35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Analysis of Epileptic EEG Signals in Children by Symbolic Dynamics

Luca Paternoster, Montserrat Vallverdú, Umberto Melia, Francisco Clariá,
Andreas Voss, and Pere Caminal

Abstract—**Epilepsy is one of the most prevalent neurological disorders among children. The study of surface EEG signals in patients with epilepsy by techniques based on symbolic dynamics can provide new insights into the epileptogenic process and may have considerable utility in the diagnosis and treatment of epilepsy. The goal of this work was to find patterns from a methodology based on symbolic dynamics to characterize seizures on surface EEG in pediatric patients with intractable epilepsy. A total of 76 seizures were analyzed by their pre-ictal, ictal and post-ictal phases. An analytic signal envelope algorithm was applied to each EEG segment and its performance was evaluated. Several variables were defined from the distribution of words constructed on the EEG transformed into symbols. The results showed strong evidences of detectable non-linear changes in the EEG dynamics from pre-ictal to ictal phase and from ictal to post-ictal phase, with an accuracy higher than 70%.**

I. Introduction

Epilepsy is responsible for high levels of suffering, affecting more than 50 million people worldwide, thus making it an important public health problem. It is a chronic disorder of the central nervous system that predisposes individuals to episodic interruptions of cerebral electrical activities recurrent referred to as seizures. Sufferers can be of all ages, but epilepsy especially affects children, adolescents and the aged.

The surface electroencephalogram (EEG) remains the most useful and cost effective tool in the diagnosis of epilepsy. Unfortunately, the occurrence of an epileptic seizure is not predictable and its process is not completely understood. In this way, diagnostic evaluations of EEG recordings of patients are necessary for better understanding the process leading to the seizure generation. Also, automated methods are necessary for analysis of epilepsy events because the EEG visual inspection results time intensive.

This work was supported within the framework of the CICYT grant TEC2010-20886 from the Spanish Government and the Research Fellowship Grant FPU AP2009-0858 from the Spanish Government. L. Paternoster, U.S. P. Melia, M. Vallverdú and P. Caminal are with Dept. ESAII, Centre for Biomedical Engineering Research, Universitat Politècnica de Catalunya, CIBER of Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Barcelona, Spain; email: {luca.paternoster, umberto.melia, montserrat.vallverdu, pere.caminal}@upc.edu. F. Clariá is with Dept. IIE, Lleida University, Spain; email: claria@diei.udl.es, A. Voss is with the University of Applied Sciences Jena, Department of Medical Engineering, Jena, Germany.

Investigations have confirmed that the transitions in and out of an abnormal ictal state are not a random process; instead they are dynamical transitions [1]. During the transitions of a seizure, several features are traditionally evaluated by describing the static behavior of the signals [2] within a short time interval, such as signal energy or frequency changes. However, since EEG signal exhibits complex behavior with nonlinear dynamic properties, the analysis of these signals with techniques that measure the chaoticity can improve the accuracy of the seizure event detections. Iasemidis *et al*. [3, 4] based their first studies on the temporal evolution of the short-term largest Lyapunov exponent for patients with temporal lobe epilepsy. The results showed that the EEG activity becomes less chaotic as the seizure approaches and provided evidences to the hypothesis of a route to seizure, a more order state of the brain. The feasibility of using trends in Kolmogorov entropy to anticipate seizures in pediatric patients with intractable epilepsy was demonstrated in [5]. It was concluded that the Kolmogorov entropy was as effective as the correlation dimension in anticipating seizures [2, 5].

In this sense, our analysis will address the quantification of EEG of intractable epilepsy in pediatric patients by the methodology of symbolic dynamics (SD). This methodology has been selected for this study since it is able to suit for the duration variability of the time series that last from seconds up to several minutes and proves to be adequate to describe complex and nonlinear systems [6-8]. The selection of the critical parameters and variables involved in this methodology will be carefully determined to better characterize state transitions of the system over time.

