CILAMCE 2017
Florianópolis - Santa Catarina - Brazil

LEARNING EFFICIENT CODES FROM EPILEPTIC EEG

Letícia Correia

leticia.correia@ifma.edu.br

Federal Institute of Maranhão

São José de Ribamar, Maranhão, Brazil

Gean Carlos Sousa

Carlos Magno Sousa

Marta Barreiros

Allan Kardec Barros

Federal University of Maranhão

São Luís, Maranhão, Brazil

Noboru Ohnishi

Nagoya University

Nagoya, Japan

Fausto Lucena

CEUMA University

Imperatriz, Maranhão, Brazil

Abstract. *Epilepsy is a neurological condition characterized by recurrent seizures, caused by brief disturbances in electrical functions of the brain. It affects people in all age groups, regardless of gender, race or financial condition. Several studies are developed in the neuroscience field aiming the most diverse aspects of epilepsy, from its causes to its treatment. This work analysis the electroencephalography (EEG) signal based on a theoretical viewpoint from a strategy called redundancy reduction. Our results show that the efficient coding might play a role on the neural genesis of epileptic seizures. The results show that efficient coding is able to characterize and differentiate the interictal and preictal stages in the Electroencephalogram of an epileptic patient.*

Keywords: *Epilepsy, seizures, efficient coding.*

INTRODUCTION

According to the World Health Organization, epilepsy is one of the most common among the serious neurological diseases; it is classified as a recurrent seizure disorder and affects about 50 million people worldwide. Epileptic people may suffer all kinds of prejudice and stigma because of a lack of knowledge about the disease, facing difficulties in education, the labor market, marriage, and other areas of personal life.

The electroencephalography (EEG) is a record of the electrical activity of the brain, used by the diagnosis of epilepsy, that can be done by electrodes implanted in the scalp or directly into the brain. According to Jouny and Bergey (2012), the characteristic bands of the brain waves are in low frequency frequency δ: [0-4 Hz [, θ: [4-8 Hz [, α: [8-12 Hz, β: [12-30 Hz [, γ: [30-58 Hz [.

The Intracranial EEG (iEEG), used in this paper, provides information on which region of the brain abnormal activities are occurring. Many studies use externally implanted electrodes. However, Lachaux et. al. (2003) in his work suggests they may be contaminated by artifacts such as: eye movements or muscle activity.

According Javed et. al. (2015), it is possible to distingue four stages in the electrical brain activity of an epileptic patient:

(a) Interictal – period between seizures, in which there is no seizure activity;

(b) Preictal – the period immediately preceding a seizure;

(c) Ictal – it is the period where most of the seizure activity is present;

(d) Postictal – Short time period just after a seizure. It is characterized by the recovery, a moment when the patient is not fully recovery in it consciousness and attention.

Based on the study of the stages found in the EEG signal of an epileptic patient, many researches are done to detect the onset of a seizure. Several models have been proposed for detection of convulsions in iEEG signals: models based on the Hilbert and wavelet transform proposed by Polat and Ozerdem (2016); Models that use only statistical data of the signal: standard deviation, skewness and kurtosis proposed by Behara et. al. (2016); Systems based on discrete wavelet transform according Sharmila and Geethanjali (2016); Using the multi-level wavelet decomposition proposed by Bugeja et. al. (2016) and etc.

All cited approaches are based on extraction and combination of electroencephalogram signal characteristics intending to improve the accuracy of a classifier. Therefore, the detection of a crisis becomes dependent on the number of characteristics and the classifier chosen. Some of these characteristics do not have any physiological significance or feasible explanation, which makes it difficult to understand how they may aid in a possible diagnosis of the crisis.

Efficient coding hypothesis proposed by Horace Barlow in 1961 suggests that neurons are adapted to reduce the statistical redundancy of the signals to which they are subject.

In other words, neurons have evolved in order to minimize the statistical dependencies of neural responses and in order to maximize information transmission capacity according Barlow (1961).

Olshausen and Field (1996) have shown that computational models based on efficient coding principles are able to predict receptive fields properties of cells in the visual cortex and cochlear neurons (auditory system).

