parameters used in experiments were C=500 and $\sigma = 0.8$.

4 The Results of the Numerical Experiments

The numerical experiments have been performed for 15 patients of the Banach Hospital in Warsaw. Seven of them suffered from the temporal lobe epilepsy and eight from frontal lobe epilepsy. The EEG waveforms have been registered at 8 electrodes placed on the scalp with the sampling rate of 250 samples per second. They have been filtered by the low pass filter of the cut-off frequency 70Hz.

The estimations of the Lyapunov exponent values have been performed by dividing the recorded signal into the segments of T=10 seconds. The embedding dimension p applied in the experiments was equal 8, while the other parameters: $\tau = 0.016$s, $\Delta t = 45$s. The whole segment of the length T was divided into 270 subsegments and each of them was associated with the vector **x** for the largest Lyapunov exponent estimation.

To assess the results we have defined some measures of the prediction quality. Let p be the number of data segments forming the testing data, from which p_s denote the number of segments indicating the incoming seizure and p_n the number of segments outside the seizure region ($p = p_s+p_n$). By

$$\varepsilon = \frac{\Delta p}{p} \tag{6}$$

we denote the global misclassification rate, as the ratio of erroneous predictions (Δp) to the total number (p) of segments.

For the learning purposes we have assumed the period of 10 minutes of the preictal stage just before the seizure, for which the assumed destination was 1 (warning of the incoming seizure). The interictal period used in learning has covered the time span up to 20 minutes prior to seizure and these patterns have been associated with the alternative destination (no seizure warning). The learning and testing data have been generated for each patient separately and the separate SVM networks have been trained for each patients (15 SVM networks). In the testing phase we have observed the real advance time T_p of prediction and the total testing error ε. Additionally for each patient we have noticed the total time of recorded EEG and the number of segments belonging to the preictal (p_s) and interictal (p_n) stages, used in experiments.

Table 1. The results of testing the SVM system for seizure prediction

Patient	Time of EEG recording *hour:min:sec*	p_s	p_n	ε [%]	T_p *min:sec*
1	1:55:49	137	134	16.54	9:06
2	1:56:34	134	138	0.73	10:04
3	0:58:28	137	134	11.39	10:08
4	058:33	138	134	1.84	10:12
5	0:54:24	136	136	14.34	9:45
6	0:30:43	70	72	17.6	14:25
7	1:38:10	140	132	10.3	10:07
8	0:58:57	136	136	8.83	8:54
9	1:47:00	134	136	8.89	9:52
10	1:41:03	136	136	9.19	9:40
11	2:01:36	139	133	4.04	10:12
12	1:21:39	137	135	12.13	9.35
13	0:42:08	136	136	16.19	9:57
14	2:01:14	132	138	1.85	10:05
15	0:52:25	133	139	6.98	10:13

Table 1 depicts the results of prediction of the seizure for 15 patients suffering from epilepsy. For each patient there was one seizure moment determined by the neurologist. All results depicted in the table correspond to the testing data, not taking part in learning. Observe that the prediction problem has been transformed to the classification task.

The average misclassification rate ε for 15 patients was equal 9.39% and the advance prediction time was in a very good agreement with the time assumed in learning phase (10 minutes). This confirms our conjecture of strict connection of the patterns of change of the largest Lyapunov exponents with the incoming seizure.

5 Conclusions

The paper has presented the application of the analysis of the EEG waveform and the SVM network to the prediction of the epileptic seizure of the brain. We have assumed the chaotic model of the brain activity and applied the largest Lyapunov exponent for its characterization. The employed measures do not assume any particular nonlinear model of the process.

The results of numerical simulations have confirmed that this model of prediction of the incoming seizure has great potentiality of application for individual patients. The obtained average accuracy of prediction was approximately 91% at the advance time span of prediction close to 10 minutes.

Acknowledgements. The paper has been supported by the Brain Science Institute, RIKEN, Japan.

References

1. Babloyantz, A., Destexhe A.: Low dimensional chaos in an instance of epilepsy. Proc. Natl. Acad. Sci, USA 83 (1986) 3513-3517
2. Freeman W. J.: Strange attractors that govern mammalian brain dynamics shown by trajectories of EEG potentials. IEEE Trans. CaS. 35 (1988) 781-784
3. Gevins A., Remond A.: Handbook of EEG and clinical neurophysiology. (1987) Elsevier, Amsterdam
4. Iasemidis L. D., Principe J. C., Sackellares J. C.: Measurement and quantification of spatio-temporal dynamics of human seizures. (in "Nonlinear Signal Processing in Medicine", ed. M. Akay) IEEE Press, (1999) 1-27
5. Iasemidis L. D., Shiau D. S., Chaovalitwogse W., Sackellares J. C., Pardalos P. M., Principe J. C., Carney P. R., Prasad A., Veeramani B., Tsakalis K.: Adaptive epileptic seizure prediction system. IEEE Trans. Biomed. Eng. 50 (2003) 616-627
6. Iasemidis L., Sackellares J. C.: Chaos theory in epilepsy. The NeuroScientist, 2, (1996) 118-126
7. Matlab with toolboxes. (2002) MathWorks, Natick, USA
8. Nikias C., Petropulu A.: Higher order spectral analysis. (1993) Prentice Hall, N. J.
9. Osowski S.: Neural networks for signal processing. (2000) OWPW Warsaw
10. Palus M., Albrecht V., Dvorak I.: Information theoretic test for nonlinearity in time series. Phys. Lett. A 175, (1993) 203-209
11. Schölkopf B., Smola A.: Learning with kernels. (2002) Cambridge MA, MIT Press
12. Swiderski B., Osowski S., Rysz A.: Lyapunov Exponent of EEG Signal for Epileptic Seizure Characterization. (2005) IEEE Conf. European Circuit Theory and Design Cork