II. Material and Methods

A. Analyzed EEG Data and Preprocessing

The analyzed EEG database belongs to PhysioBank, a collection of well-characterized biomedical signals, containing validated seizure references [9]. For the present study, the selected database consists of surface EEG recordings from pediatric subjects with intractable seizures [10] ranging in age from 0 and 10 years. A total of 11 patients (2 males and 9 females) have been analyzed.

All EEG signals, with a sampling rate of 256 Hz and 16-bit resolution, were down sampled at 128 Hz. Channel CZ-PZ was selected for obtaining the EEG signal patterns, since midline placements are useful for the detection of residual low-voltage physiologic activity and are relatively free from artifacts [11]. In accordance with the minimum seizure

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

duration of the EEG signals, segments of N = 800 samples were selected before (pre), during and after (post) ictal (I) phases. Moreover, the selected segments were not overlapped and always a gap of at least 5 seconds was considered between pre-ictal and ictal and also between ictal and post-ictal, in order to avoid the segments to be part of the ictal phase. In this way, three segments were selected for each seizure event, and a total of 76 seizures were analyzed.

A high-pass FIR filter of 480th order with normalized cut-off frequency of 0.4 Hz was applied. Then, a filter based on analytic signal envelope (ASEF) [12] was applied to each EEG segment. This filter decreases the intensity of the peaks of the signal but keeps the frequency information. The EEG signal segments were analyzed with and without ASEF filtering. Moreover, each selected segment was filtered into the main EEG frequency bands: δ, <4 Hz; θ, 4-8 Hz; α, 8-12 Hz; and β, >12 Hz. Also, the EEG in the total (*tot*) frequency band was considered in this study.

B. Time and Frequency Domain Analysis

Traditional variables of time and frequency domain analysis were calculated for each EEG segment. In this way, the standard deviation (*STD*) of each EEG segment and its power spectral density (*PSD*) were evaluated.

C. Symbolic Dynamics Analysis

The non-linear EEG dynamics were analyzed by means of symbolic dynamics. Each EEG segment was transformed into four symbols from a given alphabet [6], in order to preserve the essential and robust properties of the dynamics in the EEG signal. Equation (1) indicates the signal transformation into four non-equidistant levels, where $EEG_m(i)$ is the $EEG(i)$ series after adding the mean value of its amplitude range.

$$S_i = \begin{cases} 0: & \overline{EEG_m} \le EEG_m(i) < (1+a)\,\overline{EEG_m} \\ 1: & (1+a)\,\overline{EEG_m} \le EEG_m(i) < \infty \\ 2: & (1-a)\,\overline{EEG_m} \le EEG_m(i) < \overline{EEG_m} \\ 3: & 0 \le EEG_m(i) < (1-a)\,\overline{EEG_m} \end{cases} \quad (1)$$

where $\overline{EEG_m}$ is the mean value of $EEG_m(i)$.

Then, a new series S_i was obtained for each EEG segment, where i=1, 2, 3, …, N samples. The parameter a was set to {0.01, 0.025, 0.05, 0.0625, 0.075, 0.0875, 0.1, 0.125} in order to match the standard deviation of the $EEG(i)$ series. From the symbol string S_i, M=64 word types $w_{ijk} = \{000, 001, 002, ..., 331, 332, 333\}$ consisting of three successive symbols with an overlap of 2 symbols were defined.

Five types of variables based on the distribution of words w_{ijk} were estimated: $P(w_{ijk})$, occurrence probability of each one of the word types; $fw(THf)$, number of forbidden words whose probability of occurrence $P(w_{ijk})$ is lower than a probability threshold THf={0.01, 0.005, 0.001, 0.0005}; $pw(THp)$, number of words whose probability of occurrence $P(w_{ijk})$ is higher than a probability threshold THp ={0.025, 0.05, 0.1, 0.2, 0.3, 0.5}; SH, Shannon entropy (2); $RE(q)$, Rényi entropy (3) with q = {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.5, 2, 2.5, 3}.

$$SH = -\sum_{i=1}^{M} P(w_{ijk})_i \, log_2\left(P(w_{ijk})_i\right) \quad (2)$$

$$RE(q) = \frac{1}{q-1} log_2\left(\sum_{i=1}^{M} P(w_{ijk})_i^q\right) \quad (3)$$

This procedure was applied to the pre-ictal, post-ictal and ictal segments.