Lewicki (2002) applied efficient coding to speech, animal vocalizations, and environmental sounds. His work suggests that the auditory system also evolved to make efficient use of all hearing aids.

Lucena et. al. (2011) assumed that the heart efficiently decodes stimuli of the central nervous system. Their results suggest that this system works according to the principles of information theory, minimizing statistical redundancies.

In efficient coding the set of encoder filters can be found using Independent Component Analysis (ICA), according to Comon (1994). In a classical work, Hyvarinen et. al. (2001) suggest that this method searches for characteristic filters, also called base functions, that transform a set of observed data into a set of elements whose components are considered statistically independent.

In this context, this work proposes the use of efficient coding, through the ICA using iEEG signals from interictal and preictal periods.

1 PROPOSED METHOD

The method proposed in this work is described by the block diagram in Fig. 1, which consists of the following steps:

1. Filtering the iEEG signals to remove artifacts and noise: we used a Notch filter in 50 Hz.
2. Signal segmentation at periods corresponding to the stationary of the iEEG, in this case one second is used.
3. Applying the FastICA algorithm to the windowed data for extraction of the base functions. We used 1,000, 5,000 and 10,000 iterations.
4. Realization of approximation in the bases functions with Gabor wavelets. For this, we used a MMQ algorithm.
5. Analysis of the characteristics obtained from the interictal and preictal signals using time-frequency analysis.

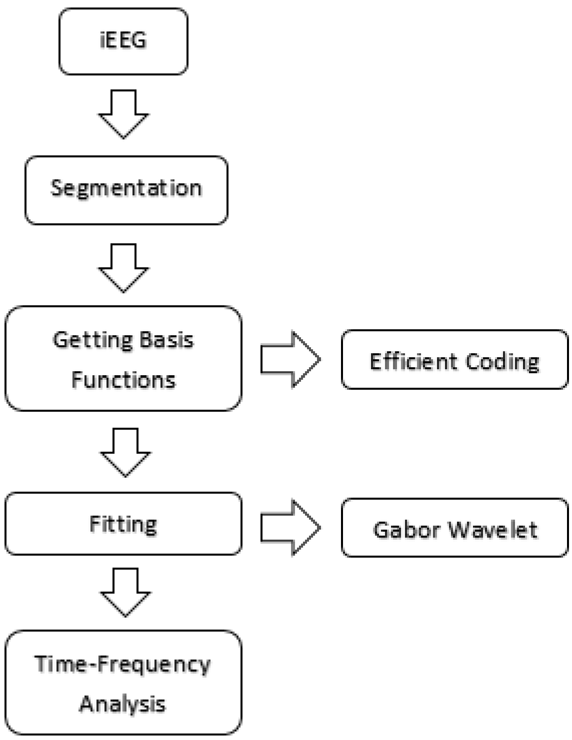

Figure. 1. Flowchart of the proposed method divided by its stages.

1.1 Database

The database used in this work was obtained from the NIH-sponsored International Epilepsy Electrophysiology site (https://www.ieeg.org/). This base is composed of several signs of Electroencephalogram obtained in various centers of research focused on the study of epilepsy.

In this work, we selected IEEG signals from the interictal and preictal periods with a sampling rate of 500 Hz. In total, we used the corresponding 15-hour interictal signal and 6-hour preictal signal of two patients. A segment of one second of each set is represented in Fig. 2.

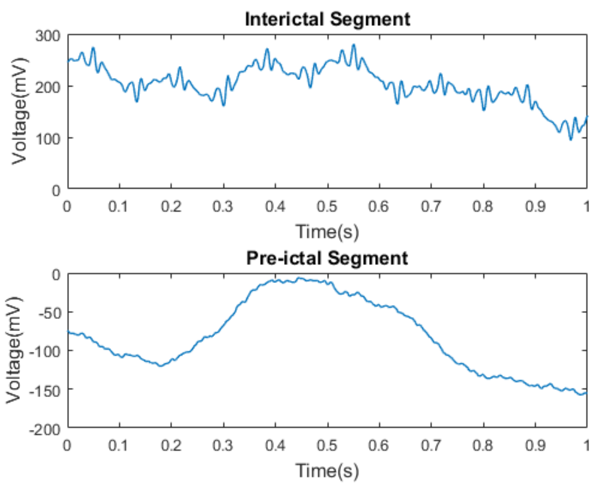

Figure. 2. One-second sample of the signal: (up) interictal and (down) preictal.

