

Exploratory Analysis of Motor Imagery local database for BCI systems

D. M. Marín-Castrillón, S. Restrepo-Agudelo, H. J. Areiza-Laverde, A. E. Castro-Ospina
Research Center of the Instituto Tecnológico Metropolitano
Calle 73 No 76A-354, Medellín, Colombia
dianamarin121064@correo.itm.edu.co

L. Duque-Muñoz
Universidad de Antioquia
Calle 70 No. 52 - 21, Medellín, Colombia

Abstract—Brain-computer interfaces (BCIs) are communication systems that translate signals from brain activity to a computer for further processing, aiming to control electronic devices external to the human body and provide communication to people with severe motor disabilities. An usual approach for BCI implementations is to discriminate motor imagery tasks through analyzing time-series, for this aim, in this paper an experimental protocol for brain signal acquisition and a database construction is proposed. Electroencephalography (EEG) is widely used to record motor imagery databases, whereby EMOTIV Epoc+ device was used to acquire brain signals along with BCI2000. For processing EEG signals, Wavelet decomposition, power spectral density (PSD) and empirical mode decomposition (EMD) were performed. Finally, we obtained an maximum accuracy of 91.8% with the implementation of well-known k -nearest neighbors (k -NN) classifier.

I. INTRODUCTION

Brain Computer Interfaces (BCI) are hardware and software communication systems, which allows the interpretation of triggered brain signals during certain tasks, aiming to control electronic devices external to the human body. BCI systems help people with severe motor disabilities to execute movements and improve their communication skills, leading to a better life quality, whereby these systems are used in applications such as Motor Restoration, Environmental Control, Locomotion y Communication [1], [2].

Measuring electrical brain activity is one main component in any BCI system, currently, there are invasive and non-invasive methods for brain signal acquisition. Electroencephalography (EEG), a non-invasive method, is the most widely used, owing to its high temporal resolution and low cost, further it is easy to use because it uses surface electrodes placed on the scalp, according to the International 10-20 System [3].

Negrete et.al [4] described that real movement and motor imagery generate similar patterns in brain activity, specifically in the sensorimotor rhythms μ [8-12Hz] and β [13-30Hz]. The analysis of these signals are performed by two approaches, analysis of evoked potentials generated by visual, auditory or somatosensory stimuli [5], and detection of oscillations in brain signals, i.e. increasing or decreasing of the amplitude in sensorimotor rhythms, i.e. events related with synchronization and desynchronization. Evoked visual potentials have been used in motor imagery dataset performed for some researchers and databases of the Graz University of Technology are commonly used for discriminate imagination motion [6], [7].

Considering variability due to inherent features of the brain according to target population, a database was build implementing a motor imagery paradigm which displayed the necessary instructions to imagine movement of the left and right hand to the subject.

Currently a lot of techniques have been applied for feature extraction in brain signals through of temporal [8], spectral and spatial information [9], [10], using statistical metrics, power spectral density, energy and wavelet transform [11], among others. Therefore, in order to perform an exploratory analysis for feature extraction over proposed local database these techniques were considered. In addition, k -nn classifier was implemented achieved an maximum accuracy of 91.8% on a leave-one-out strategy, giving insight that the constructed dataset could lead to outstanding discrimination results given a deeper analysis, both at feature extraction and classification stages.

II. MATERIALS AND METHODS

A. Hardware and software tools

EMOTIV Epoc+: Neuroheadset EMOTIV Epoc+ is a wireless device for acquisition of Electroencephalography signals. This hardware has 14 electrodes laid according to the international 10-20 system as shown in Figure 1, where red-shaded circles represent the electrodes used to record signals and the green-shaded circles are reference electrodes. Its sampling rate is 128Hz and has a bandwidth of 0.2 to 43Hz. Due to its high temporal resolution, portability, low cost and non-invasive nature EMOTIV Epoc+ can be used to recording brain activity for brain computer interfaces ¹.

