A Method Based on FFT and ACP to EEG Signal Classification

Beghdadi Hadj Ali, Senouci Mohamed
Computer science department
Oran University
Oran, Algeria
{h_beghdadi, msenouci}@yahoo.fr

ABSTRACT: *This article discusses a study to help analyze EEG signals. This study is based on two essential tools for the extraction of signal characteristics. Our tests were conducted on the basis of a 32-channel EEG acquired using the Neuroscan software. The example is referenced CZ. The EEG is sampled at 1000 Hz the main purpose of this study is to reduce the large volume of data from an EEG signal. This study is essentially based onresearching for relevant information in an EEG signal. We start from the spectral representation of the signal into a visual interpretation of the segments constituting the original signal.*

Keywords: Signal, Signal Transform, Matrix, EEG, Classification, FFT

Received: 1 December 2011, Revised 11 January 2012, Accepted 19 Janurary 2012

1. Introduction

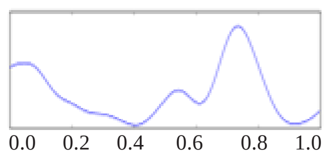
The human brain consists of a vast network of hundreds of billions of neurons whose connections and activity are complex processes and poorly understood at both the microscopic and macroscopic. This massive network is the support of brain activity that governs the overall operation of the human body [1]. Pyramidal cells of the cerebral cortex, whose major axis is perpendicular to the surface, are considered the elementary generators of EEG activity [2]. The electroencephalograph is the recording of brain electrical activity, usually undertaken by several electrodes to the scalp. EEG is to record the electrical activity of the brain using electrodes placed on the scalp surface. This system is used clinically primarily in the context of epilepsy and this is often the only system to monitor anomalies [3].The EEG records the electrical activity of cortex. Since its discovery in 1929 by Hans Berger, clinicians and researchers have used many methods to analyze the EEG signals in order to describe brain activity. The analysis of the EEG involves questions of quantification: dominant frequency, similarity, evolution, etc..on a large set of signals. But the EEG is a complex signal whose properties vary spatially and temporally. In addition to clinical approaches, two types of approaches are possible to analyze these signals: nonparametric methods, which consider the signal as a stochastic signal and parametric models, which envisages the EEG from a specific model. [4]. We approach this problem in this paper a nonparametric method that is based essentially on the theory of signal processing, which is then adapted to the analysis of EEG signals. Several studies in the fields beyond classification get the feature extraction of EEG signals were achieved. We present some work as citation:

Elif Derya Beyla (2009) presented his paper An integrated vision of automated diagnosis combined with spectral analysis techniques in the classification of electroencephalogram. The document includes detailed information on the artwork and the implementation of automated diagnostic systems and feature extraction or selection fromthe EEG signals. The main objective of this paper is to guide readers who want to develop an automated diagnostic classification of EEG signals. Ling Guo, Daniel Rivero, Juli'an Dorado, Cristian R. Munteanu, Alejandro Pazos (2011) presented their paper with rich a state of the art in this subject, they have cited the work on a: Epilepsy and the electroencephalogram (EEG), Discrete wavelet transform, Genetic programming, K-nearest neighbor classification, Previous work of genetic programming application is feature extraction [5].

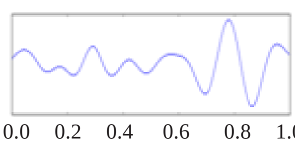
2. EEG signal frequencies

We present in this section we see some frequencies in the EEG.

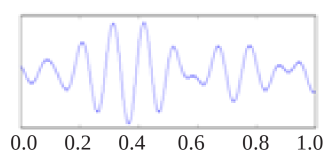
- Deltais the frequency range up to 4Hz tends to be higher in amplitude and slower waves. We see it normally in adults in the slow-wave sleep. It is also seen normally in babies.



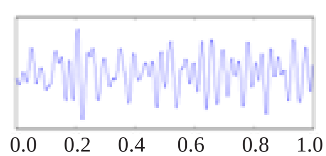
- Theta is the frequency range 4 Hz to7HzTheta is seen normally in young children.



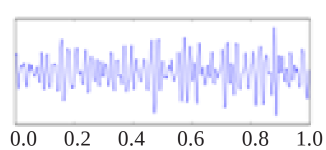
- Alpha is the frequency range 8 Hz to 12 Hz Hans Berger named the first rhythmic EEG activity he regarded as the "*alpha waves*".



