34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Epileptic Seizure Prediction based on a bivariate spectral power methodology

Mojtaba Bandarabadi, Cesar A. Teixeira, Bruno Direito, and Antonio Dourado, *Member IEEE*

Abstract—**The spectral power of 5 frequently considered frequency bands (Alpha, Beta, Gamma, Theta and Delta) for 6 EEG channels is computed and then all the possible pairwise combinations among the 30 features set, are used to create a 435 dimensional feature space. Two new feature selection methods are introduced to choose the best candidate features among those and to reduce the dimensionality of this feature space. The selected features are then fed to Support Vector Machines (SVMs) that classify the cerebral state in preictal and non-preictal classes. The outputs of the SVM are regularized using a method that accounts for the classification dynamics of the preictal class, also known as "Firing Power" method. The results obtained using our feature selection approaches are compared with the ones obtained using minimum Redundancy Maximum Relevance (mRMR) feature selection method. The results in a group of 12 patients of the EPILEPSIAE database, containing 46 seizures and 787 hours multichannel recording for out-of-sample data, indicate the efficiency of the bivariate approach as well as the two new feature selection methods. The best results presented sensitivity of 76.09% (35 of 46 seizures predicted) and a false prediction rate of $0.15h^{-1}$.**

I. Introduction

During recent years, several methods have been proposed for epileptic seizures prediction. Success would improve the living expectations of over 50 million patients suffering from ictal events. Despite the published performances of such methods, when are applied to new long-term EEG recordings, usually the presented results are not as expected. Among the features currently published, the studies on the spectral power of raw EEG signal have demonstrated the ability to track the transient changes from the normal state (interictal) to the ictal state [1] [2] [3] [4].

Mormann et al. [1] described a relative decrease in the power of the Delta band in preictal period in comparison with the interictal period. Additionally, this decrease was accompanied by a relative increase of the power in the remaining bands. Netoff et al. [2], proposed a patient-specific algorithm, based on the features obtained from spectral powers in the sub-bands: delta (0.5-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz), four gamma (30-50Hz, 50-70Hz, 70-90Hz, 90Hz-), and total power of the six EEG electrodes, 3 over the seizure focus and 3 distant from the focus. They reported an average sensitivity of 77.8% (35 of 45 seizures), and a false positive rate per hour (FPR) of zero. They argued that the spectral power in certain sub-bands of the iEEG (intracranial, invasive), specifically in higher frequency sub-bands, may play a key role in seizure prediction.

Later the same authors [3] proposed a patient-specific seizure prediction algorithm using four different methods to compute spectral power of the iEEG: raw, time-differential, space-differential, and time/space-differential, and used them as features. The proposed algorithm was applied on recordings containing a total of 80 seizures and 437-hour of inter-ictal data. The best results were achieved with the features obtained from time/space-differential approach with 86.25% sensitivity and 0.1281 false positives per hour in out-of-sample data.

These studies presented the spectral power in different sub-bands extracted from one electrode or from differential electrodes. The objective of this work is to compare the different spectral power in different sub-bands and electrodes, and derive relations that will then be used as features. Two new approaches are also introduced for selecting features from the high dimensional bivariate features set. The seizure prediction problem is faced as a binary classification problem: the preictal and non-preictal states. Preictal is the state just before one seizure that one wants to predict.

Section II presents the methodological aspects related to feature computation and selection. The results in a set of 12 patients are presented in Section III. Finally in Section IV the main conclusions are drawn.

II. Methodology

A. Spectral Power of Bands

The EEG signal has usually been expressed in terms of particular frequency sub-bands: Delta 'δ' (less than 4 Hz), Theta 'θ' (4-8 Hz), Alpha 'α' (8-15 Hz), Beta 'β' (15-30 Hz), and Gamma 'γ' (greater than 30 Hz). The above classification of the sub-bands is not unique, and a variety of other classifications have been presented, with many similarities. Spectral power of raw EEG is performed by using discrete Fourier transform (DFT) of windowed EEG signal. The DFT is calculated based on the stationarity assumption of the EEG signal, thus the raw EEG signal is first segmented to minimize the effect of non-stationarity.

Spectral power of a sub-band can be expressed as absolute

Acknowledgments: this work was partially funded by the European Project EPILEPSIAE (EU FP7 Grant 211713). MB and BD would also like to acknowledge the Portuguese Foundation for Science and Technology (SFRH / BD / 71497 / 2010 and SFRH / BD / 47177 / 2008). CT is supported by a Postdoc contract under the Ciencia 2007 program of FCT.

