VOL.2 NO.2September 2013 - ISSN 2233 – 1859

Operations Research Society in Bosnia and Herzegovina

Southeast Europe Journal of Soft Computing

Available online:www.scjournal.com.ba

Localization of the epileptogenic foci using Support Vector Machine

Samed Jukic, Abdulhamit Subasi

International Burch University, Faculty of Engineering and Information Technologies, FrancuskeRevolucije bb.Ilidza, Sarajevo, 71000, Bosnia and Herzegovina

Article Info

Article history:
Received 17 Sep.2013
Received in revised form 17 Oct 2013

Keywords:
Multi-scale Principal Component Analysis (MSPCA), Support Vector Machine (SVM), Electroencephalogram (EEG), Epilepsy, Source Localization

Abstract

Epileptic foci localization is a crucial step in planning surgical treatment of medically intractable epilepsy. The solution to this problem can be determined by the detection of the earliest time of seizure onset in electroencephalographic (EEG) recordings. This study presents the application of support vector machine (SVM) for localization of the focus region at the epileptic seizure on the basis of EEG signals. We used intracranial EEG recordings from patients suffering from pharmacoresistant focal-onset epilepsy. We have been investigating a localization of the focus region at the epileptic seizure based on SVM to detect the onset of seizure activity in EEG data. The SVM is trained on sets of intracranial EEG recordings from patients suffering from pharmacoresistant focal-onset epilepsy. The performance of SVM is measured by using accuracy obtained from a fit between the target value and network output. Our EEG based localization of the focus region at the epileptic seizure approach achieves 97.4% accuracy with using 10-fold cross validation. Therefore, our method can be successfully applied to localization of the epileptogenic foci.

1. INTRODUCTION

Common chronic neurological disorder is called epilepsy. Characteristics of epilepsy are the repeated unprovoked seizures. Due to abnormal, excessive or synchronous neuronal action in the brain these seizures are transient symptoms [5]. The noninvasive localization of the early seizure discharge is the significant problem for neurosurgery. Determining the area of the brain which is the source of that abnormal activity is also part of that [1].The epileptic focus is determined theoretically and confirmed by comparison with the clinical diagnosis and surgical results according to the consistent location of those channels that produced abnormal signals. Because of that, a crucial approach to evaluate epilepsy is EEG. The EEG is an essential means of detection and diagnosis while the diagnosis of epilepsy depends on discrete medical history.The function of EEG is additional prominent and sometimes plays a conclusive part especially when taking into account the difficulties of clinical diagnosis of a characteristic seizures. Seizure detection activity in the pre-seizure component of a mutli-electrode EEG records is crucial in the localization of epileptogenic foci. We can use only for seizure detection and prediction most of the parametric and nonparametric methods applied to ictal EEG analysis [2, 3, 4, 5, 6], since they don't allow the detection of activity.Offensive observing of seizures with subdural grids or depth electrode arrays can assist in seizure localization. Seizure sources cannot always be completely determined by visual analysis of EEG recordings even with these methods. The spread of activity is of excessive importance for localization of the epileptic focus on the foundation of EEG analysis.Recently in the localization of active areas of the brain have been applied the blind source separation and deconvolution techniques [1]. Intracranial electroencephalographic (EEG) recordings from patients suffering from pharmacoresistant focal-onset epilepsy are used. As fragment of the epilepsy diagnostics these recordings had been performed prior to and independent from our study. In order to realize seizure control the

clinical determination of these recordings was to define the brain areas to be surgically removed in each specific patient. According to this we can define two different sets of signals: All those channels that noticed first ictal EEG signal changes we define as "focal EEG channels" and "nonfocal EEG signals" as set of signals recorded from brain areas that were not involved at seizure onset. For both the randomness and the nonlinear-independence test we catch more rejections for focal vs. nonfocal signals [7]. It has been proposed that ''when strictly localized unifocal interictal spike activity can be clearly defined on a surface EEG, a high rate of ictal origin and good surgical outcome prediction can be achieved'' (Holmes et al., 2000). Actually the definition of spike focality relies on inadequate information that has been addressed through the use of computer analysis of EEG sources (Ebersole and Hawes-Ebersole, 2007). Over the last two decades EEG source localization (ESL) of interictal spikes in epilepsy surgery has been studied expansively (Ebersole and Wade, 1990, 1991; Ebersole, 1994; Merlet and Gotman, 2001; Ochi et al., 2000; Michel et al., 2004; Bast et al., 2004, 2006; Ray et al., 2007; Plummer et al., 2007, 2010).Even so, in most routine performs these methods have not been approved; and are still observed as just "holding promise" in a current review of the practicality of interictal EEG in epilepsy surgery (Dworetzky and Reinsberger, 2011). We still need clinical consequences studies for indication based-medicine assessments.

