BULETINUL INSTITUTULUI POLITEHNIC DIN IAŞI
Publicat de
Universitatea Tehnică „Gheorghe Asachi” din Iaşi
Tomul LX (LXIV), Fasc. 2, 2014
Secţia
ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ

DETECTION AND CLASSIFICATION OF MU RHYTHM FOR MOTOR MOVEMENT/IMAGERY DATASET

BY

OANA-DIANA EVA[*,1], ROXANA ALDEA[1] and ANCA LAZĂR[2]

[1]"Gheorghe Asachi" Technical University of Iaşi
Faculty of Electronics, Telecommunications and Information Technology
[2]University of Medicine and Pharmacy, Iaşi
Faculty of Biomedical Engineering

Received: May 26, 2014
Accepted for publication: June 30, 2014

Abstract. The brain–computer interfaces (BCI) provide a new way of communication between humans and computers by using brain activity as input signals. The electroencephalography (EEG) is the most used method for brain activity monitoring in BCI systems. The acquired signals from the brain have to be proper interpreted through feature extraction and classification. The classification was performed using three classifiers: linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and Mahalanobis distance (MD) using EEG Motor Movement/Imagery database. The differences between LDA and QDA errors are generally small and depend on trails and pairs of electrodes.

Key words: brain–computer interface (BCI); electroencephalogram (EEG); event related (de)synchronization (ERD/ERS); movement imagery.

1. Introduction

The technology through which human brain communicates with an external device without the involvement of peripheral nerves or muscles is called *Brain–Computer Interface* (BCI). The communication is ensured by a complex system which reveals and identifies a specific brain activity in order to

---

[*]Corresponding author : *e-mail*: oeva@etti.tuiasi.ro

be associated with a set activation, which after that will be sent to a computer for a specific activity. The purpose of a BCI is to allow individuals with severe motor disabilities to have effective control over devices such as computers, speech synthesizers, assistive appliances and neural prosthesis (Bashashati *et al*., 2007).

A BCI system records brain “waves”, extracts key patterns from signals then translates them into commands that are sent to an external device. The brain activity produces a variety of phenomena that can be measured with adequate sensors. Monitoring methods include electroencephalography (EEG), magnetoencephalography (MEG), electrocorticography (ECoC), functional magnetic resonance imaging (fMRI) and positron-emission tomography. Choosing the method based on the EEG signal is most suitable for the following reasons: low cost, which can be used in real-time applications.

The BCI interface is composed of a data acquisition module, a pre-processing block for EEG signals, an extraction block, a block for feature classification and other one for control (Fig. 1). These systems can be assumed as different software modules which are running on a fast PC (Lazăr *et al*., 2009).

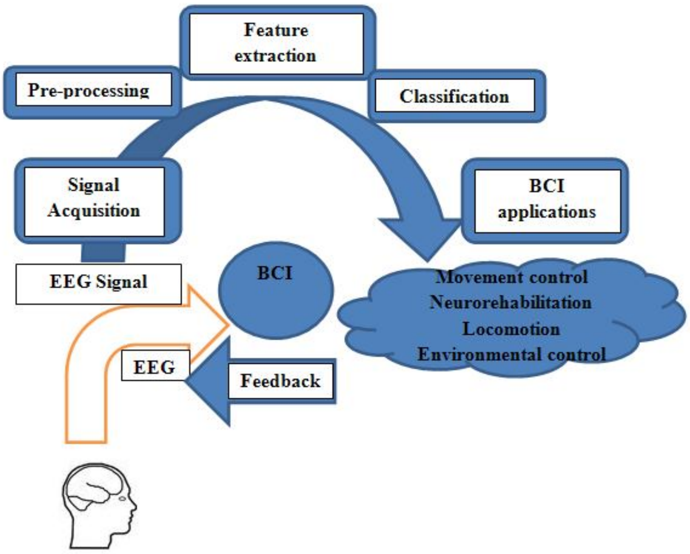

Fig. 1 – Components of a BCI system.

Various neurological phenomena such as: visual evoked potentials (VEP), slow cortical potentials (SCP), steady state visual evoked potentials (SSVEP), P300, modulations in sensorimotor rhythms (Mu, Beta, Gamma) are exploited by the BCI systems (Nicolas-Alonso & Gomez-Gil, 2012).