1.2 Efficient Coding

According to information theory, there are two main hypotheses for model construction: simple codes, which presuppose the choice of the simplest and consistent explanation with the data with the least possible assumptions. In his work, Softky (1995), suggests that efficient codes assume that organisms are optimized to improve their performance.

The efficient coding hypothesis supposes sparsity in the activation of the neurons. Spasm is directly related to statistical independence, so that when one increases the other decreases. A sparse model, according Oliveira (2012), for a random vector x can be written as a linear combination of a set of base functions A and a set of sparse coefficients:

$$x = \sum_i a_i s_i \tag{1}$$

The idea of reducing the statistical dependence between vectors is one of the features of ICA. ICA is a method to find multidimensional statistical data elements or components, so that these components are simultaneously statistically independent and not Gaussian according Hyvarinen et. al. (2001). Using the standard ICA representation in Eq. (1), it is obtained as:

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s}. \tag{2}$$

Assuming A invertible, the solution s of the above system is given by:

$$\boldsymbol{s} = \boldsymbol{W}\boldsymbol{x} \tag{3}$$

where W is the matrix to be defined and can be writen as $\boldsymbol{W} = \boldsymbol{A}^{-1}$. However, it is the A matrix the one that matters to the efficient coding, it is estudy that will provide the characteristics sought in the signs object of this article. The matrix A is a neural population and its columns are basic functions to be used for the study proposed in this work.

1.3 Algorithm FastICA

Proposed by Hyvarinen et. al. (2001), FastICA is used to determine matrices A and S using a negentropy approach to maximize the non-Gaussianity of S. A detailed description of FastICA is found in Gotman (2010) described as follows [14]:

1. Centralize the observed data vector x by making it a zero-mean vector:

$$x' = x - m_x. \quad (4)$$

2. Multiply the vector by a whittening matrix, obtaining a new vector, which is given by:

$$z = Vx'. \quad (5)$$

3. Choose the number m of codes to be estimaded.

4. Randomly initialize the matrix W.

5. Update the components of the matrix W, using the following equation:

$$w_i = E\{z\varphi(w_i^T z)\} - E\{z\varphi'(w_i^T z)\}w_i \quad (6)$$

where $\varphi(.)$ is the hyperbolic tangent.

6. Normalize W using:

$$W = (WW^T)^{-1/2}W. \quad (7)$$

7. Verify the convergence of W.

In this paper, FastICA was tested using three different numbers of iterations to evaluate the accuracy of the results. We used 1,000, 5,000 and 10,000 iterations each producing its own matrix A.

1.4 Fitting

The basic functions obtained with FastICA were approximated by Gabor wavelets, whose mathematical expression is given in 8. Gabor is selected by visual inspection, the literature, Lewicki (2002) and Lucena (2011), related in their results also find Gabor functions for biological signals. This function is expressed as a cosine modulated by a Gaussian envelope according Durka et. al. (2001):

$$v(t) = e^{-\pi\left(\frac{t-t_0}{s}\right)} \cos(\omega(t - t_0) + \gamma). \quad (8)$$

where t_0 is the Gaussian envelope mean, s is envelope standard deviation, ω and γ are the fundamental frequency and phase of the cosine, respectively.

Using of least squares algorithm, FastICA base functions were approximated by Gabor wavelets. Observing Eq. (1), we have:

$$x(t) = \sum_i v_i(t,\omega)s_i \quad , \tag{9}$$

in other words, after the fitting there is an expression in which a signal as a function of time can be analyzed in the time domain and the frequency simultaneously.

1.5 Time-frequency Analysis

In sequence to the fitting performed with the bases functions, part is taken for the time-frequency analysis of the Gabor functions obtained. Time-frequency analysis is a technique that consists of studying a signal in time and frequency simultaneously. Considering that there is relation in the temporal behavior and the frequency behavior of a signal, these properties can be better understood if analyzed together.