BCI2000: is a general-purpose software for BCI research, that has libraries and executable files for free and among its applications can be used for data acquisition and stimulus presentation through the design of paradigms or settings of parameters in the batches that are included in the platform. Therefore, for our purposes Stimulus Presentation batch is used for motor imagery trials. During acquisition, BCI2000 stores data with along with its relevant event triggers and information about used configuration. ²

B. Discrete Wavelet Transform

The discrete wavelet transform (DWT) subdivides a signal into its frequency bands without temporal information loss. This time- frequency decomposition generates two sets of basis functions i.e Wavelets and scaling functions, mathematically DWT can be seen as a set of wavelet transform coefficients, computed as inner product between a wavelet basis and a finite-length sequence of the signal. These allow to analyze the spectrum of the signals as function of time. Additionally, the DWT is also known as a filter bank due to the application of high-pass and low pass filters to the EEG time-series in the decomposition process. DWT can be expressed as:

$$Wf(j, k) = \sum_{n=0}^{N-1} f(n) \cdot \psi_{j,k}^*(n) \quad (1)$$

with $Wf(j, k)$ a DWT coefficient and $f(n)$ is a sequence with length N . Wavelet basis is expressed as:

$$\psi_{j,k}(n) = \frac{1}{\sqrt{s_o^j}} \psi\left(\frac{n - s_o^j \cdot k}{s_o^j}\right) \quad (2)$$

where s_o^j and $s_o^j \cdot k$ refers to the discretized versions of the DWT scale and translation parameters, respectively.

C. Power Spectral Density

The power spectral density (PSD) is the distribution of the energy at a frequency band. The Fourier transform (FT) is used to become the signals from time domain to frequency. The PSD is defined as the squared magnitude of the FT of a given signal [12]. Mathematically, PSD is expressed as follows:

$$P(\omega_i) = \frac{1}{N} |X(\omega_i)|^2 \quad (3)$$

where $X(\omega)$ stands for the FT of signal $x(t)$.

D. Empirical mode decomposition

Empirical mode decomposition is an adaptive data driven approach that allows the representation of a nonlinear and non-stationary signal, which decomposes the signal into a sum of basis components, each one called Intrinsic Mode Function (IMF). During implementation of EMD high frequencies of signals are removed, repeating the process in the new signal also called residue [13], [14], expressed as:

$$x(t) = \sum_{n=1}^N c_n(t) + r_N(t) \quad (4)$$

with N is the total number of IMFs, $c_n(t)$ for $n = 1, \dots, N$ are each IMF and $r_N(t)$ is known as the residual or the last IMF. It is worth to state that lower IMFs contains high frequency components of the signal $x(t)$.

For a function to be considered as an IMF must satisfy two conditions: (i) the number of local extrema and the number of zero crossings must be equal to each other or different by at most one (ii) the mean value of the envelope defined by the local maxima and the local minima should be zero.

EMD algorithm consists on the following steps:

1. Determine all local maxima of the data and join it by cubic spline line as the upper envelope $x_{up}(t)$.
2. Determine all local minima of the data and join it by cubic spline line as the lower envelope $x_{low}(t)$.
3. The mean $m_1 = \frac{x_{up}(t) + x_{low}(t)}{2}$ is computed and subtracted from the original data to obtain the first component:

$$h_1(t) = x(t) - m_1$$

if h_1 meets the two criteria of an IMF it is considered as the first IMF $c_1(t)$.

¹<https://www.emotiv.com/>

²<http://www.schalklab.org/research/bci2000>

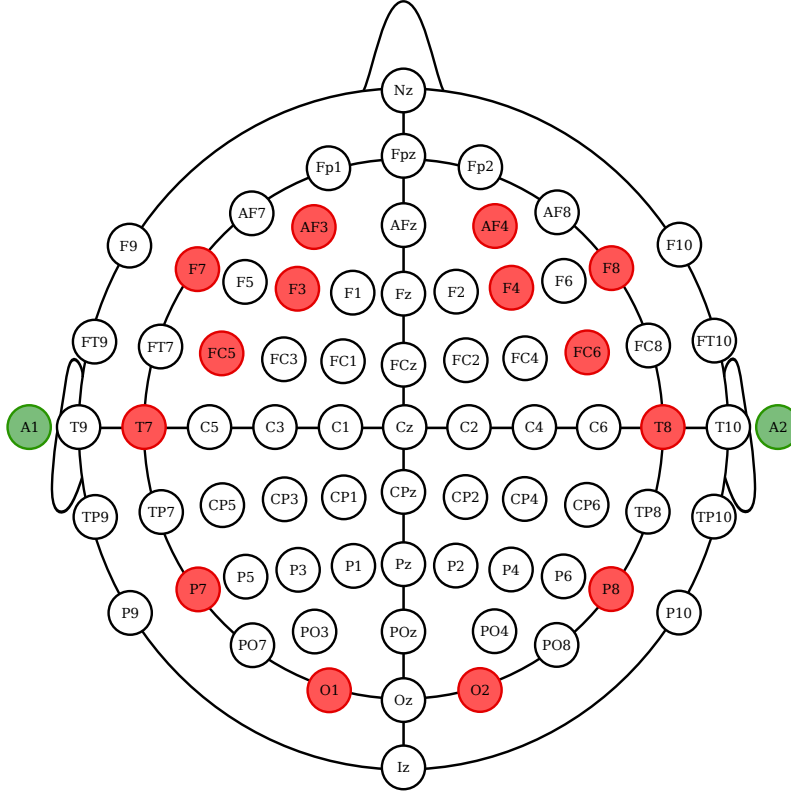


Figure 1. Emotiv Epoc+ electrodes layout

4. If h_1 does not satisfy IMF criteria then it is treated as the data and steps 1 through 3 are repeated until an IMF is found.
5. After an IMF is obtained, it is then subtracted from the original data to get a residue:

$$r_1(t) = x(t) - c_1(t)$$

Such residue is treated as a new signal and the process is repeated to obtain the next IMF ($c_2(t)$) and residue signal $r_2(t)$. Therefore, each residue signal can be expressed as:

$$r_j(t) = r_{j-1}(t) - c_j(t)$$

6. EMD process of decomposition is ended if $r_j(t)$ is a constant or a monotonic function.

E. Linear Discriminant Analysis

Linear discriminant analysis (LDA) searches a linear combination of features or attributes that best separates the classes. It has been also seen as a dimensionality reduction

technique due to the transformation of the p -measurements in a new subspace with lower dimension. In this way, the complexity of the models is reduced and therefore the overfitting is minimized in the classification stage. Finally, LDA assumes that each class data points can be represented by normal density functions and these have the same covariance [15].

F. k -nearest neighbors (k -NN)

This classification method represents all training samples in a p -dimensional space, where p is the number of features that describes each training sample. To classify a new sample into one known class, this method measures the distance between the new sample and all the samples in the training set, the decision is taken considering the most frequent class membership of its k nearest neighbors. The application of feature normalization is necessary to guarantee a better performance of the k -NN method [16].

III. EXPERIMENTAL FRAMEWORK

Experimental work flow is depicted in Figure 2, where the first step is the database construction by a visual paradigm

using Emotiv Epoc+ neuroheadset, then an exploratory analysis is performed over such database by using well-known feature estimation and selection techniques to discriminate among classes with a k -NN classifier.

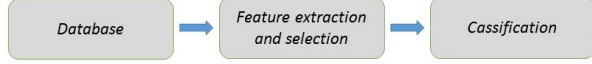


Figure 2. Block diagram of experimental framework

A. Dataset construction

To acquire brain signals with the Emotiv Epoc device, a module of BCI2000 called *Stimulus Presentation* was used for setting parameters. Such module allows the design of the paradigm, number and type of stimuli, duration of stimulus, directory where data are saved, number of channels, type of sequence, among others. At the beginning of the recording, BCI2000 captures data that are transmitted to the computer through of Emotiv Epoc+ USB port.

Fifteen untrained healthy subjects for motor imagery test were chosen, the motor imagery database was obtained with two types of tasks, imagination motion of left and right hand, each participant accomplished a total of 20 trials in a session. By the above, database is presented as a three-dimensional array of $12 \times 256 \times 300$ corresponding to channels, samples and trials respectively.

Test is explained in detail to each person and they are asked to fill informed consent. During test, the subject remains seated in a comfortable chair looking at computer screen, before the device positioning at the subject head, the electrodes must be moistened with saline solution and the quality contact of electrodes is verified with Test Bench application of Emotiv. To begin test, designed parameters are loaded to BCI2000 and information about the participant is entered, when test starts, a prompt with the phrase "waiting to start" is displayed on screen, after a second of recording, 20 arrows are displayed rightward or leftward in random order during two seconds each one, it indicates the direction of imagery movement to subject, between arrows a white background is displayed during two seconds and at the end of the test, a sign with phrase "Time out" is displayed to indicate that the test is over. An example of the whole process is shown in Figure 3.

As an exploratory analysis and according to the state of the art, wavelet decomposition is applied using the mother Wavelet Daubechies 4 [11] with four levels of decomposition and detail coefficients D3 and D4 related with μ and β rhythms were used to extracting energy and statistical features for each electrode, mean, standard deviation and

kurtosis [17].

Other computed features were mean, maximum and minimum of PSD of each signal. Furthermore, each signal was decomposed by means of EMD, then intrinsic mode functions one and two (IMF_1 and IMF_2) were summed to compute the band power and it was used as a feature, as proposed in [18]. Later, z -score normalization was performed and LDA was implemented aiming reduce the feature dimensionality from 180 to 10 dimensions. Finally, k -NN technique was used for class discrimination with Euclidean distance and the leave-one-out validation method.