D. Statistical Analysis

All statistical analysis was performed using signed Wilcoxon rank test. The univariate statistical analysis was applied to each variable in order to determine statistical significant ($p<0.05$) differences between pre-ictal, post-ictal and ictal phases. The best variables were selected and enrolled for univariate and multivariate analyses on the basis of discriminant function analysis. The leave-one-out procedure was used as the cross-validation technique.

III. Results and Discussion

A. Time and Frequency Domain

Table I contains the mean values of the analyzed time and frequency domain variables obtained from the EEG series filtered by ASEF procedure and without filtering. During the seizure, the δ rhythm remained the most present rhythm described by the *PSD*, also observed in [13]. The evaluation of the *PSD* in the δ band statistically discriminated between ictal and post-ictal phases, increasing these differences when ASEF filter was applied. In this way, the mean value of PSD_δ during the ictal phase was statistically lower than post-ictal phase ($p<0.0005$; accuracy=65.1%), that was accompanied by a relative increase of the ictal phase in the remaining bands. Between ictal and pre-ictal phases no statistically significant differences were found when considering *PSD*. However, for these cluster of seizures it can be observed a continuous decrease of the low frequency content (δ power) while approaching the next seizure, from post-ictal to pre-ictal ending at ictal phase. The evaluation of *STD* showed higher values during ictal than pre-ictal ($p<0.01$). This result was accordant with [13].

TABLE I. Time and Frequency Domain Indices

Variables	*I Phase mean (SE)*	*Pre-I Phase mean (SE)*	*Post-I Phase mean (SE)*
	Without ASEF Filtering		
STD	76.8 (4.7)*	58.9 (3.8)*	66.9 (4.7)
PSD_δ	58.5 (2.6)†	67.6 (1.4)	72.9 (1.8)†
	With ASEF Filtering		
STD	71.5 (4.0)*	56.2 (3.5)*	62.5 (4.2)
PSD_δ	58.9 (2.6)††	68.4 (1.4)	73.3 (1.7)††

I: Ictal; SE, standard error; *p <0.01; †p <0.001; †acc.=61.8%; ††p <0.0005; ††acc.=65.1%.

B. Symbolic Dynamics

Figs. 1 and 2 show the number of variables obtained from symbolic dynamics in function of the parameter a in the *tot* band, without and with ASEF filtering, respectively, that well classify ($p<0.05$ and accuracy>60%) the seizure segments in pre-ictal, ictal and post-ictal phases. It can be observed that ASEF filtering (Fig. 2) permits a higher number of variables to better classify the seizure segments from $p<0.05$ till $p<0.0005$ than without ASEF filtering (Fig. 1). In this way, the parameter a= 0.125 (see Fig. 2) allowed 9 variables to classify the seizure segments with $p<0.0005$ and 16 variables with $p<0.05$. Without ASEF filtering (Fig. 1) no variables were found with $p<0.0005$ and only 6 variables presented $p<0.05$. Then, these results confirmed to apply ASEF algorithm to the SD analysis of EEG filtered at the *tot* frequency band. A similar study was performed on each EEG frequency band, giving a particularized value of parameter a for each band, as it can be seen in Tables II and III.