One of the possible ways to obtain the time-frequency diagram of a signal is through the Hilbert Transform, it can be found in Lucena (2011). The Hilbert transform of a signal $\varphi(t)$, denoted by $\mathrm{H}[\varphi(t)],$ gets the imaginary part of the signal:

$$A[\varphi] = \varphi(t) + iy(t) . \tag{10}$$

That is given by:

$$y(t) = \mathrm{H}[\varphi(t)] = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{\varphi(t)}{t-\tau}d\tau . \tag{11}$$

The envelope of the signal is expressed by:

$$A[\varphi] = \varphi(t) + iy(t) . \tag{12}$$

2 RESULTS

Using the iEEG signals, from the base described above, windowed in periods of one second corresponding to the stationarity period of the iEEG, the base functions of the interictal and preictal periods were obtained using FastICA.

Some base functions of the interictal period are shown in Fig.3. Frequency-time diagram of the base functions obtained by the use of interictal signals and preictal signals are shown in Fig. 4(up) and Fig. 4(down), respectively.

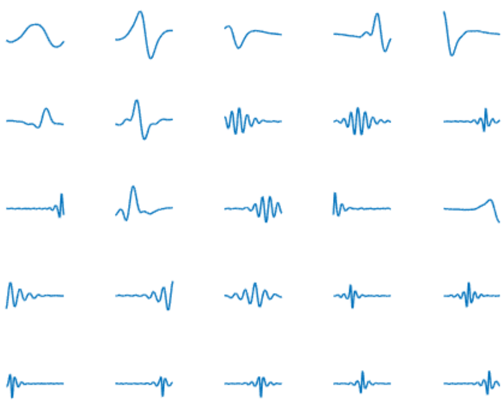

Fig. 3. Base functions obtained by FastICA to interictal period sign.

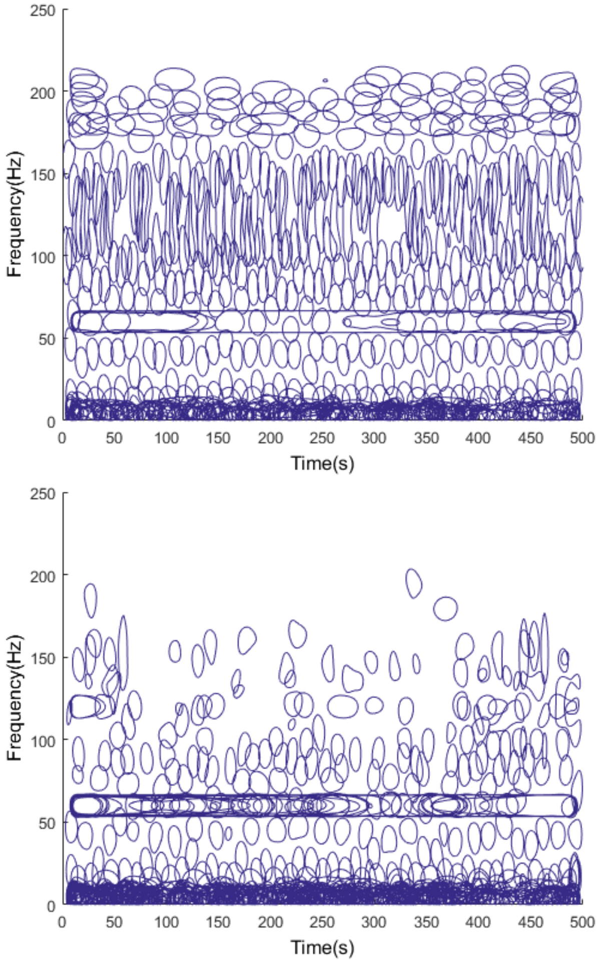

Fig. 4. Time-frequency diagram of the basic function set of: (up) interictal and (down) preictal.

