

## A review of the state of the art of feature extraction of electroencephalographic signals

### Revisión del estado del arte de la extracción de características de señales electroencefalográficas

ESQUEDA-ELIZONDO, José Jaime†\*, JIMÉNEZ-BERISTÁIN, Laura, CHÁVEZ-GUZMÁN, Carlos Alberto and GUERRA-FRAUSTRO, Ricardo Jesús Renato

*Universidad Autónoma de Baja California, Faculty of Chemical Sciences and Engineering, Faculty of Engineering and Business*

ID 1<sup>st</sup> Author: *José Jaime, Esqueda-Elizondo* / ORC ID: 0000-0001-8710-8978, CVU CONACYT ID: 90966

ID 1<sup>st</sup> Coauthor: *Laura, Jiménez-Beristáin* / ORC ID: 0000-0002-9362-5450, CVU CONACYT ID: 91007

ID 2<sup>nd</sup> Coauthor: *Carlos Alberto, Chávez-Guzmán* / ORC ID: 0000-0002-2850-3676, CVU CONACYT ID: 94692

ID 3<sup>rd</sup> Coauthor: *Ricardo Jesús Renato, Guerra-Fraustro* / ORC ID: 0000-0001-5491-4105, CVU CONACYT ID: 162677

DOI: 10.35429/EJE.2019.11.6.21.27

Received September 18, 2019; Accepted December 20, 2019

#### Abstract

We present a review of the state of the art of the techniques and algorithms most used in the selection and detection of characteristics of electroencephalographic signals of people when consciously performing activities. These features are numeric parameters that describe the behavior of the signal and are the basis of patterns. In addition, previous experiences in the acquisition of electroencephalographic signals using the Epoc brain-computer interface manufactured by Emotiv are presented. First, some techniques used to eliminate artifacts (disturbances) present in the signal generated by blinking, strong breathing or other movements that contaminate the signal are presented. Later, the algorithms most frequently used in the processing of electroencephalographic signals are shown for the extraction of characteristics that describe the behavior of these patterns and that can be used to detect and recognize patterns in other signals. Finally, we present the lessons that we have acquired as a work team in the recording of electroencephalographic signals in order to be helpful for beginners.

**Feature extraction, Electroencephalographic signals, BCI**

#### Resumen

Se presenta una revisión del estado del arte de las técnicas y los algoritmos más empleados en la selección y detección de características de señales electroencefalográficas de personas al desarrollar actividades de forma consciente. Estas características son parámetros numéricos que describen el comportamiento de la señal y son la base de los patrones. Asimismo, se presentan experiencias adquiridas en la adquisición de señales electroencefalográficas con la interfaz cerebro computadora Epoc del fabricante Emotiv. Se presentan primero técnicas empleadas para eliminar artefactos (perturbaciones) presentes en las señales generadas por parpadeos u otros movimientos que contaminan a la señal. Posteriormente se muestran los algoritmos más usados en el procesamiento de señales electroencefalográficas para la extracción de características que describan el comportamiento de esos patrones y que puedan ser utilizados para detectar y reconocer patrones en otras señales. Finalmente, se presentan las lecciones que como equipo de trabajo hemos adquirido en la toma de registros de señales electroencefalográficas, esperando que esto sea útil a los que inician en el campo.

**Extracción de características, Señales electroencefalográficas, BCI**

**Citation:** ESQUEDA-ELIZONDO, José Jaime, JIMÉNEZ-BERISTÁIN, Laura, CHÁVEZ-GUZMÁN, Carlos Alberto and GUERRA-FRAUSTRO, Ricardo Jesús Renato. A review of the state of the art of feature extraction of electroencephalographic signals. ECORFAN Journal-Ecuador. 2019. 6-11: 21-27

\* Correspondence to Author (email: [jjesqueda@uabc.edu.mx](mailto:jjesqueda@uabc.edu.mx))

† Researcher contributing as first author.