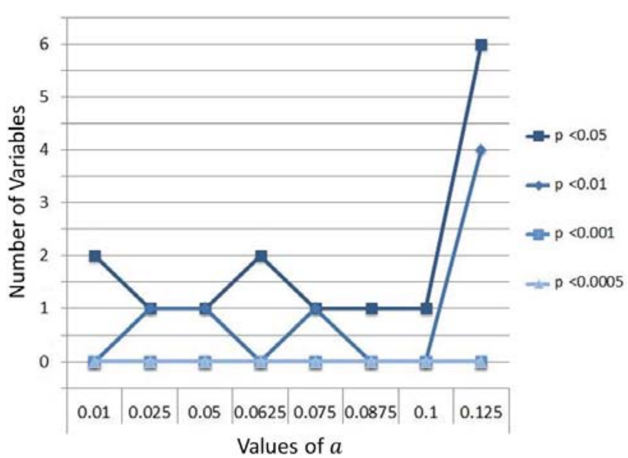

Figure 1. EEG signals without ASEF filtering: Total number of variables, as function of a values, obtained by symbolic dynamics that classify the segments in pre-ictal, ictal and post-ictal phases with statistical significance level $p<0.05$ and accuracy>60%.

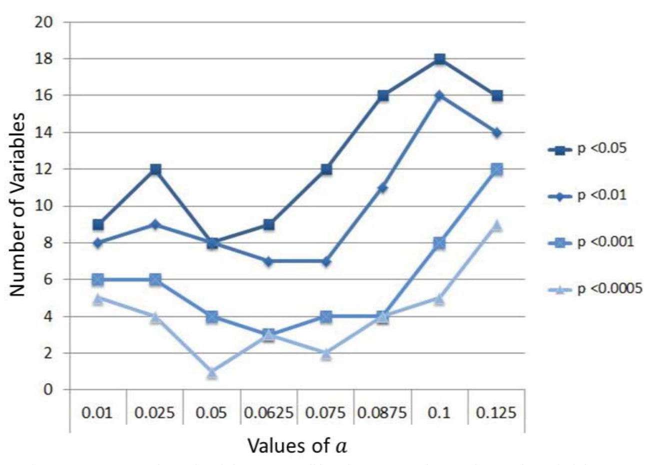

Figure 2. EEG signal with ASEF filtering: Total number of variables, as function of a values, obtained by symbolic dynamics that classify the segments in pre-ictal, ictal and post-ictal phases with statistical significance level $p<0.05$ and accuracy>60%.

Tables II and III present the best variables able to discriminate ictal from pre-ictal or post-ictal phases, only found at *tot*, α and β frequency bands. The mean values of those variables are presented in Table II and the statistical analysis results in Table III. Observing the behavior of the mean value of $P(w_{ijk})$ variables and according how the symbolic dynamics algorithm separates into levels the time signal amplitude, the words w_{ijk} can be grouped into two sets: Set 1, including words composed by two symbols from the lower and/or upper level (level 1 and 3), w_{ijk}={033, 301, 330}; Set 2, including words composed by two middle symbols (level 0 and 2), w_{ijk}= {001, 100, 200, 223}. The mean values of variables $P(w_{ijk})$ that w_{ijk} belong to the same set have a similar trend when ictal phase is compared with pre-ictal or post-ictal phases (see Table II). In particular, the words of Set 1 became much more frequent approaching the ictal phase than pre-ictal or post-ictal phases. This means that the probability of the words in Set 1 was widely higher during the ictal phase than in the other phases. On the other hand, words of Set 2 present an opposite behavior, since these words tend to become less frequent in the ictal phase than pre-ictal or post-ictal phases. This means that the probability of the words in Set 2 was lower during the ictal phase than in the other phases.