In order to better represent the functions in the time-frequency space, the base functions were approximated by Gabor’s functions. Some samples of this approximation may be seen in Fig.5. Some of the base functions of the original set were discarded because they were not localized in time and frequency. The energy of these base functions is not concentered in a single peak, so they do not represent filters.

Fig. 6. Base functions (in blue) approached by Gabor wavelets (magenta).

After fitting, the relation between bandwidth and center frequency was made graphically, to improve the understanding of the high frequencies states. Fig. 6 (up) shows interictal set and in Fig. 6 (down) preictal period. In Fig. 6 the characteristic frequencies of the iEEG signal (high and low frequency) and the very high frequencies encountered by fastICA are shown.

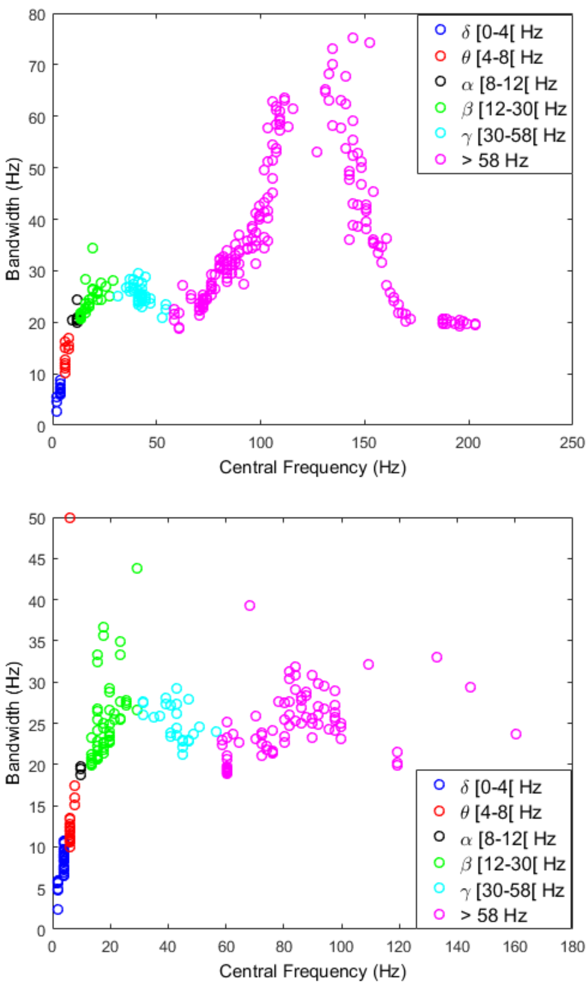

Fig. 6. Relation between the bandwidth and the center frequency for the approximate bases functions of: (up) preictal set and (down) interictal set.

3 DISCUSSION

In this paper, a model was presented for the study of seizures characteristics in an epileptic patient based on the efficient coding of electroencephalogram signals, of the interictal and preictal periods.

By analyzing the time-frequency diagrams of the interictal and preictal periods, shown in Fig. 4, it is noticed that there is a significant difference between the two sets. All signals of the pacients analyzed presents frequencies above 58 Hz in the interictal period. These results suggest that the very high frequencies (above 58 Hz) are present in a larger number and with more regularity in the interictal period. This difference is emphasized when we observe the frequency range above 100 Hz.

At low and high frequencies (below 58 Hz), the bandwidth vs. center frequency ratio, presented in Fig. 6, is similar for the two sets. However, for central frequencies greater than 58 Hz, Fig. 6. confirms the conclusion obtained in the analysis of Fig. 5.

The presented results suggest that the hypothesis raised by Gotman (2012) that high-frequency oscillations can be used as a marker for epilepsy. According Andrade-Valena et. al. (2012) and Jacobs et. al. (2008) the high frequencies are related to the interictal period and may be an indicator of the cerebral epileptic zone.

4 CONCLUSION

Our model demonstrated that the characteristic frequencies of the iEEG signal in the interictal period are represented at high frequency, differently, these frequencies are smoothed when observed in signals of the pre-ictal period, in other words, they occur less regularly.