## Introduction

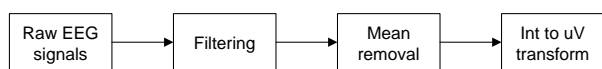
With the appearance of a variety of commercial Brain Computer Interfaces (BCI) systems, there has been a need to carry out research in the EEG pattern detection. These BCI systems can be used in applications for medical, smart environments, detection, rehabilitation, gaming, education, authentication and other areas (Abdulkader, Atia, & Mostafa, 2015). This paper has the intention to help other people that are beginning in EEG signal processing. BCI systems use electroencephalographic (EEG) sensors (electrodes placed on the scalp) to acquire these signals and send them to an external device where it can be transformed into instructions or commands. Generally, a BCI system is divided in the following processes: signal acquisition, pre-processing, feature extraction and classification (Miao, Wang, Zhao, & Liu, 2016). The key factors that define the effectiveness of a BCI system are the feature extraction and the classification stages.

## Signal acquisition

Many BCI systems are commercially available, like those presented in (Ramadan & Vasilakos, 2017). The BCIS are made with one to 256 electrodes and there are wired or wireless. There are open source BCIs like OpenBci or others like Epop Emotiv, that require a license. Also, in (Ramadan & Vasilakos, 2017) some software used with BCIs or for signal analysis are presented. With this software some preliminary signal analysis can be done.

## Pre-processing

The preprocessing stage generally includes eliminating the isoelectric lines of each channel, filtering and converting data from integer to a floating point that represents the real voltage level measured. Using digital signal processing techniques, it is possible to reduce noise, interference and the artifacts present in the signal. This must be done before the feature extraction stage. This procedure is shown in figure 1.



**Figure 1** Signal preprocessing  
Source: Self-Made

## Features

The features are numeric parameters used to describe signals, in this case, the EEG signals. Some common features are: amplitude of the EEG signals, band power, power spectrum density, autoregressive parameters, time-frequency features and features based in inverse models.

## Feature considerations

When a BCI system is designed, there are some points that need to be considered:

- Noise and atypical values: EEG signals have noise and they have atypical values because of the low signal to noise ratio.
- High dimensionality: feature vectors normally have high dimensionality because they come from different channels and from several time segments before they can be concatenated in only one feature vector.
- Time information: the extracted features have time information like the patterns related to some variations in a specific time window of the EEG.
- Non stationarity: EEG signals are non-stationary because they change rapidly with respect to time.
- Small training sets: datasets are relatively small because the training process requires time and data.

## EEG time variations considerations

Most of the brain activity patterns used in a BCI system are related to particular time variations of the signal, possibly in specific frequency bands. For this reason, it is necessary to consider the timeline during the feature extraction stage, (Lotte, Congedo, Lécuyer, Lamarche, & Arnaldi, 2007) propose:

- Concatenation of different time segments features: this consists of feature extractions from different time segments and concatenating them into one single features vector.

- Combination of classificatory algorithms in different time segments: this consists of making the feature extraction and the classification in different time segments and then combining the results of the different classifiers
- Dynamic classification: this consists of extracting features from different time segments in order to construct a temporal sequence of feature vectors.

### Feature extraction

Feature extraction allows us to obtain useful or descriptive information hidden in a signal by decreasing the unnecessary or redundant information. In the feature extraction stage, some algorithms are used in order to obtain the useful information hidden in the signal that describe its behavior. It usually involves a dimensionality reduction or a data compression process that also reduces the amount of data needed to process. The feature extraction evolved from time processing (1960s), frequency processing (1960s-1980s), time-frequency (1980s-2000+) and sparse (2000+-today) (Krishnan & Athavale, 2018). With this classification, the feature extracting techniques can be grouped in:

- Time domain
- Frequency domain
- Joint time – frequency domain
- Signal decomposition and sparse domains

### Time domain feature extraction techniques

This kind of techniques extract characteristic properties or features for a specific time window containing  $N$  samples. As we know, biomedical signals are non-linear and non-stationary, so the time window can be placed anywhere. The underlying patterns or descriptive information could remain the same for specific phenomenon presented in the EEG signals. The basic features that can be extracted are their statistical properties, such as mean, variance, standard deviation, quadratic mean value, etc. There are also some other feature extraction time domain techniques, such as cross correlation, Autoregressive modelling (AR), Linear Predictive Coding (LPC), Cepstrum analysis and Kernel based modelling.

