Nonparametric Statistical Analysis of EEG Signals with Wavelet Detection: Seizure Detection

Mahamat Ali Issaka

Laboratoire de Modélisation Mathématiques, Informatique, Applications et Simulations (L2MIAS) Université de N’Djamena, BP 1027, Tchad

&

Laboratoire d’Etudes et de Recherche en Statistique et Développement (LERSTAD) Université Gaston Berger, BP 234, Saint-Louis, Sénégal

This article is distributed under the Creative Commons by-nc-nd Attribution License. Copyright © 2019 Hikari Ltd.

Abstract

In this manuscript, we consider a study based on the statistical analysis of EEG signals with wavelet tools combined with segmentation criterion in order to analyze electroencephalograph (EEG) recording data collected from normal and epileptic subjects. Our study is based on the localization of epilepsy seizure with none parametric tools which give a better resolution in the frequency and time domain as the examination of EEG signals is often done with visual inspection of the rhythm (delta, theta, alpha, beta, gamma) by neurologist practitioners. The accuracy of the detection of our method is estimated to 74.31% with a sensitivity of 86.08% according to the interpretation of neurologist.

Keywords: statistical data analysis, wavelet analysis, non parametric model, time-frequency domain, EEG

Mathematics Subject Classification: 62P10, 62F03, 62-07, 68

1 Introduction

The mathematical theory of wavelets was developed by Yves Meyer [13] and many collaborators in the 1980s. However, wavelets were introduced into statistics during the late 1980s and early 1990s.
Wavelet transform is rapidly surfacing in fields as diverse as telecommunications and biology. Because of their suitability to analyze non-stationary signals, those whose statistical properties change over time, they have become a powerful framework who can provide both very good time resolution at high frequencies and good frequency resolution at low frequencies.

Nowadays, because of their long time of recording, varying twenty minutes to one hour, EEG signal has become an important clinical tool for the evaluation and treatment of epilepsy [9]. This is because the EEG can yield information about the timing and location of all kind of electrical activity.

The examination of EEG signals is often reviewed with visual inspection of the rhythm (delta, theta, alpha, beta, gamma) by neurologist practitioners and trained technician. However, an automatic method for EEG analysis can provide an efficient tool to visual analysis procedures, and offer several advantages over visual scoring [10]. The automatic method have several advantages over the visual method. For example, it results in considerable saving by highlighting likely areas of interest. An other important example is that the automatic method also makes sure that nothing of importance is missed by allowing recordings to be reviewed much more quickly, and so more patients can be treated in the time available.

As automatic methods proposed in the literature, we can cite parametric model in [1, 2, 7] which treat the detection of seizure with segmentation criterion. As non-linear and non-parametric model based essentially on wavelet analysis, we have in the literature [3, 8, 11]. Finally, we can enumerate [12] as semi-parametric method. All those proposed methods in detecting seizures are effective but generate after evaluation a high specificity. Therefore, it will be necessary to develop new non-parametric methods considering the time and frequency approach.

In this paper, we propopose a novel model based on the applications of wavelets in statistics with the linear regression in order to indicate automatically the epileptic seizures and other abnormal events in EEG. we proposed a new non-parametric model based on the localization of epilepsy seizure with non-parametric tools which give a better resolution in the frequency and time domain. The rest of the paper is organize as follows. In section 2, we first describe the data used, next revise the statistical model for wavelet methods (DWT) and finally propose our approach of detection. Section 3 is dedicated to our experimental results and evaluation of the proposed method. Finally, section 4 is dedicated to further discussions and concluding remarks.
2 Proposed method

2.1 Clinical Data and brief description

All of the EEG data used in this study were recorded in University Hospital Center of Dakar in the service of neurology. We considered ten patients among them patients suffering from epileptic discharge, and three with a normal EEG signals.

2.2 Wavelet Transform

A wavelet is, as the name suggests, a small wave. Many statistical phenomena have wavelet structure. Often small bursts of high frequency wavelets are followed by lower frequency waves or vice versa. The theory of wavelet reconstruction helps to localize and identify such accumulations of small waves and helps thus to better understand reasons for these phenomena. Wavelet theory is different from Fourier analysis and spectral theory since it is based on a local frequency representation [15].