The results of the experiments showed that the efficient coding is able to point out differences between the signs of the interictal and preictal stages of the disease. By the analysis of the time-frequency diagram, the results suggest that the appearance of components at high frequency (above 58 Hz) is related to the interictal period as described in the literature Andrade-Valena et. al. (2012) and Jacobs et. al. (2008).

As an extension of this work, we intend to analyze the energy behavior of the iEEG in the interictal and preictal periods in order to obtain a classifier that can distinguish these two periods.

REFERENCES

________http://apps.who.int/gb/ebwha/pdf_files/EB136/B136 R8–en.pdf

Andrade-Valena, L., Mari, F., Jacobs, J., Zijlmans, M., Olivier, A., Gotman, J. & Dubeaua, F., 2012. Interictal high frequency oscillations (HFOs) in patients with focal epilepsy and normal MRI. Clinical Neurophysiology. 123(1): 100105.

Barlow, H. B., 1961. Possible principles underlying the transformation of sensory messages. In Sensory Communication (M. G. Rosemblum, ed.), pp. 217234, Cambridge, MA: MIT Press.

Behara D. S. T., Kumar A., Swami P., Panigrahi B. K. and Gandhi T. K., 2016. Detection of Epileptic Seizure Patterns in EEG through Fragmented Feature Extraction. 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 2539-2542.

Bugeja S., Garg L., and Audu E. E., 2016. A novel method of EEG data acquisition, feature extraction and feature space creation for Early Detection of Epileptic Seizures. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 837-840.

Comon P., 1994. Independent Component Analysis, a new concept? Signal Processing, Elsevier 36: 287314

Durka, P. J., Ircha, D. & Blinowska, K. J., 2001. Stochastic time-frequency dictionaries for matching pursuit. IEEE Transactions on Signal Processing pp. 507-510.

Gotman, J., 2010. High frequency oscillations: The new EEG frontier? Epilepsia. 51(0 1): 6365.

Hyvarinen, A., Karhunen, J., Oja, E., 2001. Independent component analysis. John Wiley & Sons.

Jacobs, J., LeVan, P., Chander, R., Hall, J., Dubeau, F. & Gotman, J., 2008. Interictal high-frequency oscillations (80-500 Hz) are an indicator of seizure onset areas independent of spikes in the human epileptic brain. Epilepsia; 49(11):1893-907.

Javed, M., Akhtar, A., Ahmed, I. & Faisal, R., 2015. A novel method for seizure detection In intracranial EEG recordings. In Proc 2015 International Conference on Computational Intelligence and Communication Networks, pp. 237-241

Jouny, C. C. & Bergey, G. K., 2012. Characterization of early partial seizure onset: frequency, complexity and entropy. Clinical Neurophysiology, Vol. 123, Issue 4, pp. 658-669.

Lachaux , J. Ph., Rudrauf , D. & Kahane , P., 2003. Intracranial EEG and human brain mapping. Journal of Physiology-Paris, Vol. 97, pp 613-628.

Lewicki, M. S., 2002. Efficient coding of natural sounds. Nature Neuroscience, vol. 5, no. 4, pp. 356363.

Lucena, F., Barros, A. K., Principe, J. C. & Ohnishi, N., 2011. Statistical coding and decoding of hearteat intervals. PLoS One, vol. 6, no. 6, p. e20227.

Oliveira, F. L., 2012. HRV Estimation from BP and Coding Strategy Analysis of the Cardiovascular Regulation. Doctoral Thesis, Nagoya University, Japan.

Olshausen, B. A. & Field, D. J., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, vol. 381, no. 6583, pp. 607609.

Polat H., Ozerdem M. S., 2016. Epileptic Seizure Detection from EEG Signals by Using Wavelet and Hilbert Transform. 2016 XII International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH), pp. 66-69.

Sharmila A. and Geethanjali P., 2016. DWT Based Detection of Epileptic Seizure From EEG Signals Using Naive Bayes and k-NN Classifiers. IEEE Access, vol.4, pp. 7716-7727, 2016.

Softky, W. R., 1995. Simple codes versus efficient codes. Curr Opin Neurobiol, vol. 5, no. 2,pp. 239247.