M. Bandarabadi, C. A. Teixeira, B. Direito, and A. Dourado are with the Centre for Informatics and Systems (CISUC), University of Coimbra, Portugal (phone:+351-239-790000; e-mail: {mojtaba, cteixei, brunodireito, dourado}@dei.uc.pt).

or normalized values. The normalized spectral power (NSP) feature for a given sub-band is computed by dividing the spectral power of the sub-band by total power (1),

$$NSP_i = \frac{\sum_i |DFT(x)|}{\sum_{tot} |DFT(x)|} \quad (1)$$

where x is a portion of raw EEG, i and tot index a given frequency sub-band and total frequencies respectively. $|DFT(x)|$ is the absolute values of Fourier transform. The advantage of using the *NSP* values instead of absolute values is that they decrease the effect of changes in total power on the values of spectral power of sub-bands [5]. Thus the *NSP* values of five sub-bands (δ, θ, α, β, and γ) are investigated in this study.

B. Bivariate Approach

Around 60% of the epileptic patients suffer from partial seizures that are related to a specific brain region [6]. Comparing the EEG from the focal region with the EEG from other regions it is likely to find significant signal differences, at least during the ictal phase. So, features that explore this spatial dependence should be considered (Fig.1). Multivariate features developed in this way have other advantages, such as the rejection of common mode interferences that are not related with ictogenesis.

With the previous assumptions a bivariate approach based on univariate normalized power in different sub-bands is presented in this paper. Mathematically the new feature is described by (2):

$$RNSP_{i_1 j_1 i_2 j_2} = \frac{NSP_{i_1 j_1}}{NSP_{i_2 j_2}} \quad , \begin{cases} i_1, i_2 = 1, \dots, No.\ of\ Subbands \\ j_1, j_2 = 1, \dots, No.\ of\ Channels \\ \text{if} \quad j_1 = j_2 \quad \text{then} \quad i_1 \neq i_2 \end{cases} \quad (2)$$

where *RNSP* stands for Relative Normalized Spectral Power, which is given by the ratio between the normalized spectral power for the i_1-th band in the j_1-th channel, and the normalized spectral power for the i_2-th band in the j_2-th channel. So, this feature gives the cross-power information not just between two channels but also between two frequency bands. For instance, if six channels and 5 sub-bands are considered independently, 30 features will be achieved. The combination of the 30 features set by the proposed bivariate approach will lead to a total of 435 new features (C_2^{30}). These new features are targeted to find preictal trends that can be used to predict seizures.

Since the number of bivariate features is very high, the individual study using long-term recordings presents a computational challenge. So, feature selection methods were designed to select the most promising feature subset. In the next sub-sections two new approaches for feature selection are introduced, and will be then compared with the existent mRMR [7] method.

C. Feature selection based on amplitude distribution

A new supervised feature selection method is introduced based on amplitude distribution histograms (ADH). An ADH is the histogram of the samples of a given feature associated with one class. For a two-class problem two different ADH are considered.

The basic idea of the method is the selection of the features that have the Maximum Difference of ADHs (mDAD). The difference of ADHs for a two-class problem is defined as (3):

$$DAD = 1 - CA \quad (3)$$

where *CA* is the common area of two normalized ADHs (Fig.2) and is calculated as:

$$CA = \sum_i minimum(ADH_{norm1}, ADH_{norm2}) \quad (4)$$

where ADH_{norm1} is the normalized ADH of the class 1, and i indexes the interval that the values of two classes are distributed. In order to achieve the normalized ADHs, the original histograms are divided by the number of samples in each class (5).

$$ADH_{norm} = \frac{ADH}{Number\ of\ samples} \quad (5)$$

The total area under each normalized ADH is one, and the common area is a value in the interval [0 1]. In summary, lower *CA* values represent higher separability between samples of different classes, in a given feature. So, features with low *CA* are more likely to improve seizure prediction performances.

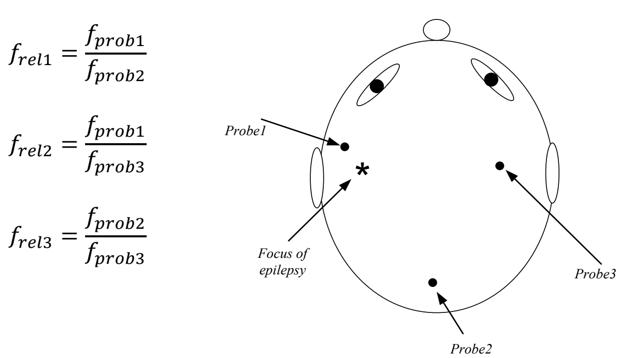

Fig.1 The bivariate approach. Relative features are achieved by dividing the features extracted from each channel (specifically focus channels) by the features of the other channels (as reference channels).