The EEG signal contains a fairly wide range of frequencies. However, the relevant frequency range from physiological viewpoint lies between 0.1 Hz and 100 Hz. The most important 6 EEG rhythms characterized by their frequency band or their location are: Alpha (8…12 Hz), Beta (13…30 Hz), Theta (4…7.5 Hz), Delta (0.1…3.5 Hz), Gamma (over 30 Hz) and Mu, which is in the same frequency band as Alpha. Mu wave has a spatial distribution essentially confined to the pre-central-post-central region (activity focused over motor cortex) (Kachenoura *et al*., 2007). The rhythms are associated with specific activities: Beta is associated with active thinking and attention, Theta rhythm is associated with emotional process and Mu is affected by movements and movement imagery.

The Mu rhythm was first reported in the 1950s but has not received the kind of attention than other EEG oscillations, most likely because until recently it was thought to occur infrequently and only in a small percentage of the population. New and more sophisticated techniques, such as independent component analysis (ICA), have shown that Mu rhythm are found with scalp EEG in most, if not all, healthy adults (Pineda, 2005).

The oscillations of Mu rhythm are limited to smaller time range between 0.5…2 s and are localized in sensorimotor cortex in the absence of movement (Pfurtscheller & Silva, 1999).

Sensory stimulation, motor behavior and mental imagery can generate modifications in the functional connections of the cortex causing decreasing in amplitude called *event-related desynchronization* (ERD) or increasing in amplitude – *event related synchronization* (ERS) of MU or Beta bands. For example: planning, preparation or imagining left or right hand movement produces an ERD for a short period of time for Mu or Beta in the right and in the left hemisphere in the cortex, respectively.

The motor imagery has become the newest trend in BCI research since the movement imagery appears to recruit neural mechanism in the brain, which are similar or the same to those used to perform actually the same movement (Stavrinou *et al*., 2007). Some BCI systems can describe if the user thinks about moving left or right hand or foot. Others BCI systems rely on a certain specific movement but more abstract. During training sessions, people can develop their own motor imaginary strategy. For example, during the movement of a cursor, people can learn which types of motor imagery movements are suitable for BCI control by moving it up or down. A *prior* training can help the subjects to move the cursor in a 2- or 3-dimensional space.

2. Methods

The EEG signals have random fluctuations and the analysis based on Fourier transform cannot be applied directly. A statistical point of view must be adopted. In the time domain the estimate of the autocorrelation function is used. Its Fourier transform, the power spectral density (PSD), characterizes the signal in frequency domain. PSD is important in the analysis of stationary random processes, quantifying the total distribution of power depending on frequency.

Classical methods for estimating the spectrum power (non-parametric methods) are obtained directly from the signal and make no assumption about how the data are generated. They are based on periodogram introduced by Shuster (1898) and assume obtaining consistent estimates.

The estimate of the spectral density of a signal (periodogram), $\hat{P}_{xx}(f)$, for a sequence, $x(n)$, of length N, is given by (Proakis & Manolakis, 2007):

$$\hat{P}_{xx}(f) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n) e^{-j2\pi fn}\right|^2 = \frac{1}{N}\left|X(f)\right|^2, \tag{1}$$

where $X(f)$ is the discrete Fourier transform for sequence $x(n)$.

Periodogram is an inconsistent estimate of the actual power spectral density. For a finite sequence, the mean value of $\hat{P}_{xx}(f)$ is the Fourier transform of the autocorrelation function of the signal $x(n)$ multiplied with a triangular window (Bartlett) which leads to a distortion of the power spectrum density. Spectrum is influenced by the smoothing and spreading effects embedded in Bartlett window, which makes the estimation of the appropriate spectrums to be limited. In case of the periodogram, dispersion is not cancelled when data with long length are processed (Proakis & Manolakis, 2007).

Others non-parametric methods are obtained through smoothing operations applied directly to periodogram or autocorrelation function. The effect is decreasing of the estimate dispersion with the effect of resolution reduction. In order to minimize periodogram dispersion, the data sequence can be divided into k segments that do not overlap and the periodogram is calculated for each segment. The estimate results as the average of the k segments (averaged periodogram or Bartlett estimate) is

$$\hat{P}_{xxB}(f) = \frac{1}{k}\sum_{k=0}^{k-1} \hat{P}_{xx}^{i}(f), \tag{2}$$

where $\hat{P}_{xx}^{i}(f)$ is the estimate of each segment of length $M = N/k$, given by

$$\hat{P}_{xx}^{i}(f) = \frac{1}{M}\left|X(f)_M^i\right|^2. \tag{3}$$

The averaged periodogram has a smaller dispersion than periodogram.