In this paper we applied the multi-scale principal component analysis (MSPCA) for de-noising and AR method for the extraction of the features, on the basis of which the machine learning methods localizes the epileptic focus region in the brain. Support vector machines (SVM) is used to realize the imperative changes of the flow activity of each channel in the time span of the seizure onset. The electrode of its greatest change of activity at the time of the seizure incidence leads to the region of the brain, containing the epileptic focus. The experimental results of localization done for epileptic patients have been compared to their real location in the brain which is determined by the medical neurosurgery. This comparison has shown close coincidence of the increased activity of the focal electrode with the location of the real epileptic focus in the brain.

2. MATERIALS AND METHODS

2.1. *Dataset*

In this paper, we used intracranial EEG recordings from five epilepsy patients described in REF. These recordings were performed as part of the epilepsy diagnostics. All five patients, being the candidates for epilepsy surgery, had longstanding pharmacoresistant temporal lobe epilepsy. Each patient went through long-term intracranial EEG recordings at the Department of Neurology of the University of Bern because noninvasive studies did not result in localization of the seizure onset zone or epileptic focus. We used intracranial strip and depth electrodes multichannel for recording EEG signals. Extracranial reference electrode was placed between 10/20 positions Fz and Pz. Those parts are manufactured by AD-TECH (Racine, WI, USA). Depending on whether they were recorded with more or less than 64 channels, EEG signals that we use were tested at 512 or 1024 Hz. The brain areas where seizures started could be contained for all five patients based on these intracranial EEG recordings. In fact, we found these areas in parts of the brain that could be surgically resected. This is without the risk of neurological deficits that would be unacceptable for the patients. Surgical outcome is good for all five patients. Three of five patients attained complete seizure freedom. Corresponding to class 1 and 2 according to the "International League Against Epilepsy" classification of surgical outcome. Two patients had auras only. They didn't have other seizures following surgery. The ethics committee of the Kanton of Bern retrospective approves EEG data analysis. In fact, all patients gave written acceptance that their data from long-term EEG can be used for research purposes. Using a fourth-order Butterworth filter all EEG signals were digitally band-pass filtered in the range 0.5 to 150 Hz. To minimize phase distortions we used forward and backward filtering. Those EEG signals that had been recorded with a sampling rate of 1024 Hz were down-sampled to 512 Hz prior to further analysis. Then we re-referenced EEG signals against the median of all the channels that is judged by visual inspection free of artifacts that is permanent. On general estates it is impossible to find reference that can be considered as "best". The impact of six common EEG references Rummel et al. [8] investigated on bi-and multivariate correlation actions. We used this for the sample of scalp montages. In a controlled way correlation is obviously introduced by the global average. In this publication we are not examine the median reference. Though, it has the extra advantage that the rank of the correlation matrix remains full as compared to the mean. In significance, the amount of global average reference is even bigger than artificially introduced correlation. All those channels that noticed first ictal EEG signal changes we define as "focal EEG channels". More than two neurologists who are also board-certified electroencephalographers judge those changes by visual inspection. One of the specialists was always KS. A fully objective is not though visual analysis, reducing subjective interpretation permitted by joint-analysis with fellow neurologists. Additionally, the most important method for clinical decision making is still visual EEG analysis. "Nonfocal EEG channels" included all other channels in the recordings were classified. From the pool of all signals measured, at focal EEG channels randomly selected 3750 pairs of concurrently recorded signals x and y. For that purpose, at first, corresponding to 10 240 samples, we separated the recordings into time windows of 20 seconds. We excluded recordings of seizure activity and three hours after the last seizure. After that, for each individual signal pair we randomly chose one

of the five patients. One of this patient's focal EEG channels and one of this channel's neighboring focal channels and in this patient's recordings one time window included. The signal pair was visually reviewed, after this we included into the database. In instance, the signal pair was discarded, it contained prominent measurement artifacts. Clinical selection standards such as the presence or absence of epileptiform activity is not applied. Lastly, in the order in which they were drawn, the focal EEG signal pairs were kept. In the same way 3750 pairs of nonfocal signals randomly selected and measured at non-focal EEG channels.