- Beta is the frequency range 12 Hz to 30 Hz are often seen on both sides in symmetrical distribution and is the most obvious front.

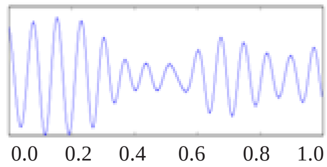


- Gamma is the frequency range 30-100 Hz. Gamma rhythms are supposed to represent the binding of different populations of neurons as well as a network in order to achieve a certain cognitive function.



- Mu range 8-13 Hz, and partly overlap with other frequencies. It reflects the - synchronous firing of motor neurons in the resting state.

$-\infty$



3. Fourier transforms

Fourier transforms analysis of the "*frequency content*" of a signal. Its many properties make it suitable for the study of stationary linear operators, including the diversion. It is a global representation of the signal.

$$\hat{f}(w) = \int_{-\infty}^{+\infty} f(t)\, e^{-wt} dt$$

The inverse Fourier transform is f as a sum of sinusoids :

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(w)\, e^{wt} dw$$

The Fourier transform will serve us to study the spectrum of the EEGsignal.

The decomposition of our EEG signal into a set of elementary signals represents our raw data.

3. Principal component analyze

In most situations, there are several observations on each signal component signals study. We therefore take into account variables p signal, p is strictly greater than 1. The separate study of each of these variables gives some information but is insufficient because it ignores the connections between them, although this is often what we want to study. It is the role of multifactor statistical data analysis as a whole, taking into account all variables. The Principal Component Analysis is then a good method to study the multi dimensional data, where all observed variables are digital, preferably in the same units, and we want to see if there are links between these variables. One starts with a rectangular array of data representing all data, by placing individuals in line (signals) and column variables.

$$X = \text{indivuals} \begin{array}{c} \\ 1 \\ \cdots \\ n \end{array} \begin{array}{c} \text{Variables} \\ \begin{array}{ccc} 1 & \ldots & p \end{array} \\ \left[\begin{array}{ccc} x_1^{\,1} & \ldots & x_p^{\,1} \\ \cdots & \cdots & \cdots \\ x_1^{\,n} & \ldots & x_p^{\,n} \end{array} \right] \end{array}$$

4. Treatment approach

Our study is focused on a sample that is part of an EEG sample of 32 treated by Canales and Neuros can referencias CZ. Figure1.

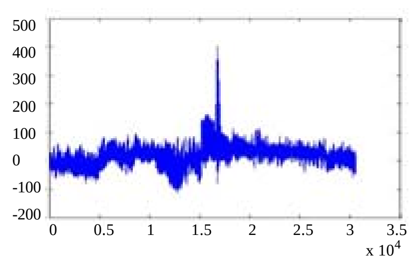


Figure 1. Figal test of chanel 1

Segmentation of the signal: the original signal will be segmented into a set of fragment. These sets a re the result of sampling. Not arbitrarily wasset experimentally at 1000 values. Figure 2 shows the first sample.

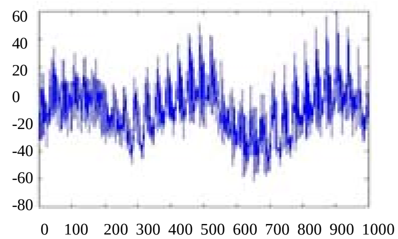


Figure 2. First segment of the signal

Fourier Transform segments: the segment will be transformed by FFT. Decomposition or spectral shape creates default 1000 sinusoids. The calculation modules: this step is to calculate the modules corresponding to different sinusoids created. Figure 3 shows the modules of different sub-signals.

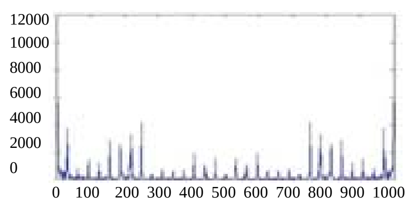


Figure 3. Modules of signals

Quantification of the modules: this step aims to quantify the level modules. The bearing is an experimental variable has been set for our case to 900. Figure 4 shows the distribution of modules for different levels.

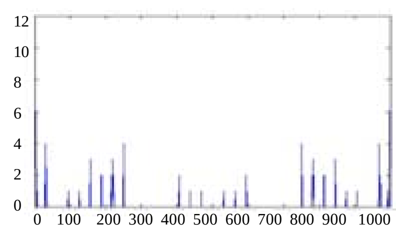


Figure 4. Quantification of modules

The aim of these treatments makes the main features for the power processed by the methods ACP. To simplify the treatment we see necessary to represent information such as a histogram. Figure 5 shows the histogram on the first fragment.