Welch method is an improved estimator of the power spectral density which is based on periodogram. The method consists of dividing the data into segments that overlap, the calculation of modified periodogram for each segment (because the data sequence is weighted with other window than rectangular one) and then mediating power spectral density estimates. The result is the Welch estimate of the power spectral density; $\hat{P}_{xxW}(f)$ is given by the following relantionship:

$$\hat{P}_{xxW}(f) = \frac{1}{k}\sum_{k=0}^{k-1}\hat{P}^{i}_{xxM}(f), \tag{4}$$

where $\hat{P}^{i}_{xxM}(f)$ is the modified periodogram calculated for each segment,

$$\hat{P}^{i}_{xxM} = \frac{1}{M}\left|\sum_{n=0}^{M-1} w(n)x_i(n)\mathrm{e}^{-\mathrm{j}2\pi fn}\right|^2; \tag{5}$$

$w(n)$ is a distinct to rectangular window. Hamming, Hanning or Blackman windows may be used (Ifeachor & Jevis, 1993).

Averaging the modified periodogram leads to a decrease for estimate dispersion towards estimate dispersion given by the simple periodogram of the entire record. Although the overlap for data segments tends to increase the redundancy this effect is reduced by using a non-rectangular window which reduce the weight given to the samples situated at the end of the segments (samples that overlap). By overlaping the sequences, it results longer segments then in the case of averaged periodogram, which can decrease the dispersion. If the same number of segments is maintained, due to overlapping, longer segments are obtained which leads to resolution increase. In conclusion, a compromise between decreasing dispersion and resolution must be realized. In order to get improved estimates towards periodogram, especially when signal to noise ratio is small, we must find the most suitable parameters for Welch method.

The linear discriminant analysis (LDA, also known as Fisher's LDA), is a classification method originally developed in 1936 by R.A. Fisher. It is simple, mathematically robust and often produces models whose accuracy is as good as more complex methods. LDA is based upon the concept of searching for a linear combination of variables (predictors) that best separates two classes (targets). The aim of LDA is to use hyperplanes to separate the data representing the different classes (Lotte *et al*., 2007).

One way of assessing the effectiveness of the discrimination is to calculate the Mahalanobis distance (MD) between two groups. MD is simple and robust and has shown good performance in BCI research. The MD measures the dissimilarity between feature vectors from different classes and can be used to remove outliners. A distance greater than 3 means that two averages differ by more than 3 standard deviations. It means that the overlap (probability of misclassification) is quite small.

Another standard approach for classification problems is quadratic discriminant analysis (QDA), which models the likelihood of each class as a Gaussian distribution, then uses the posterior distributions to estimate the class for a given test point. The estimates for QDA are similar to those for LDA, except that separate covariance matrices must be estimated for each class (Hastie *et al*., 2009).

3. Dataset and Results

The dataset we used is EEG Motor Movement/ Imagery Dataset (eegmmidb) with 109 subjects and was downloaded from http://www.physionet.org (Mark *et al*., 2014). Subjects had performed different motor/imagery tasks. The EEGs were recorded from 64 electrodes in the extended international 10-20 system (excluding electrodes Nz, F9, F10, FT9, FT10, A1, A2, TP9, TP10, P9, and P10). For testing the method we have taken into consideration the first 30 subjects and only the channels Fc1, Fc2, Fc3, Fc4, C1, C2, C3, C4, Cp1, Cp2, Cp3, Cp4, as shown in Fig. 2.

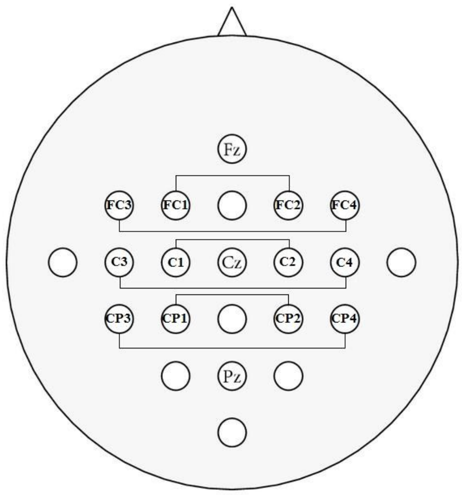

Fig. 2 – Depiction of electrode placement in the standard 10-20 System.