IV. RESULTS AND DISCUSSION

From the low-dimensional space achieved after LDA processing, class overlapping is low as shown in Figure 4, where can be seen three dimensions in a scatter plot. It is clear how left hand imagery movement, represented as blue dots ($\bullet$), and right hand movement imagery ($\bullet$) are sort of separable.

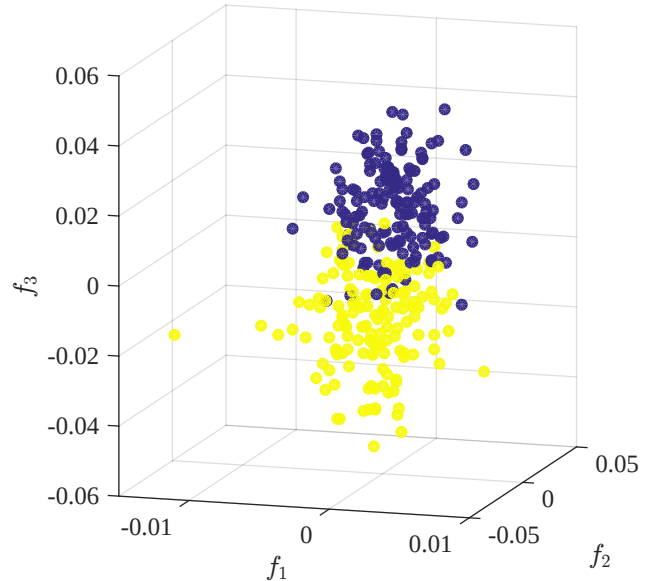


Figure 4. Scatter plot for left and right hand imagery movement

Since this work presents an exploratory analysis, the simple and effective k -NN classifier was used. Also, to demonstrate the database is promising and the results could be improved with different and more elaborated feature estimators and classification techniques. Performing a leave-one-out strategy over constructed dataset and varying k values as odd numbers between 1 and 15, a maximum

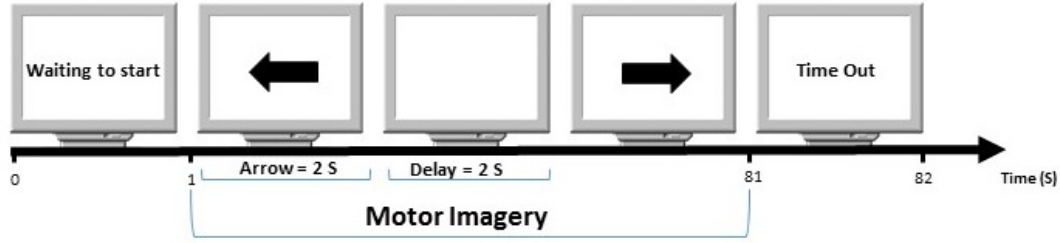


Figure 3. Motor Imagery Paradigm

accuracy 91.8% was achieved for $k = 9$. Besides, as shown in Figure 5 for selected values of k , a minimum of 86% classification accuracy was achieved, which is a suitable result given the low deepness of the proposed analysis.

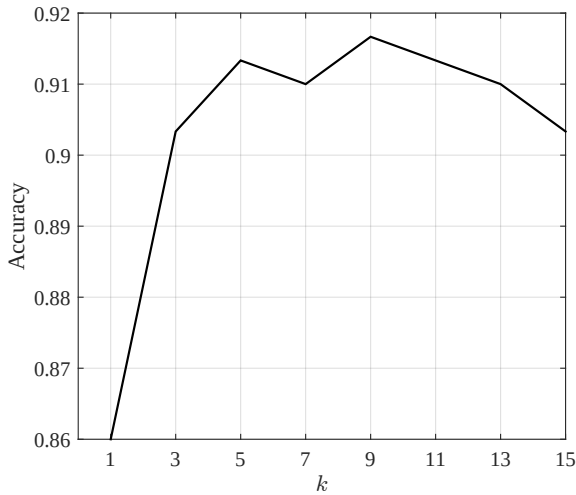


Figure 5. Achieved k -nearest neighbors accuracy results

V. CONCLUSIONS

An experimental protocol for BCI-based brain signal acquisition was proposed, followed by a local database construction for left and right hand movement imagery. Both paradigm design and steps of the test were described. Features based on discrete wavelet transform, empirical mode decomposition and power spectral density were estimated. Then, a linear discriminant analysis was used as dimensionality reduction technique to improve the separation between the classes, leading to high classification accuracy results by means of k -NN.