2.2. Multiscale Principal Component Analysis (MSPCA)

Principal Component Analysis (PCA) transforms an n x p data matrix, **X** by computing the variables as a linear weighted sum as,

$$X = TP^T \quad (1)$$

where, **P** are the principal component loadings, **T** are the principal component scores, n and p are the number of measurements and variables individually. The eigenvectors of the covariance matrix of **X** are the loadings, and the eigenvalues specify the variance taken by the related eigenvector. The data matrix can be decomposed by singular value decomposition as,

$$X = U\Lambda^{1/2}V$$

where, Λ is a diagonal matrix of the eigenvalues, $P^T = V$, and $T = U\Lambda^{1/2}$.

A significant conclusion in PCA is to choice the suitable number of principal components that capture the fundamental relationship, while removing the errors. For this task are available several techniques, and are reviewed by Jackson (1991) and Malinowski (1991). Methods such as the scree test, parallel analysis or crossvalidation may be used if an estimation of the error is not obtainable.

Multiscale PCA (MSPCA) combines the capability of PCA to separate the crosscorrelation or relationship between the variables, with that of orthonormal wavelets to extract deterministic features from stochastic processes and approximately decorrelate the autocorrelation among the measurements. The measurements for each variable (column) are decomposed to its wavelet coefficients using the same orthonormal wavelet for each variable to achieve the benefits of PCA and wavelets.

2.3. AR Burg method for spectral analysis

The model-based methods are based on modeling the data sequence $x(n)$ as the output of a linear system described by a rational system. We have to steps in the model-based methods for the spectrum estimate procedure. From given data sequence $x(n)$, $0 \le n \le N-1$ the parameters of the method are estimated. The PSD estimate is computed after this from these estimates. The most commonly used parametric method is AR method. With this method by solving linear equations, estimation of the AR parameters can be done without difficulty. We can model data can as output of a causal, all-pole, discrete filter whose input is white noise in the AR method. The AR method of order p is expressed as the following equation:

$$x(n) = -\sum_{k=1}^{p} a(k)x(n-k) + w(n) \quad (2)$$

where $a(k)$ are the AR coefficients and $w(n)$ is white noise of variance equal to σ^2.

The AR(p) model can be described by the AR parameters $\{a[1], a[2], \ldots, a[p], \sigma^2\}$. The PSD is

$$P_{AR}(f) = \frac{\sigma^2}{|A(f)|^2} \quad (3)$$

where

$$A(f) = 1 + a_1 e^{-j2\pi f} + \ldots + a_p e^{-j2\pi fp} \quad (4)$$

We must take into consideration some factors, such as selection of the optimum estimation method, selection of the model order, the length of the signal which will be modeled, and the level of stationary of the data to obtain stable and high performance AR method [8, 9, 10, 11].

The method developed by Burg for AR parameter estimation is based on the minimization of the forward and backward prediction errors and on calculation of the reflection coefficient. The forward and backward prediction errors for a *p*th-order model can be defined as

$$\hat{e}_{f,p}(n) = x(n) + \sum_{i=1}^{p} \hat{a}_{p,i} \; x(n-i). \cdots\cdots n = p+1, \ldots N \quad (5)$$

$$\hat{e}_{b,p}(n) = x(n-p) + \sum_{i=1}^{p} \hat{a}^*_{p,i} . \; x(n-p+i). \cdots\cdots n = p+1, \ldots N \quad (6)$$

The AR parameters related to the reflection coefficient $\hat{k}_{p,}$ are given by

$$\hat{a}_{p,i} = \begin{cases} \hat{a}_{p-1,i} + \hat{k}_p \hat{a}^*_{p-1,p-i,} & i = 1 \ldots\ldots\ldots p-1 \\ \hat{k}_{p,} & i = p \end{cases} \quad (7)$$

From the estimates of the AR parameters, PSD estimation is defined as

$$\hat{P}_{BURG}(f) = \frac{\hat{e}_p}{\left|1 + \sum_{k=1}^{p} \hat{a}_p(k) e^{-j2\pi fk}\right|^2} \quad (8)$$

where$\hat{e}_p=\hat{e}_{f,p}+\hat{e}_{b,p}$ is the total least-squares error [8, 9, 11, 12, 13, 14].