Each subject performed 14 experimental runs: two one-minute baseline runs (one with eyes open, one with eyes closed) and three two-minute runs of each of the four following tasks: open and close left or right fist, imagine opening and closing left or right fist, open and close both fists or both feets, imagine opening and closing both fists or both feets. Only two tasks were considered: movement of the fist and fist movement imagery.

Each signal was sampled at 160 Hz and annotated with the following codes (T0, T1, T2): T0 corresponding to the resting period, T1 corresponding to the onset of motion (real or imagined) of the left fist and T2 corresponds to onset of motion (real or imagined) of the right fist. There are available three data sets for wrist movement (trials 3, 7, 11) and three imaginary data sets (trials 4, 8, 12).

Signals are filtered with a band pass filter 8...12 Hz corresponding to the Mu rhythm range of frequency. We have selected portions from the signals (2 s after stimulus appearance) according to the notation for each mental task (T2, T1), extracting the data for right or left wrist real motion. Similar for the relax period (T0) were extracted sequences of 2 s following right or left wrist movement, while the subject is relaxed. Signals were not filtered by artefacts.

In order to assess the existence of ERD (desynchronization) during real motion performing the power spectral density (PSD) was calculated for all useful channels and for all trails related with right or left fist real motion. The average for these trails was calculated. The MATLAB used function is pwelch with a Hanning window, by specifying the averaging each trial with a factor explicitly (Schmid, 2012). Window length is calculated as the ratio between the length trial and this factor. The function returns frequency components up to 80 Hz (Fs/2), but we have extracted only those components with range of frequency 8...12 Hz. The same procedure was applied for calculating the power spectral density for the relaxation period following right or left wrist opening.

A quantity is calculated, noted as ERD, to assess desynchronization/synchronization which appears on pair of electrodes on the left hemisphere and right hemisphere in right/left wrist real movement,

$$\mathrm{ERD} = \frac{\mathrm{PSD}_{\mathrm{REST}} - \mathrm{PSD}_{\mathrm{MOVEMENT}}}{\mathrm{PSD}_{\mathrm{REST}}}. \tag{6}$$

Feature vector was formed from each pair of electrodes on the left/right hemisphere in the following way: ERD calculated for wrist right movement for the signal recorded from left hemisphere (FC1, FC3, C1, C3, CP1 or CP3), ERD calculated for wrist left movement for the paired electrode from left hemisphere (FC2, FC4, C2, C4, CP2 or CP4), ERD calculated for wrist left movement for the electrode from left hemisphere, ERD calculated for wrist left movement for the electrode from right hemisphere. For the classification step we have used LDA, QDA and MD for all six pairs of electrodes. The steps described above were followed by the trials corresponding to right/left wrist movement imagery. The classification error obtained for the test set was followed for movement/ imagery of movement, pair of electrodes and classifier, as shown in Table 1.

Table 1

Test Error Rate Performance for Classifiers and Pair of Electrodes

Task	Technique	Pair of electrodes					
		FC3-FC4	FC1-FC2	C3-C4	C1-C2	CP3-CP4	CP1-CP2
Movement	LDA	11.98	12.35	11.73	13.89	15.68	13.83
	QDA	9.26	8.89	8.40	10.25	10.31	9.94
	MD	12.28	11.73	12.35	14.26	14.01	14.81
Movement imagery	LDA	14.32	14.57	16.54	15.68	15.93	15.86
	QDA	10.86	11.05	12.78	11.54	11.54	12.84
	MD	15.68	15.31	18.02	14.94	15.86	15.99

The smaller error (8.40) was obtained on electrodes C3/C4 for movement with QDA classifier and for movement imagery the smaller error (10.86) was attained on pair FC3/FC4 for the same classifier. The higher error on movement task was achieved on CP1/CP2 for MD classifier and for movement imagery task on pair C3/C4 also for MD classifier.

The overall errors obtained after applying the quadratic classifier are better than those observed for linear classifier and MD, but the differences between LDA and QDA errors are generally small.