AR modeling and LPC calculate the future values as a function of the current and past values. The Cepstrum analysis is based on the rate of change of different frequency spectrum bands of the analyzed signal. It is similar to the homomorphic filtering, where the signals are transformed by joint additions and multiplication operations. One of the advantages of the AR modelling is that it makes an enhanced data compaction and for spectral peaks detection, and also reduces signal noise and it provides a very good resolution. The order of the AR model cannot be determined a priori.

Cepstrum does not need to calculate the Fourier Transform of the signal, so it is considered a time domain processing. Cepstrum requires that the analyzed signal must be stationary over the time interval. This technique has been used in seismic, voice and geophysical signals.

### Frequency domain feature extracting techniques

These techniques transform the signals from time domain to frequency domain in order to extract low level features of a windowed signal. But sometimes these characteristics may not represent the real stationarity of the signal and only obtain global information that approximately classifies the signals.

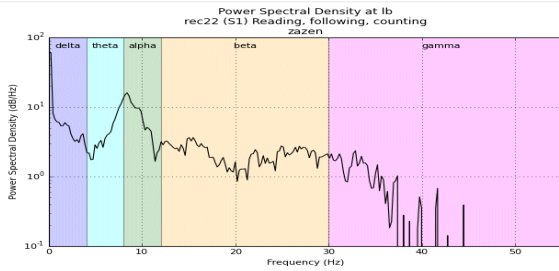
### Power Spectrum Density

The power spectrum density (PSD) function shows how the power of a signal is distributed in the frequency. This can be obtained with the Wiener-Khinchine Theorem or with the Short Time Fourier Transform. The Wiener-Khinchine Theorem express that the PSD and the Correlation are Fourier Transform pairs, as it is shown in

$$S_{xx}(f) = \mathfrak{F}[R_{xx}(\tau)] \quad (1)$$

$$R_{xx}(\tau) = \mathfrak{F}^{-1}[S_{xx}(f)] \quad (2)$$

Where  $S_{xx}(f)$  is the self-spectrum or PSD of the signal  $x(t)$  and  $R_{xx}(\tau)$  is the autocorrelation of the signal  $x(t)$ . Figure 2 shows an example of a PSD of an EEG signals and their bands.



**Figure 2** EEG Power Spectral Density  
 Source: (Trowbridge, 2015)

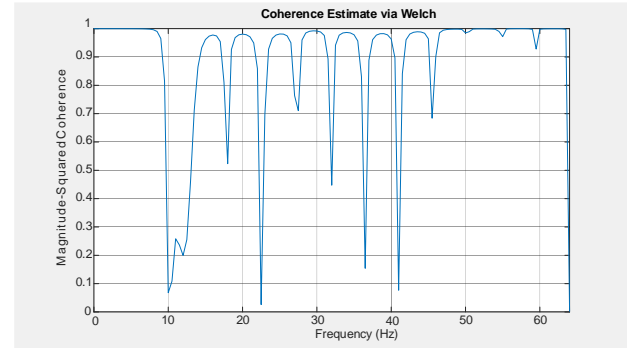
## Coherence

Coherence is a frequency function given in normalized units between 0 and 1, that shows how much one signal  $x(t)$  corresponds in power to another signal  $y(t)$  for each frequency. In other words, coherence is a quadratic correlation coefficient that estimates the amplitude and phase consistence between two signals for every frequency. When the signal  $x(t)$  entirely corresponds to the signal  $y(t)$ , that means that are the same signal. Coherence is given by

$$\Gamma^2(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)}; \quad 0 \leq \Gamma(f) \leq 1. \quad (3)$$

Any pair of signals can be coherent in some frequency bands and have no coherence in another bands. Contrasting with amplitude measurements, coherence is the measurement of the synchronization between two signals principally based in the phase consistence. This represents that if two signals have different phase, like in simple linear circuits, high coherence is presented when the phase difference tends to remain constant.