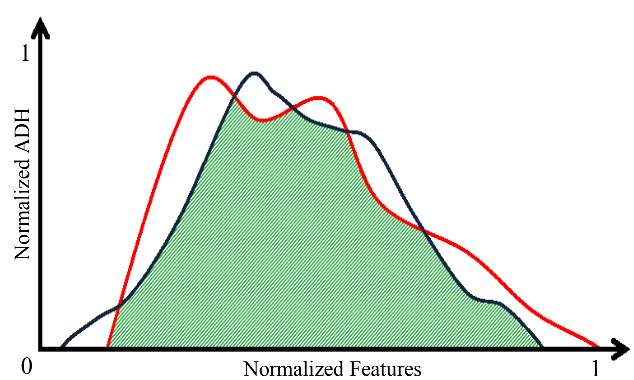

Fig.2 Common area of the ADHs of two classes (green hachure), the normalized ADH of the preictal samples (red curve) and the normalized ADH of the non-preictal samples (blue curve)

Since each sample is labeled, this method is supervised method, and currently it is developed for two classes problems such as in seizure prediction (preictal and non-preictal classes).

D. Feature selection based on percentiles

A different measure is introduced to determine the separability between two classes. The idea is based on the percentiles of the samples of two classes. Since in the real classification problems, the values of the samples of the different classes have some overlap, searching for the features that have the minimum overlap could be used for selecting the features. In order to quantify the overlap value, a new measure is introduced as (6):

$$NDP = \frac{p_{n1}(class\ 1) - p_{n2}(class\ 2)}{p_{50}(class\ 2) - p_{50}\ (class\ 1)} \tag{6}$$

In (6), *NDP* is normalized difference of the percentiles, $p_n(class\ m)$ is the n-th percentile of values of the features of the class m (Fig.3). The samples of the classes have minimum overlap when the above measure is minimized (*mNDP*). According to this principle, the features presenting the lowest *NDP* should be selected. The 70^{th} and 30^{th} percentiles are selected for class1 and class2 respectively by trial and error.

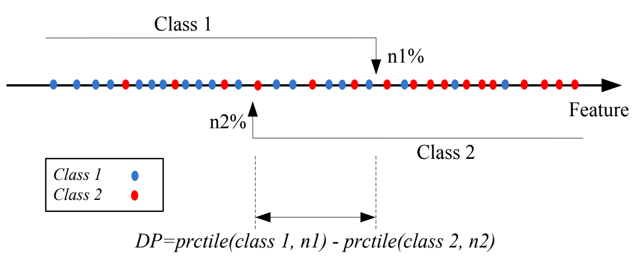

Fig.3 Difference of percentiles of n1 and n2 of two classes samples

E. Support Vector Machines

Support vector machines (SVMs) are a set of commonly used supervised learning methods employed for classification problems [8]. SVM classifiers in their simplest form use linear boundaries to classify binary data. To classify datasets with nonlinear boundaries, SVM uses kernel functions representing the data in a higher feature space where linear boundaries may separate data. The popular Gaussian Radial Basis Function (RBF) kernel (7) is used,

$$K(x,y) = exp(\frac{-|x-y|^2}{2\sigma^2}) \tag{7}$$

where σ is the scale parameter (openness of the Gaussian), *x*, *y* are feature vectors in the input space. The Gaussian kernel has two hyper parameters to control classification performance: the cost *C* and the scale parameter σ.

Parameter *C* controls the tradeoff between maximization of the margin width and the minimization of the number of misclassified samples in the training set [8]. Also, the σ parameter in (7) controls the width of the Gaussian surface of the RBF kernel. These two parameters are optimized through a grid search method.

F. Output regularization

In order to reduce the number of false alarms, the classification output of the SVM classifiers was subjected to regularization by a method that accounts for the classification dynamics in the preictal class, the firing power (FP). This methodology is explained in detail in [9].

G. Proposed Algorithm

Initially, the normalized spectral power features are extracted from 6 channels which 3 of them are related with the focal area and remained belong to areas far from the focal region. Then the relative features using the proposed bivariate approach are computed. Since the number of achieved bivariate features is very high, feature selection methods are used to find the best subset of features. Classification is carried out using LibSVM toolbox [10], SVMs are trained in a part of the data and tested in a different one. The outputs of the SVM classifier are regularized using the FP method.

III. Experimental Results

Data from 12 epileptic patients from the EPILEPSIAE database [11] with long-term continuous multichannel EEG recordings were used to evaluate the proposed method (Table I). The total number of considered seizures is 82 from which 36 for SVM training (3 for each patient) and the remaining 46 are used for SVM testing.