There exist many different types of wavelet transform among them, we will distinguish

* The continuous wavelet transform
* The discrete wavelet transform

2.2.1 Continuous Wavelet Transform (CWT)

Let $f \in L^2(\mathbb{R})$ a continuous and integrable function i.e. $\int_{-\infty}^{+\infty} |f(t)|^2 dt < \infty$, then the continuous wavelet transform of the function $f$ is defined by:

$$ W_{a,b} = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{|a|}} \psi^* \left( \frac{t - b}{a} \right) dt $$

(2.1)

The parameters $a \in \mathbb{R}_+^*$ and $b \in \mathbb{R}$ are called dilation (who express the frequency) and translation (via location) parameters, respectively. The function $\psi \in L^2(\mathbb{R})$ is called mother wavelet and defined by :

$$ \psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi^* \left( \frac{t - b}{a} \right) $$

(2.2)

Therefore $\psi(.)$ is a real function satisfying two basic properties [5]

i) the integral of $\psi(.)$ is zero:

$$ \int_{-\infty}^{+\infty} \psi(t) dt = 0 $$

(2.3)
ii) the square of $\psi(.)$ integrates to unity:

$$\int_{-\infty}^{+\infty} \psi^2(t)dt = 1$$  \hspace{1cm} (2.4)

Now, as the main idea of the paper is the reconstruction of the original signal $f$ with the wavelet, it convince to set the invert of the $W_{a,b}$ which is

$$f(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_{a,b}\psi_{a,b}(t)dadb$$  \hspace{1cm} (2.5)

### 2.2.2 Discrete Wavelet Transform

The CWT places redundant information on the time-frequency plane. To overcome these deficiencies, the CWT is discretized and algorithms equivalent to the two-channel filter bank have been developed for signal representation and processing. However, here we take the discrete values of the scale parameter $a$ and the translation parameter $b$ in a different way. So the discrete form of those parameters will be:

$$a_j = 2^{-j}, b_{j,k} = k2^{-j}, \forall j, k \in \mathbb{Z}$$  \hspace{1cm} (2.6)

With these values of $a$ and $b$, the equation (2.2) becomes

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^jt - k), \forall j, k \in \mathbb{Z}$$  \hspace{1cm} (2.7)

Therefore the integral of (2.1) can be write

$$W_{2^{-j},k2^{-j}} = 2^{j} \int_{-\infty}^{+\infty} f(t)\psi(2^jt - k)dt$$  \hspace{1cm} (2.8)

Let us now discretize the function $f(t)$. For simplicity, assume the sampling rate to be 1. In that case, the integral of (2.8) can be written as

$$W_{2^{-j},k2^{-j}} \approx 2^j \sum_{n} f(n)\psi(2^jn - k)$$  \hspace{1cm} (2.9)

Finally, as the function $\psi_{j,k}(t)$ is orthogonal, the original signal can be reconstruct by using the wavelets coefficients by:

$$f(t) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} W_{j,k}(t)\psi_{j,k}(t)$$  \hspace{1cm} (2.10)
2.2.3 The wavelet series expansion

We try to define the special structure of wavelets forms which is based on the space of square integrable real functions i.e \( L^2(\mathbb{R}) \). Let the function \( f \in L^2(\mathbb{R}) \). Considering the wavelets definition of Daubechies [4] which is the most approved in the applications of wavelets in statistics, we can have two mutually orthonormal functions \( \phi \) (the scaling function) and \( \psi \) respectively called father and mother (defined above) wavelets. By translations of the \( \phi \) and dilations of \( \psi \) using the relationships, we have [6]

\[
\phi_{j_0,k}(t) = 2^{j_0/2} \psi(2^{j_0}t - k), \quad \psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), \quad j = j_0, j_0 + 1, \ldots, \quad k \in \mathbb{Z} \tag{2.11}
\]

for some fixed \( j_0 \in \mathbb{Z} \).