2.4. Support Vector Machines (SVM)

Vapnik introduced SVMs in 1992 as a new supervised machine learning formulation useful to the well-known problem of function estimation (Vapnik, 2000). It was first presented as a binary classifier established on statistical learning theory and structural risk minimization (SRM) theory developed to solve pattern recognition problems by Vapnik since the 1960s. It was future protracted to general nonlinear problems where mapped data to a feature space by an implicit nonlinear mapping. The framework to regression for estimating real-valued functions extended in 1995. The basic idea behind the SVMs is to maximize the hyperplane margin to get respectable classifier performance. We express a nonlinear mapping from input space to higher dimensional feature space, if training set is not linearly separable. In expressions of the Lagrangian variables the SVM problems are constrained quadratic programs. Actually, any optimization technique appropriate for solving these kinds of problems can be used, for example, projected gradient descent, projected Newton method, and conjugate gradients (Fletcher, 1981; Gill et al., 1981). Because data sets can be very large, problems with computational memory invariably surface according to the storage requirements for the dense kernel matrix. The decomposition method has commonly been usefull to solve the SVM by decomposing the key problem into smaller subproblems. With the available computer memory those problems can be solved. The subproblems are categorized by their variables often denoted to as blocks, chunks, or working sets. Decomposition algorithms operating by sequentially solving subproblems until all KKT situations of the problem have been satisfied or until some stopping standard has been met (Lin, 2002).

3. RESULTS AND DISCUSSION

In this study, we used EEG signals taken from epileptic patients in order to determine the focus region at the epileptic seizure. Intracranial EEG recordings from patients suffering from pharmacoresistant focal-onset epilepsy are used. We have been investigating a localization of the focus region at the epileptic seizure based on SVM classifier to detect the onset of seizure activity in EEG data. EEG recordings are first pre-processed with the multi-scale PCA in order to reduce the noise level. After de-noising, autoregressive (AR) Burg method is used to extract features. The formed feature vector is processed with an SVM for localization of the focus region at the epileptic seizure using intracranial EEG recordings. The visual analysis of the EEG signals does not provide the solution to the localization of the epileptic focal region, as it is difficult to notify in a visual analysis any substantial increase of the equivalent flow at the seizure onset [1].

But our proposed method based on the application of the MSPCA de-noising with **SVM** is able to analyse as many patterns of signals as needed. On the basis of MSPCA de-noising of EEG signals we have formed the feature vectors by using AR Burg modelling and used them in learning and then testing the **SVM** classifier.

Table 1 shows the high de-nosing performance on an EEG signal before and after the multi-scale PCA is applied to the for focal and non-focal EEG time series. From table 1, it can clearly be seen that the rate of correctly classified cases increases significantly when applying the MSPCA, compared to utilising only the raw original EEG signals in the classification task. The MSPCA increases the classification performance nearly 20% when distinguishing between focal and non-focal signals. This improvement is a demonstration of the MSPCA ability to de-noise the EEG time series. A true classification rate of 96.4% focal and 96% for non-focal are achieved as well as 96.2% for the total accuracy measurement. This demonstrates that the filtering properties of the multi-scale PCA itself can produce sufficient results. Based on the results of the present study and experience in the EEG time series classification problem, we would like to put emphasis on the following:

1. There is an improvement in terms of classification accuracy when the MSPCA is applied to de-noise EEG signals. This indicates that the MSPCA is a suitable de-noising method for localization of the focus region at the epileptic seizure.
2. The classification results indicated that the SVM with the MSPCA has considerable success in the EEG time series classification. The proposed MSPCA de-noising approach with an SVM classifier can be used in classification of EEG signals for localization of the focus region at the epileptic seizure.
3. The computational complexity of the MSPCA is in the range of O(log n), when the algorithm is implemented in parallel [15]. This makes the de-nosing step very efficient with regards to computational complexity. With highly competitive classification accuracies, the MSPCA is a well suited technique compared to other de-noising methods.

Table 1: Performance of the classification task

Performance of the classification task			
	Accuracy (focal)	Accuracy (non-focal)	Overall Accuracy
PBURG + SVM	71,7	78,4	75,1
MPCA + PBURG + SVM	96,4	96	96,2

4. CONCLUSION

In this study we presented the new method for the localization of the epileptic focal region by applying the MSPCA de-noising. Features were calculated by using AR Burg method to represent their frequency distribution. Features extracted from EEG time series were used with SVM classifier and cross-compared in terms of their

accuracy relative to the patterns. The comparisons were based on scalar performance measures (which is accuracy) derived from the confusion matrix. On the basis of the performed numerical experiments we have found that the highest accuracy occurs for the region of the brain containing the epileptic focal region. This observation may be used for supporting the localization of the epileptic focal region in the brain. These patients were of well-defined localization of the focus of epilepsy, determined on the basis of comprehensive neurological examination. The result of EEG signal classification using SVM classifier shows that MSPCA signal de-noising improve the performance of classifier. Consequently, the proposed method may be of potential application for determining the region of the epileptic focus in the brain on the basis of the analysis of the EEG signals from the scalp.