The normalized spectral power features were computed using a time window of 5 seconds. Four different preictal times: 10, 20, 30, and 40 minutes before each onset are used to label the preictal samples for training and testing the SVM classifier. The number of preictal samples is related to the preictal time, for instance, 10 minutes preictal time will provide 120 samples (600s/5s) for each seizure and totally 360 samples for 3 training seizures. The remaining samples outside the preictal times are used as non-preictal samples. Then three feature selection methods (mDAD, mNDP, and mRMR) are applied on the labeled features to rank them.

Afterward, since the number of non-preictal samples is much more than the preictal samples, and usually classifiers tend to produce high accuracy over the class with more training samples, thus the number of non-preictal samples of the training set is reduced by resampling to achieve a balanced number of samples for the two classes.

For each feature selection method 5 runs were performed selecting in each one a different size of the subset: 3, 5, 10, 20 and 40 features. The outputs of the classifier are regularized by FP method with the predefined threshold of 0.5. The best results for each patient in terms of alarm sensitivity (SS) and false prediction rate (FPR) of alarms are summarized in table 1. The best results are selected so that the SS and FPR are closed to the optimal performance points, i.e., *SS=100%* and *FPR=0* h^{-1}.

IV. Conclusion

A new bivariate approach was introduced for seizure prediction problem based on the normalized spectral power. Discriminative features were selected by two proposed methods. Both introduced feature selection methods have shown improved performance in comparison to the well-known mRMR feature selection method. For instance, the mDAD method has provided a higher performance of seizure prediction with lower number of features (8.75 features in average) in comparison to the mRMR method (9.91 features in average). The achieved results using just first 3 high ranked features for each feature selection method are presented in table 2 for comparison.

In average 76.09% of the seizures in the testing set were predicted with an average FPR of 0.15 h^{-1}. Results show the ability of the proposed techniques to predict the epileptic seizures. However further work is needed to improve the performance of the feature selection methods as well as the bivariate approach. One way for bivariate approach would be to use differential features instead of relative features. Our research aims at good seizure predictors with a low number of channels (equal or less than 6) in order to allow the development of transportable devices for incoming seizure warning.

TABLE 1
INFORMATION AND RESULTS FOR THE 12 STUDIED PATIENTS

Patient Info.							mDAD				mNDP				mRMR			
ID	Type	Samp. Rt. Hz	No. Seiz. Total	No. Seiz. Test	Rec.time (h) Total	Rec.time (h) Test	SS%	FPR (h^{-1})	No. S.F.	SOP (min)	SS%	FPR (h^{-1})	No. S.F.	SOP (min)	SS%	FPR (h^{-1})	No. S.F.	SOP (min)
1	Invas	400	5	2	401	110	100	0.17	5	40	**100**	**0.09**	**3**	**40**	100	0.1	3	30
2	Invas	400	7	4	209	18	75	0.22	3	40	**75**	**0.16**	**5**	**40**	50	0.06	40	40
3	Invas	400	8	5	239	80	40	0.23	10	40	40	0.20	5	40	**40**	**0.19**	**5**	**40**
4	Scalp	400	6	3	140	40	**100**	**0.04**	**10**	**40**	100	0.1	10	30	100	0.17	10	40
5	Scalp	400	3	1	163	43	**100**	**0.02**	**3**	**10**	100	0.09	5	40	100	0.07	10	40
6	Invas	400	12	9	217	157	**77.7**	**0.24**	**5**	**40**	44.4	0.1	10	40	22.2	0.09	3	30
7	Scalp	512	7	4	120	37	75	0.24	40	40	50	0.13	20	10	**75**	**0.24**	**10**	**30**
8	Scalp	512	8	4	120	62	**100**	**0.13**	**3**	**30**	75	0.09	5	40	75	0.19	10	40
9	Scalp	512	7	4	125	23	**100**	**0.04**	**10**	**40**	100	0.08	10	40	100	0.13	10	40
10	Scalp	512	6	3	124	65	**66.6**	**0.08**	**3**	**10**	33.3	0	5	30	66.6	0.11	10	20
11	Scalp	256	6	3	118	68	66.6	0.24	3	10	66.6	0.11	5	40	**66.6**	**0.04**	**3**	**40**
12	Scalp	512	7	4	117	84	50	0.01	10	40	**75**	**0.05**	**3**	**20**	50	0	5	40
Tot./ Avg.			82	46	2093	787	**76.09**	**0.15**	**8.75**	**31.6**	65.22	0.1	7.16	34.1	60.87	0.11	9.91	35.8