Given the above wavelet basis, a function \( f \in L^2(\mathbb{R}) \) is then represented in a corresponding wavelet series as

\[
f(t) = \sum_{k \in \mathbb{Z}} c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j=j_0}^{\infty} \sum_{k \in \mathbb{Z}} w_{j,k} \psi_{j,k}(t), \quad \forall t \in [0,1] \tag{2.12}
\]

with

- \( c_{j_0,k} = \langle f, \phi_{j_0,k} \rangle \): discrete scaling coefficients
- \( w_{j,k} = \langle f, \psi_{j,k} \rangle \): discrete wavelet coefficients
- \( \langle.,.\rangle \): is the standard \( L^2(\mathbb{R}) \) inner product of two functions

\[
\langle f_1, f_2 \rangle = \int_{\mathbb{R}} f_1(t)f_2(t)dt
\]

2.3 Statistical model for wavelet methods DWT

Here we revise some relatively well-established wavelet methods in statistical applications. These include nonparametric regression, density estimation, inverse problems, change-point problems and some specialized aspects of time series analysis. We consider nonparametric regression in the most detail, since the basic scheme used in that case recurs with appropriate modifications in the other applications.

2.3.1 Linear Estimation of Coefficients

Consider the standard univariate nonparametric regression model

\[
y_i = f(t_i) + \varepsilon_i, \quad \forall i = 1, \ldots, n \tag{2.13}
\]
where $\varepsilon_i \sim N(0, \sigma^2)$ are independent random variables. The goal is to recover the underlying function $f$ from the noisy data $y_i$ without assuming any particular parametric structure for $f$.

Suppose that $f$ is expanded as a wavelet series on the interval $[0, 1]$ with $j_0 = 0$. So we have:

$$f(t) = c_0 \phi(t) + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} w_{j,k} \psi_{j,k}(t)$$

(2.14)

where $c_0 = \langle f, \phi \rangle$ and $w_{j,k} = \langle f, \psi_{j,k} \rangle$.

Of course, we cannot estimate an infinite set of coefficient $w_{j,k}$ from the finite sample, so it is commonly supposed that $f$ belongs to a class of functions with a certain regularity. The corresponding norm of the sequence of $w_{j,k}$ is therefore finite and $w_{j,k}$ must decay to zero. Then, considering the Fourier standard method, it is assumed that

$$f(t) = c_0 \phi(t) + \sum_{j=0}^{L} \sum_{k=0}^{2^j-1} w_{j,k} \psi_{j,k}(t)$$

(2.15)

for some $L < J$.

Thus, a corresponding wavelet estimator of $f(t)$ is defined by

$$\hat{f}_L(t) = \hat{c}_0 \phi(t) + \sum_{j=0}^{L} \sum_{k=0}^{2^j-1} \hat{w}_{j,k} \psi_{j,k}(t)$$

(2.16)

With this form, the original nonparametric problem essentially transforms to linear regression and the sample estimates or empirical estimations of the scaling coefficient and the wavelet coefficients are given by

$$\hat{c}_0 = \frac{1}{n} \sum_{i=1}^{n} \phi(t_i) y_i$$

(2.17)

$$\hat{w}_{j,k} = \frac{1}{n} \sum_{i=1}^{n} \psi_{j,k}(t_i) y_i$$

(2.18)

### 2.4 Proposed approach of detection

Once estimate the wavelets coefficients and the scaling coefficients with the linear regression in the previous section, we proposed a model based on those coefficients with a segmentation criterion. So we can localize the epileptic area
of the brain by combination of the mean and variance of the estimated coefficients.