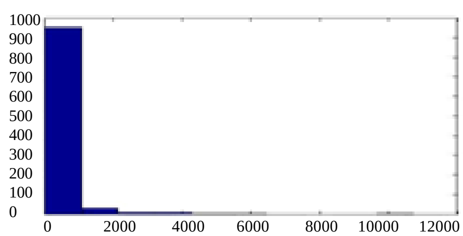


Figure 5. Histogram of the first fragment

The aim of these treatments is to extract the main features to deal with the PCA method.

After quantification, five variables were identified (the variables that define the landing).

- Var 1 : 0 α 200
- Var 2 : 200 α 400

- Var 3 : 400 α 600
- Var 4 : 600 α 800
- Var 5 : 800 α 1000.

- Treatment of all segments: the same procedure is applied to all segments from 1 to 30,504 (the actual size of the recording).
- The implementation of the PCA: the purpose of this step is to reduce the size and draw the relevant information. The input matrix isdefinedin Tables 1 & 2

-0,16	0,34	0,04	0,05	0,66	0,06	-0,06	-0,29
-0,15	0,12	-0,27	-0,14	0,42	0,13	0,25	0,26
-0,19	-0,20	-0,18	0,02	0,01	-0,11	-0,04	-0,24
-0,18	-0,03	-0,05	0,16	-0,12	-0,10	-0,12	0,01
-0,20	-0,02	0,22	-0,07	-0,07	0,02	0,07	0,33
-0,19	-0,03	0,10	0,09	-0,11	0,95	-0,05	-0,04
-0,20	-0,19	0,02	0,13	-0,03	-0,04	0,91	-0,10
-0,22	-0,31	0,16	0,07	0,30	-0,07	-0,14	0,70
-0,17	0,18	0,01	0,17	0,06	-0,07	-0,03	-0,03
-0,18	-0,04	0,00	-0,16	-0,13	-0,02	-0,03	-0,02
-0,14	0,18	-0,36	0,08	-0,14	0,00	-0,02	0,12
-0,19	0,18	0,36	-0,08	-0,12	-0,08	0,01	-0,03
-0,17	0,07	0,00	-0,25	-0,04	-0,02	-0,01	-0,02
-0,19	-0,16	-0,16	0,20	-0,02	-0,03	-0,09	-0,06
-0,17	-0,05	-0,10	-0,47	0,01	0,00	-0,03	-0,06
-0,16	0,34	0,05	0,15	-0,08	-0,07	0,01	0,06

Tableau 1. Répartition des modules sur paliers (1000 jusqu'a 6000)

0,16	0,34	0,05	0,15	-0,08	-0,07	0,01	0,06
-0,17	-0,05	-0,10	-0,47	0,01	0,00	-0,03	-0,06
-0,19	-0,16	-0,16	0,20	-0,02	-0,03	-0,09	-0,06
-0,17	0,07	0,00	-0,25	-0,04	-0,02	-0,01	-0,02
-0,19	0,18	0,36	-0,08	-0,12	-0,08	0,01	-0,03
-0,14	0,18	-0,36	0,08	-0,14	0,00	-0,02	0,12
-0,18	-0,04	0,00	-0,16	-0,13	-0,02	-0,03	-0,02
-0,17	0,18	0,01	0,17	0,06	-0,07	-0,03	-0,03
-0,22	-0,31	0,16	0,07	0,30	-0,07	-0,14	-0,30
-0,20	-0,19	0,02	0,13	-0,03	-0,04	-0,09	-0,10
-0,19	-0,03	0,10	0,09	-0,11	-0,05	-0,05	-0,04
-0,20	-0,02	0,22	-0,07	-0,07	-0,06	-0,03	-0,08
-0,18	-0,03	-0,05	0,16	-0,18	-0,03	-0,05	0,02
-0,19	-0,20	-0,18	0,02	-0,02	-0,01	-0,09	-0,07
-0,15	0,12	-0,27	-0,14	-0,05	0,00	-0,02	0,05

Tableau 2. Répartition des modules sur paliers (6000 jusqu'à 12000)

5. Resultant and discussion

The result of this study represents 12 classes of signals which we call families of relevant information. This was achieved by applying the PCR method. Figure 5 shows a grid of 31 modules treated segments. The legend shows the different levels used modules.