Type: Type of EEG recording; Invas: Invasive recording, Scalp: Scalp recording
SS: Sensitivity of the raised alarms in percent
FPR: False prediction rate per hour
No. S.F.: Number of selected features
SOP: The preictal time in minute

TABLE 2
RESULTS OF THE 3 HIGH RANKED FEATURES

ID	mDAD			mNDP			mRMR		
	SS%	FPR (h^{-1})	SOP (min)	S%	FPR (h^{-1})	SOP (min)	SS%	FPR (h^{-1})	SOP (min)
1	50	0.16	40	**100**	**0.09**	**40**	100	0.1	30
2	**75**	**0.22**	**40**	75	0.28	40	25	0	20
3	20	0.21	40	**20**	**0.07**	**20**	20	0.16	40
4	66.6	0.19	40	**66.6**	**0.07**	**40**	66.6	0.24	10
5	**100**	**0.02**	**10**	100	0.11	30	100	0.25	20
6	44.4	0.18	30	**44.4**	**0.17**	**30**	22.2	0.09	20
7	75	0.38	30	25	0	30	**50**	**0.19**	**30**
8	**100**	**0.12**	**30**	75	0.11	40	50	0.08	40
9	**66.6**	**0.04**	**40**	66.6	0.08	40	33.3	0.04	40
10	**66.6**	**0.07**	**10**	33.3	0.03	30	33.3	0.06	20
11	66.6	0.24	10	66.6	0.13	40	**66.6**	**0.04**	**40**
12	25	0.01	40	**75**	**0.05**	**20**	25	0.04	20
Tot. Avg	58.7	0.15	30	**56.52**	**0.1**	**33.3**	41.3	0.1	27.5

References

[1] F. Mormann, T. Kreuz, C. Rieke, R.G. Andrzejak, A. Kraskov, P. David, C.E. Elger, and K. Lehnert, "On the predictability of epileptic seizures," *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology,* vol. 116, no. 3, pp. 569-587, Mar. 2005.

[2] T. Netoff, Y. Park, and K. Parhi, "Seizure prediction using cost-sensitive support vector machine," *IEEE Engineering in Medicine and Biology Society Conference,* vol. 2009, pp. 3322-3325, 2009.

[3] Y. Park, T. Netoff, and K. K. Parhi, "Seizure prediction with spectral power of time/spacedifferential EEG signals using cost-sensitive support vector machine," *in 2010 IEEE International Conference on Acoustics, Speech and Signal Processing,* Dallas, TX, USA, 2010, pp. 5450-5453.

[4] R. Cerf and E. H. el-Ouasdad, "Spectral analysis of stereo-electroencephalograms: preictal slowing in partial epilepsies," *Biological Cybernetics,* vol. 83, no. 5, pp. 399-405, Nov. 2000.

[5] J. O. E. H. van Laar, C. H. L. Peters, S. Houterman, P. F. F. Wijn, A. Kwee, and S. G. Oei, "Normalized spectral power of fetal heart rate variability is associated with fetal scalp blood pH," *Early Human Development,* vol. 87, no. 4, pp. 259-263, Apr. 2011.

[6] J. F. Tellez-Zenteno, and L. Hernandez-Ronquillo, "A Review of the Epidemiology of Temporal Lobe Epilepsy," *Hindawi Publishing Corporation, Epilepsy Research and Treatment,* Volume 2012, Article ID 630853, 5 pages.

[7] H. Peng, C. Ding, and F. Long, "Minimum Redundancy-Maximum Relevance Feature Selection," *IEEE Intelligent Systems,* pages 70–71, 2005

[8] B. Schlkopf and A.J. Smola, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond," *The MIT Press,* 2001.

[9] C. Teixeira, B. Direito, H. Feldwisch-Drentrup, M. Valderrama, R. Costa, C. Alvarado-Rojas, S. Nikolopoulos, M. L. V. Quyen, J. Timmer, B. Schelter, and A. Dourado, "Epilab: A software package for studies on the prediction of epileptic seizures," *Journal of Neuroscience Methods,* vol. 200, no. 2, pp. 257 – 271, 2011.

[10] C.-C. Chang and C.-J. Lin, "LIBSVM: a Library for Support Vector Machines," *ACM Transactions on Intelligent Systems and Technology,* 2:27:1--27:27, 2011. *Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm*

[11] M. Ihle, H. Feldwitch-Drentrup, C.A. Teixeira, A. Witon, B. Schelter, J. Timmer, and A. Schulze-Bonhage, "EPILEPSIAE - A common database for research on seizure prediction," *Computer Methods and Programs in Biomedicine*, 2010.