Let consider the estimate coefficients \( \hat{c}_0, \hat{w}_{j,k} \) \( \forall X_i, i = 1, \ldots, n \) derivation of EEG signals.
We can define the criterion:

\[
C = \max_{1 \leq i \leq n} |\hat{c}_0| - |\mu|, \quad (2.19)
\]

and

\[
D = \max_{1 \leq i \leq n} |\hat{w}_{j,k}| - |\sigma_X|, \quad (2.20)
\]

with \( \mu \) and \( \sigma_X \) are respectively the empirical mean and the standard deviation of the wavelets and scaling coefficients.
So the related segmentation criterion can be defined by

\[
1_{\hat{c}_0} = \begin{cases} 
1 & \text{if } \hat{c}_0 > C \\
0 & \text{if } \hat{c}_0 \leq C 
\end{cases}
\]

and

\[
1_{\hat{w}_{j,k}} = \begin{cases} 
1 & \text{if } \hat{w}_{j,k} > D \\
0 & \text{if } \hat{w}_{j,k} \leq D.
\end{cases}
\]

3 Experimental Results

We consider four derivations from a patient suffering to partial epilepsy on fronto central party of the brain. The derivation corresponding are the Fp2F4, F4C4, Fp1F3 and F3C3. Fig. 1 display the original EEG recordings for an unhealthy patient.
3.1 Choice of wavelet resolution level

In this party, we focus on the wavelet resolution level. By using the Daubechies wavelets, you choose the level for which the energy of the filter attend the maximal value.

In the Fig. 2, we can remark that the maximal energy is at level 13. So we use the level 13 in order to estimate the wavelets and the scaling coefficients for each derivation.
3.2 MODWT

Fig. 3 and Fig. 4 display the detection results of patient with the above approach of discrete wavelets transform associate with linear regression. Otherwise, we have estimated the wavelets and scaling coefficients of the four area of the brain concerned by the epileptic region, and we used those empirical estimate coefficients in order to detect the onset of epileptic seizure.

![Figure 3: Coefficient seizure detection](image)

With the wavelet coefficients, as showed in the Fig. 3 and 4, we have a good resolution in temporal domain very coherent. The seizure began at minute 02-06, 16-21 and 32-35. One of the most strength seizures occurs at the 02-06 and 16-21 in the epileptic derivation.

Considering the detection in scaling coefficients, we have a better representation of beginning of seizure in the frequency domain.
3.3 Model proposed

Fig. 5 display also the detection of epileptic seizure with the proposed model of detection. We have choose this model in order to select only the epileptic time of occurs. Otherwise, we can remark very clearly the beginning of the seizure at the minute 02-06 as commented in the previous section with only the estimated coefficients.
3.4 Evaluation

\[
sensitivity = \frac{\text{Correctly detected epileptic seizures}}{\text{Total number of epileptic seizures}}, \quad (3.21)
\]

\[
specificity = \frac{\text{Correctly detected normal activities}}{\text{Total number of activities}}, \quad (3.22)
\]

\[
\text{false detection rate (FDR)} = \frac{\text{Number of false detection}}{\text{Total length of the data}} \quad (3.23)
\]

In the sample, we choose three patients suffering respectively to fronto-central, centro-temporal and right anterior temporal seizure of epilepsy for which we evaluate the accuracy of the used method in this study as shown in Table 1.

Table 1: Statistical parameters classifier

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fp2F4</td>
<td>77.56</td>
<td>65.85</td>
<td>68.22</td>
</tr>
<tr>
<td>F4C4</td>
<td>85.3</td>
<td>53.5</td>
<td>66.4</td>
</tr>
<tr>
<td>Fp1F3</td>
<td>95.4</td>
<td>67.4</td>
<td>88.32</td>
</tr>
<tr>
<td>Total</td>
<td>86.08</td>
<td>62.25</td>
<td>74.31</td>
</tr>
</tbody>
</table>

4 Conclusion

This paper address a novel model based on the applications of wavelets in statistics with the linear regression in order to indicate automatically the epileptic seizures and other abnormal events in EEG. It proposed in the phase to estimate wavelets and scaling coefficients in the empirical estimation. After, by using the characterization of the frequency bands and the energy content of a signal, the model give a temporal localization as well as a better detection of short seizures and other spikes, especially with Daubechies level 13 wavelet. The proposed approach is more effective than the power spectral detection and the former analogical EEG detection. However, this model can be implemented in Micromed Machine EEG interface as statistical tool analysis for the practitioners in the Service of Neurology.

References


Received: June 19, 2019; Published: July 30, 2019