For each frequency, the coherence measures when the signals are related by a linear time invariant transformation, which means that they have a constant amplitude and phase displacement rate (David et al., 2016; Schwartz, Kessler, Gaughan, & Buckley, 2017) (David et al., 2016) (Esqueda Elizondo José Jaime, Bermúdez Encarnación Enrique Guadalupe, Jiménez Beristáin Laura, Rojo Ramírez Yesenia, Ruiz Morales Angélica, Munguía Carrillo Paul Eriel, 2015).



**Figure 3** Coherence of two one seconds segment of an EEG signal  
 Source: Self-Made

## Discrete Cosine Transform

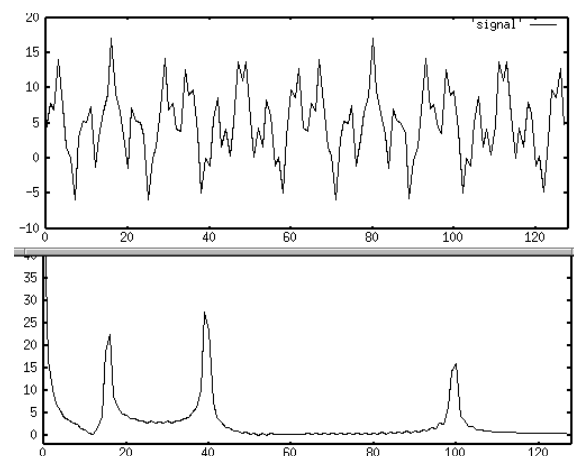
The Discrete Cosine Transform (DCT) is a variant of the Discrete Fourier Transform (DFT), in which the signal is decomposed in a sum of cosine functions (not in sine and cosine functions as in the FDT) and these cosine coefficients are the interested features that describe the signals. The Discrete Cosine Transform is defined as:

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos\left(\frac{(2x+1)u\pi}{2N}\right) \quad (4)$$

where

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{para } u = 0 \\ \sqrt{\frac{2}{N}} & \text{para } u = 1, 2, \dots, N-1 \end{cases} \quad (5)$$

The DCT is usually used to compress information because of its energy compaction property (Krishnan & Athavale, 2018) (Golmohammadi, Torbati, de Diego, Obeid, & Picone, 2017). Figure 4 shows an example of a DCT Power Spectrum of 3 128 samples sinusoidal signals with an offset.



**Figure 4** DCT Power Spectrum of three 128 samples sinusoidal signals with offset  
 Source: (Ruye Wang, 2007)

## Wavelet Transform

The Wavelet Transform is a mathematical tool that can carry out a Time-Frequency analysis of a signal. This is an advantage compared with Discrete Fourier Transform based techniques because with Fourier we can know which frequential components are present in the signal, but we do not know in which moment they appeared. With Wavelet Transform provides that information.

The Wavelet Transform is defined by:

$$W_f(s, \tau) = \int f(t) \psi_{s,\tau}^*(t) dt \quad (6)$$

Where  $\psi_{s,\tau}^*(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right)$  is the mother Wavelet.

The Continuous Wavelet Transform is used in EEG signals to reduce noise in order to obtain a clearer and more precise signal. It can also be used to observe the Time-Frequency relationship of a signal. This Transform shows the moments with higher power in function of both time and frequency, representing them graphically in scalograms. The scalograms are a representation of the Wavelet Transform that in the  $x$  axis shows time and in the  $y$  axis shows the frequency scale and is given by:

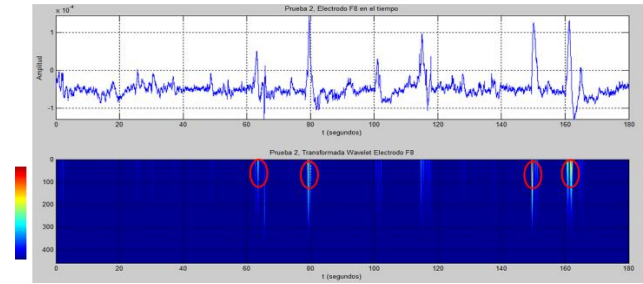
$$\iint |CWT(\tau, \alpha)|^2 \frac{d\tau, d\alpha}{\alpha^2} = E_x \quad (7)$$

Where  $E_x = \int |x(t)|^2 dt$

For a signal  $x(t)$ , if the FDT is given by the relation  $W(t, \omega)$ , so the total energy of the signal will be:

$$\iint W(t, \omega) = \int |x(t)|^2 dt = \int |X(\omega)|^2 d\omega \quad (8)$$

Figure 5 shows both the time domain EEG signal and its Continuous Wavelet Transform, with the colormap, where red color means high power and blue color means low power.