TABLE II. SYMBOLIC DYNAMICS: BEST CLASSIFICATIONS

Variables	***I Phase* mean (*SE*)**	***Pre-I Phase* mean (*SE*)**	***Post-I Phase* mean (*SE*)**
tot frequency band **(a = 0.125)**			
P(001)	0.0164 (0.0010)	0.0218 (0.0008)	0.0201 (0.0008)
P(033)	0.0127 (0.0010)	0.0066 (0.0006)	0.0063 (0.0007)
P(100)	0.0162 (0.0010)	0.0218 (0.0008)	0.0199 (0.0008)
P(301)	0.0089 (0.0008)	0.0040 (0.0005)	0.0045 (0.0005)
P(330)	0.0124 (0.0009)	0.0061 (0.0006)	0.0064 (0.0007)
fw(0.005)	29.9 (0.91)	34.4 (0.82)	35.4 (0.95)
pw(0.05)	3.1 (0.19)	4.0 (0.18)	3.4 (0.15)
α *frequency band* **(a = 0.01)**			
P(223)	0.0005 (0.0001)	0.0013 (0.0002)	0.0013 (0.0002)
fw(0.001)	33.0 (0.77)	29.0 (0.75)	29.2 (0.79)
RE(*q*=0.1)	4.74 (0.0384)	4.94 (0.0350)	4.93 (0.0389)
β *frequency band* **(a = 0.1)**			
P(200)	0.0848 (0.0060)	0.1201 (0.0058)	0.1223 (0.0054)

I: Ictal; SE, standard error.

Concerning to *fw*(*THf*) and *pw*(*THf*), these variables could differentiate between ictal phases in the *tot* frequency band (a = 0.125) and α frequency band (a = 0.01). In the *tot* frequency band, *fw*(0.005) and *pw*(0.05) presented the lowest value in ictal phase compared with pre-ictal and post-ictal phases. This means that during the ictal phase a higher number of words (approximately 50%) have probabilities between 0.005 and 0.05, contrarily in pre-ictal and post-ictal phases the words are more distributed, indicating a lower regularity. In the α frequency band, *fw*(0.001) presented the highest value in ictal phase than pre-ictal phase, presenting ictal phase more regularity since *RE*(*q*=0.1) was the lowest ($p<0.0005$). This complexity behavior was in agreement with the results found by other authors [5,14].

Combining two variables from the *tot* frequency band permitted to increase the classification accuracy: *P*(100) and *P*(301), accuracy =73.7% between ictal and pre-ictal phases; *P*(001) and *P*(033), accuracy =70.4% between ictal and post-ictal phases.

It should be noticed that variables calculated in δ and θ frequency bands could not describe the ictal phases. Also, no variables were found presenting statistically significant differences comparing pre-ictal and post-ictal phases.

TABLE III SYMBOLIC DYNAMICS: BEST CLASSIFICATIONS

	Variables	*Acc.(%)*	*p value*
tot frequency band		**(a = 0.125)**	
I vs. Pre-I	P(100)	66.4	<0.0005
	P(301)	69.1	<0.001
	P(330)	66.4	<0.0005
	fw(0.005)	64.5	<0.01
	pw(0.05)	68.4	<0.0005
I vs. Post-I	P(001)	64.5	<0.05
	P(033)	69.1	<0.0005
	P(301)	66.4	<0.0005
	P(330)	65.8	<0.0005
	fw(0.005)	63.8	<0.01
α frequency band		**(a = 0.01)**	
I vs. Pre-I	P(223)	67.8	<0.0005
	fw(0.001)	66.5	<0.0005
	RE(q=0.1)	64.5	<0.0005
I vs. Post-I	RE(q=0.1)	63.8	<0.0005
β frequency band		**(a = 0.1)**	
I vs. Post-I	P(200)	65.8	<0.0005

I: Ictal; *acc.* = accuracy.

IV. CONCLUSIONS

The present work consisted of a preliminary study of intractable seizures in pediatric patients by applying a symbolic dynamics technique upon a database of scalp recorded EEG signals. The results obtained from symbolic dynamics applied to signals filtered by ASEF algorithm evidenced the presence of non-linear patterns that permit to localize seizure events with accuracy higher than 70%. While non-linear measures appeared to be sensitive to changes in ictal phase in relation to the periods of pre-ictal and post-ictal phases, linear measures were not. Also, this analysis has allowed the selection of the parameters involved in the methodology and the definition of the variables that best characterize state transitions of the system over time. It can be concluded that the proposed symbolic dynamics methodology can be effective on recognizing differences between epilepsy EEG events.