As future work, more subjects should enlarge dataset as well as the use of feature generation techniques of different

nature or robust classifiers should be used to improve classification results. Also, a channel relevance analysis could give insights of methodology improvement.

ACKNOWLEDGMENT

This work is carried out under grants provided by “Análisis de señales de electroencefalografía orientado a la clasificación de imaginación motora” (Programa Nacional de Jóvenes Investigadores e Innovadores, convocatoria 564 de 2015 – COLCIENCIAS), and Programa de financiación de Becas de Doctorado Nacional convocatoria 647 COLCIENCIAS.

REFERENCES

- [1] L. F. Nicolas-Alonso and J. Gomez-Gil, “Brain computer interfaces, a review,” *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.
- [2] J. E. Muñoz-Cardona, C. D. Muñoz-Cardona, and O. A. Henao-Gallo, “Diseño de una estación de trabajo para personas con discapacidad en miembros superiores usando una interfaz cerebro computador,” *Tecno Lógicas*, pp. 55–66, 2013.
- [3] S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, “Brain computer interfacing: Applications and challenges,” *Egyptian Informatics Journal*, vol. 16, no. 2, pp. 213 – 230, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1110866515000237>
- [4] J. Cantillo-Negrete, J. Gutiérrez-Martínez, T. B. Flores-Rodríguez, R. I. Cariño-Escobar, and D. Elías-Viñas, “Caracterización de la actividad eléctrica cerebral relacionada con la imaginación del movimiento de la mano en sujetos sanos,” *Revista de Investigación Clínica*, vol. 66, no. S1, pp. 111–121, 2014.
- [5] Y. Wang, X. Gao, B. Hong, C. Jia, and S. Gao, “Brain-computer interfaces based on visual evoked potentials,” *IEEE Engineering in medicine and biology magazine*, vol. 27, no. 5, pp. 64–71, 2008.
- [6] L. Vega-Escobar, A. Castro-Ospina, and L. Duque-Muñoz, “Feature extraction schemes for bci systems,” in *Signal Processing, Images and Computer Vision (STSIVA), 2015 20th Symposium on*. IEEE, 2015, pp. 1–6.
- [7] B. Blankertz, K.-R. Muller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. R. Millan, M. Schroder, and N. Birbaumer, “The bci competition iii: Validating alternative approaches to actual bci problems,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, pp. 153–159, 2006.
- [8] M. H. Alomari, A. Samaha, and K. AlKamha, “Automated classification of l/r hand movement eeg signals using advanced feature extraction and machine learning,” *arXiv preprint arXiv:1312.2877*, 2013.

- [9] N. Brodu, F. Lotte, and A. Lécuyer, "Comparative study of band-power extraction techniques for motor imagery classification," in *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*. IEEE, 2011, pp. 1–6.
- [10] H. Bashashati, R. K. Ward, G. E. Birch, and A. Bashashati, "Comparing different classifiers in sensory motor brain computer interfaces," *PloS one*, vol. 10, no. 6, p. e0129435, 2015.
- [11] H. U. Amin, A. S. Malik, R. F. Ahmad, N. Badruddin, N. Kamel, M. Hussain, and W.-T. Chooi, "Feature extraction and classification for eeg signals using wavelet transform and machine learning techniques," *Australasian Physical & Engineering Sciences in Medicine*, vol. 38, no. 1, pp. 139–149, 2015.
- [12] A. Zhang, B. Yang, and L. Huang, "Feature extraction of eeg signals using power spectral entropy," in *2008 International Conference on BioMedical Engineering and Informatics*, vol. 2. IEEE, 2008, pp. 435–439.
- [13] S. Tolwinski, "The hilbert transform and empirical mode decomposition as tools for data analysis," *Tucson: University of Arizona*, 2007.
- [14] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971. The Royal Society, 1998, pp. 903–995.
- [15] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for eeg signal classification," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 11, no. 2, pp. 141–144, 2003.
- [16] A. Yazdani, T. Ebrahimi, and U. Hoffmann, "Classification of eeg signals using dempster shafer theory and a k-nearest neighbor classifier," in *2009 4th International IEEE/EMBS Conference on Neural Engineering*. IEEE, 2009, pp. 327–330.
- [17] J. Hong, X. Qin, J. Bai, P. Zhang, and Y. Cheng, "A combined feature extraction method for left-right hand motor imagery in bci," in *2015 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE, 2015, pp. 2621–2625.
- [18] D. Trad, T. Al-Ani, and M. Jemni, "Motor imagery signal classification for bci system using empirical mode décomposition and bandpower feature extraction," *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, vol. 7, no. 2, pp. 5–16, 2016.