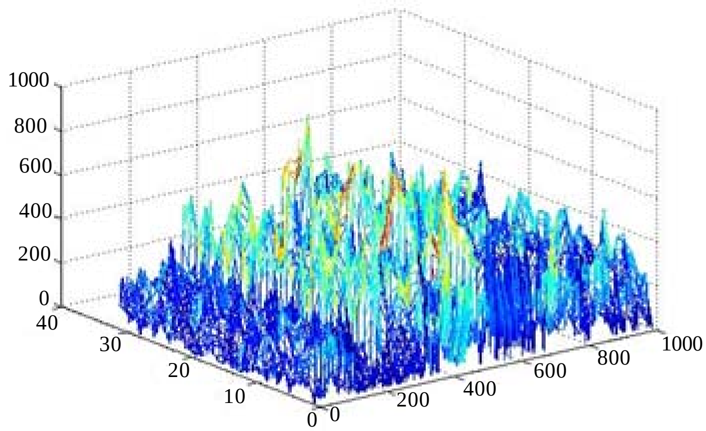


Figure 5. Modules des 31 segments

The resulting segments build the raw data of the CPA. The implementation of the CPA shows that 82% of the segments are on the first axis, 16% is on the second axis where the very good distribution on the two axes. Figure 6 shows the distribution of segments (components) on the new axes.

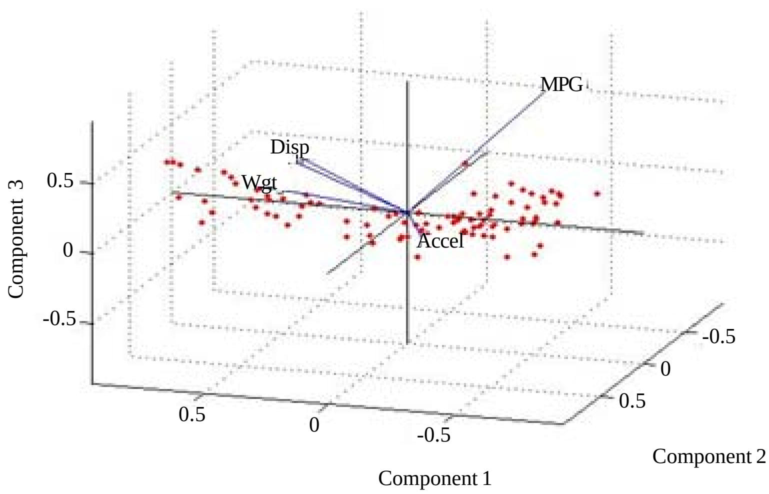


Figure 6. représentation des composantes principales

From this presentation we can say that relevant information is focused to 82% on all segments and corresponding frequencies 600 and 800 hz.

6. Conclusion

The length of EEG recording is often a problem for data analysis. This analysis is based only on relevant information makes this process very difficult either to find the appropriate signal or for analysis. This approach contributes significantly to the analysis of EEG data is reducing the volume of information to a group of families who will be the analysis. The main goal of this study and creates a single space and mostly reduced to analyze its contents. Followed in real time of 23 seconds EEG often poses problems from one side to the complexity of the latter and on the other hand the length of this signal makes the relevant information very difficult to locate. In our case a reduction of signal recording 30,504 in 12 components is a first result in the future to begin the search for the family of signals relevant to the analysis.

7. Acknowledgment

The authors thank all the researchers of the research laboratory LTE and Dr.Bouchetara, a neurologist at CHU- Oran.

References

[1] Serratrice, G., Autret, A. (1996). Neurologie : ISBN 2-7298-9656-2 Ellipses.

[2] Pulvermüller, F. (2005). Brain mechanisms linking language and action. Nature Reviews Neuroscience, 6 (7) 576-82.

[3] Boudet, S., Peyrodie, L., Galois, P., Vasseur, C. (2006). A global approach for automatic artifact removal for standard EEG record, *In*: Proceedings of 28th Annual International Conference of the IEEE-EMBS, p. 5719–5722.

[4] Montaine BERNARD. (2002). Méthodologie d'analyse des synchronisations neuronales dans les signaux EEG à l'aide de graphes d'informations temps-fréquence, École Doctorale : Sciences Pour l'Ingénieur et Aéronautique, France.

[5] Guo, L., Rivero, D., Dorado, J., Munteanu, C.R., Pazos, A. (2011). Automatic feature extraction using genetic programming: An application to epileptic EEG classification, Expert Systems with Applications, doi:0.1016/j.eswa.2011.02.118.