**Figure 5** Time domain EEG signal and its Continuous Wavelet Transform  
*Source: Self-Made*

## Average Instant Frequency

The average instant frequency is given by the first invariant time movement of the FDT trough the frequency axes:

$$E_t \omega = \frac{1}{|x(t)|^2} \int \omega W(t, \omega) d\omega \quad (9)$$

And the group delay through the time axis is:

$$E_{\omega} t = \frac{1}{|X(\omega)|^2} \int t W(t, \omega) dt \quad (10)$$

In a similar way, the instant bandwidth is:

$$\sigma_{\omega} = (\int (\omega - \langle \omega \rangle)^2 |X(\omega)|^2)^{1/2} \quad (11)$$

Where  $\langle \omega \rangle = \int \omega |X(\omega)|^2 d\omega$ .

## Entropy

The entropy is a parameter that measures the randomness of a chaotic system or signal. Due to the complexity and non-linearity of the EEG signals, entropy can be used to analyze them. There are different types of entropy. For example, Shannon or log-energy entropies are used to measure the irregularities of EEG signals. Shannon entropy ( $H_{SE}$ ) measures the average information contained in the probability distribution function (PDF). Log-energy entropy ( $H_{LE}$ ) describe the electrophysiological behavior of different kinds of EEG signals. The Renyi entropy ( $H_{RE}$ ) has been used to derive the continuous family of information measures. The Tsallis Wavelet entropy ( $H_{TE}$ ) can extract some improved features by reducing the negative effects of the wavelet aliasing.

For an N samples dataset, these entropies (Das & Bhuiyan, 2017) are given by:

$$H_{SE} = -\sum_{k=1}^N p_i^2 \times \log_2(p_k^2) \quad (12)$$

$$H_{LE} = \sum_{k=1}^N \log_2(p_k^2) \quad (13)$$

$$H_{RE} = \frac{1}{1-\alpha} \log \sum_{k=1}^N p_k^\alpha \quad (14)$$

$$H_{TE} = \frac{1}{\alpha-1} (1 - \sum_{k=1}^N p_k^\alpha) \quad (15)$$

The Renyi entropy and the Tsallis entropy are of order  $\alpha$ , where  $\alpha > 0$  and  $\alpha \neq 1$ . If  $\alpha = 2$ , the estimation equally emphasizes the sub-Gaussian and the super-Gaussian components and is known as quadratic Renyi entropy and quadratic Tsallis entropy and they are given by (Das & Bhuiyan, 2017):

$$H_{RE} = -\log \sum_{k=1}^N (p_k)^2 \quad (16)$$

$$H_{TE} = 1 - \sum_{k=1}^N p_k^2 \quad (17)$$

## Machine Learning

Machine learning, also known as automatic learning is defined as the use of algorithms for data organization, pattern recognition in order for the computer to learn with these models and solve tasks without pre-programming. Machine learning is the computer's science area that defines Artificial Intelligence. The machine learning algorithms learn from the input data, and with these data, the machines are trained in order to execute different tasks autonomously.

Deep learning is the part of the automatic learning that using high level algorithms, imitates the neural network of the brain. These are complex algorithms made from several layers of neurons fed by a huge volume of data and are capable of image and speech recognition and to execute advanced tasks without human intervention (Krishnan & Athavale, 2018; Teo & Chia, 2018).

## Experiences using Epoc BCI

During this research we worked with the Epoc+ made by the Emotiv Co. One of the lessons we learned with this project is that when we work with kids with Autistic Spectrum Disorder (ASD), first we have to spend some time trying to make the ASD user accepts and tolerates the headset.

We recommend that a visual inspection always be done first in order to eliminate and always record a video of the EEG recording process to have a timeline that can be used to locate movements that can generate disturbances. The problem with the new version of the Epoc+ is that a purchased software license is required in order to have access to the EEG data records.