ACKNOWLEDGMENT

CIBER of Bioengineering, Biomaterials and Nanomedicine is an initiative of ISCIII.

REFERENCES

[1] F. Mormann, R. G. Andrzejak, Ch. E. Elger and K. Lehnertz, "Seizure prediction: the long and winding road", *Brain*, vol. 130, 2007, pp. 314–333.

[2] A. Aarabi, R. Fazel-Rezai, and Y. Aghakhani, "EEG Seizure Prediction: Measures and Challenges", *31st Annual International Conference of the IEEE EMBS*, 2009, pp. 1864-1867.

[3] L. D. Iasemidis, J. C. Sackellares, "Chaos theory and epilepsy", *The Neuroscientist*, vol. 2, no. 2, Mar. 1996, pp. 118-126.

[4] L. D. Iasemidis, D. S. Shiau, J. C. Sackellares, P. M. Pardalos, A. Prasad, "A dynamical resetting of the human brain at epileptic seizure: applications of nonlinear dynamics and global optimizations techniques", *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 3, Mar. 2004, pp. 493-506.

[5] W. van Drongelen, S. Nayak, D. M. Frim, M. H. Kohrman, V. L. Towle, H. C. Lee, A. B. McGee, M. S. Chico, and K. E. Hecox, "Seizure anticipation in pediatric epilepsy: use of Kolmogorov entropy", *Pediatr. Neurol.*, vol. 29, 2003, pp. 207-13.

[6] A. Voss, J. Kurths, H. J. Kleiner, A. Witt, N. Wessel, P. Saparin, K. J. Osterziel, R. Schurath, R. Dietz, "The application of methods of non-linear dynamics for the improve and predictive recognition of patient threatened by sudden cardiac death", *Cardiovascular Research*, vol. 31, no. 3, Mar. 1996, pp. 419-433.

[7] A. Voss, R. Schroeder, P. Caminal, M. Vallverdú, H. Brunel, I. Cygankiewicz, R. Vázquez, A. Bayés de Luna, "Segmented symbolic dynamics for risk stratification in patients with ischemic heart failure", *Cardiovascular Engineering and Technology*, vol. 1, no. 4, Dec. 2010, pp. 290 -298.

[8] N. Tupaika, M. Vallverdú, M. Jospin, E. W. Jensen, M. M. R. F. Struys, H. E. M. Vereecke, A. Voss, P. Caminal, "Assessment of the depth of anesthesia based on symbolic dynamics of the EEG", *2010 Annual International Conference of the IEEE, Engineering in Medicine and Biology Society*, pp. 5971 – 5974.

[9] A. L. Golberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, Peng Chung-Kan, H. E. Stenley, "PhysioBank, physioToolkit and physionet: component of a new research resource for complex physiologic signal", *Circulation,* vol. 102, no. 23, 2000, pp. 215-220.

[10] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment", PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2009.

[11] C. M. Epstein, D. B. Mark, N. Foldvary-Schaefer, T. D. Lagerlund, S. L. Helmers, A. M. Husain "Guideline three: minimum technical standards for EEG recording in suspected cerebral death", *J Clin Neurophysiol*, vol. 23, no. 2, Apr. 2006, pp. 97-104.

[12] U. Melia, F. Clariá, M. Vallverdú, P. Caminal, "Removal of peak and spike noise in EEG signals based on the analytic signal magnitude", *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (*EMBS*), 2012, pp.3523-3526.

[13] F. Mormann, T. Kreuz, C. Rieke, R. G. Andrzejak, A. Kraskov, P.David, C. E. Elger, and K. Lehnertz, "On the predictability of epileptic seizures", *Clin Neurophysiol*, vol. 116, 2005, pp. 569-87.

[14] H. R. Moser, B. Weber, H. G. Wieser, P. F. Meier, "Electroencephalograms in epilepsy: Analysis and seizure prediction within the framework of Lyapunov theory", *Physica D*, vol. 130, 1999, pp. 291-305.