REFERENCES

B. Swiderski, S. Osowski, A. Cichocki and A. Rysz, "Single-Class SVM Classifier for Localization of Epileptic Focus on the Basis of EEG," in *International Joint Conference on Neural Networks*, 2008.

J. Corsini, L. Shoker, S. Sanei and G. Alarcon, "Epileptic seizure predictability from scalp EEG incorporating constrained blind source separation," *IEEE Tr. Biomed. Eng.,* vol. 56, pp. 790-799, 2006.

S. Feldt, H. Osterhage, F. Mormann, K. Lehnertz and M. Zochowski, "Internetwork and intra-network communications during bursting dynamics: applications to seizure prediction," *Phys Rev E,* vol. 76:021920, 2007.

D. L. Iasemidis, "Epileptic seizure prediction and control," *IEEE Trans. Biomed. Eng.,* vol. 50, pp. 549-558, 2003.

K. Lehnertz, F. Mormann, H. Osterhage, A. Müller, A. Chernihovskyi, M. Staniek, D. Krug, S. Bialonski and C. Elger, "State-of-the-art of seizure prediction," *J Clin Neurophysiol,* no. 24, pp. 147-153, 147-153.

B. Swiderski, S. Osowski, A. Cichocki and A. Rysz, "Epileptic seizure characterization by Lyapunov exponent of EEG signal," *Compel,* no. 26, pp. 1276-1287, 2007.

A. G. Ralph, S. Kaspar and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic," *PHYSICAL REVIEW ,* vol. E 86, 2012.

H. Adeli, Z. Zhou and N. Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform," *Journal of Neuroscience Methods,* no. 123, pp. 69-87, 2003.

A. Cohen, Biomedical Signals: Origin and Dynamic Characteristics; Frequency-Domain Analysis: Second Edition, Boca Raton: CRC Press LLC, 2000.

I. Guler, M. K. Kiymik, M. Akin and A. Alkan, "AR spectral analysis of EEG signals by using maximum likelihood estimation," *Computers in Biology and Medicine,* no. 31, pp. 441-450, 2001.

M. Akin and M. K. Kiymik, "Application of Periodogram and AR Spectral Analysis to EEG Signals," *Journal of Medical Systems,* no. 24(4), pp. 247-256, 2000.

J. Muthuswamy and N. V. Thakor, "Spectral analysis methods for neurological signals," *Journal of Neuroscience Methods,* no. 83, p. 1–14, 1998.

J. Pardey, S. Roberts and L. Tarassenko, "A review of parametric modelling techniques for EEG analysis," *Medical Eng. and Physics,* no. 18(1), pp. 2-11, 1996.

S. Goto, M. Nakamura and K. Uosaki, "On-line spectral estimation of nonstationary time series based on AR model parameter estimation and order selection with a forgetting factor," *IEEE Trans on Signal Process.,* no. 43, p. 1519–1522, 1995.

A. Subasi, E. Erçelebi, A. Alkan and E. Koklukaya, "Comparison of subspace-based methods with AR parametric methods in epileptic seizure detection," *Computers in Biology and Medicine,* vol. 36, no. 2, p. 195–208, 2006.

B. R. Bakshi, "Multiscale PCA with Application to Multivariate Statistical Process Monitoring," *AIChE Journal,* 1998.

B. Rezaul , L. T. H. Daniel and P. Marimuthu , Computational Intelligence in Biomedical Engineering, CRC Press.

T. Xiaoying , X. Li , L. Weifeng , P. Yuhua , C. Duanduan , G. Tianxin and Z. Yanjun , "Analysis of Frequency Domain of EEG Signals in Clinical Location of Epileptic Focus," *Clin EEG Neurosci,* 2012.

C.-C. E. Pedro, W. Y. Joyce , C. L. Leo , S. Kirk , D. Sandy and N. R. Marc , "Quantification and localization of EEG interictal spike activity in patients with surgically removed epileptogenic foci," *Clinical Neurophysiology,* vol. 123, p. 471–485, 2012.

B. R. Robyn , S. Mingui, S. L. Mark and S. J. Robert , "Detection of Seizure Foci by Recurrent Neural Networks," in *22‴ Annual EMBS International Conference*, Chicago IL., 2000.