## Acknowledgments

We want to thank the UABC for funding this project and to the students that participated in this research: Diana Yara Hernández Abarca, Diana Carolina Ramos Solano, Cecilia del Carmen Solano Mendivil, Rosa Itzel Ortíz Quezada, Erik Omar González Ramírez, Sara Jessica Brito Calvo y Valeria López López.

## Conclusions

There are several techniques used to extract features from the EEG signals, depending on the intended use of the EEG signals. The literature reviewed recommends to combine feature extraction techniques and to extract features in different frequency bands, to have a more complete set of signal characteristics.

## References

- Abdulkader, S. N., Atia, A., & Mostafa, M.-S. M. (2015). Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal*, 16(2), 213–230. <https://doi.org/10.1016/J.EIJ.2015.06.002>
- Das, A. B., & Bhuiyan, M. I. H. (2017). Discrimination of focal and non-focal EEG signals using entropy-based features in EEMD and CEEMDAN domains. *Proceedings of 9th International Conference on Electrical and Computer Engineering, ICECE 2016*, 435–438. <https://doi.org/10.1109/ICECE.2016.7853950>
- David, N., Schneider, T. R., Peiker, I., Al-Jawahiri, R., Engel, A. K., & Milne, E. (2016). Variability of cortical oscillation patterns: A possible endophenotype in autism spectrum disorders? *Neuroscience and Biobehavioral Reviews*, 71, 590–600. <https://doi.org/10.1016/j.neubiorev.2016.09.031>

Esqueda Elizondo José Jaime, Bermúdez Encarnación Enrique Guadalupe, Jiménez Beristáin Laura, Rojo Ramírez Yesenia, Ruiz Morales Angélica, Munguía Carrillo Paul Eriel, Z. V. J. M. (2015). Análisis de señales electroencefalográficas en un niño con autismo realizando diferentes actividades de aprendizaje. *Congreso Internacional de Ingeniería Electrónica*, 37, 181–186.

Golmohammadi, M., Torbati, A. H. H. N., de Diego, S. L., Obeid, I., & Picone, J. (2017). *Automatic Analysis of EEGs Using Big Data and Hybrid Deep Learning Architectures*. 13(March), 1–14. <https://doi.org/10.3389/fnhum.2019.00076>

Krishnan, S., & Athavale, Y. (2018). Trends in biomedical signal feature extraction. *Biomedical Signal Processing and Control*, 43, 41–63. <https://doi.org/10.1016/J.BSPC.2018.02.008>

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 4(2), R1–R13. <https://doi.org/10.1088/1741-2560/4/2/R01>

Miao, M., Wang, A., Zhao, C., & Liu, F. (2016). EEG pattern recognition based on dual-tree complex wavelet transform and particle swarm optimization. *Proceedings of the International Conference on Sensing Technology, ICST*, 1–5. <https://doi.org/10.1109/ICSensT.2016.7796311>

Ramadan, R. A., & Vasilakos, A. V. (2017). Brain computer interface: control signals review. *Neurocomputing*, 223(August 2016), 26–44. <https://doi.org/10.1016/j.neucom.2016.10.024>

Ruye Wang. (2007). Fourier Analysis and Image Processing. Retrieved September 30, 2019, from website: [http://fourier.eng.hmc.edu/e101/lectures/Image\\_Processing/node13.html](http://fourier.eng.hmc.edu/e101/lectures/Image_Processing/node13.html)

Schwartz, S., Kessler, R., Gaughan, T., & Buckley, A. W. (2017). Electroencephalogram Coherence Patterns in Autism: An Updated Review. *Pediatric Neurology*, 67, 7–22. <https://doi.org/10.1016/j.pediatrneurol.2016.10.018>

Teo, J., & Chia, J. T. (2018). Deep Neural Classifiers for Eeg-Based Emotion Recognition in Immersive Environments. *2018 International Conference on Smart Computing and Electronic Enterprise, ICSCEE 2018*, 1–6. <https://doi.org/10.1109/ICSCEE.2018.8538382>

Trowbridge, D. (2015). still breathing. Retrieved September 30, 2019, from <https://still-breathing.net/tag/psd/>