Table 2
Errors for Subject 5 for all Tasks, Trials, Classifiers and Pair of Electrodes

Task	Trial	Classifier	Pair of electrodes					
			FC3-FC4	FC1-FC2	C3-C4	C1-C2	CP3-CP4	CP1-CP2
Movement	3	LDA	0.00	33.33	0.00	33.33	0.00	33.33
		QDA	0.00	33.33	0.00	33.33	0.00	33.33
		MD	0.00	27.78	16.67	33.33	27.78	33.33
	7	LDA	5.56	0.00	0.00	0.00	0.00	0.00
		QDA	0.00	0.00	0.00	0.00	0.00	0.00
		MD	0,00	0.00	0.00	5.56	0.00	0.00
	11	LDA	5.56	11.11	38.89	27.78	27.78	55.56
		QDA	11.11	11.11	38.89	22.22	22.22	50.00
		MD	33.33	11.11	38.89	27.78	27.78	55.56
Movement imagery	4	LDA	0.00	50.00	0.00	61.11	0.00	0.00
		QDA	0.00	50.00	0.00	55.56	0.00	0.00
		MD	0.00	61.11	0.00	50.00	0.00	5.56
	8	LDA	0.00	5.56	0.00	5.56	0.00	0.00
		QDA	0.00	5.56	0.00	0.00	0.00	0.00
		MD	0.00	5.56	0.00	5.56	0.00	0.00
	12	LDA	0.00	0.00	0.00	0.00	0.00	0.00
		QDA	0.00	0.00	0.00	0.00	0.00	0.00
		MD	0.00	0.00	0.00	0.00	0.00	0.00

On subject 21 we have obtained high errors (> 50%) for all pairs of electrodes for movement imagery opposed to movement task errors was less than 1% on Fc3/FC4, FC1/FC2, C3/C4 for all classifiers. On pair CP3/CP4 we have obtained errors smaller than 1% for all movement imagery trials and for all classifiers on subject 3. For movement task, for the same subject on all pairs of electrodes errors were higher. The exception is trial 7. On pairs C1/C2, CP3/CP4, CP1/CP2 the errors are small.

Another subject with acceptable errors (< 1%) for movement and for movement imagery was 27. Depending on classifier, the errors were in the range 33%...39% for trial 7.

15 subjects performed well on movement task, 10 subjects had satisfactory results for movement imagery and at 5 subjects are obtained good errors for movement and also for imagined task. An example is depicted in Table 2.

For real motion the errors were in range between 0…55, and for movement imagery between 0…61. The higher errors obtained in trial 4 can be explained of an imperfect contact of some of the electrodes on the scalp because on the other trials the subject performed well the demanded task. Another possible explanation could be that the subject did not develop the Mu rhythm.

In other works (Sleight *et. al.*, 2009) the classification result vary significantly from one patient to another and some inaccuracies in the Physionet data were found.

4. Conclusions

Using EEG Motor Movement/Imagery databaset, power spectral density and Welch method we have assess the existence of ERD. For the classification we have used LDA, QDA and MD for all six pairs of electrodes. The results show reasonable classification accuracies, which are consistent across most subjects, for movement and movement imagery tasks, all trials and all pairs of electrodes. Classification is slightly more successful for real movement then for imagined movements.

It is important to note that results differ significantly across patients. Movement and imagery movement are challenging and a mental effort is required. Maybe some subjects could not execute the task requested, especially for imagined movement because they do not have imaginative skills or their level of concentration was below average.

The differences are generally small because both LDA and QDA perform well on an amazingly large and diverse set of classification tasks.

Future work involves testing all subjects from the database and applying statistical tests.

REFERENCES

Bashashati A., Fatourechi M., Ward R.K., Birch G.E., *A Survey of Signal Processing Algorithms in Brain–Computer Interfaces Based on Electrical Brain Signals.* J. Neur. Engng., **4**, *2*, R32–57 (2007).

Hastie T., Tibshirani R., Jerome F., *The Elements of Statistical Learning – Data Mining, Inference, and Prediction.* Standford, California, 2009.

Ifeachor E., Jevis B., *Digital Signal Processing-A Practical Approach.* Addison-Wesley Publ. Ltd., Workingham, England, 1993.

Kachenoura A., Albera L., Senhadji L., Comon P., *ICA: A Potential Tool for BCI Systems.* IEEE Signal Proc. Mag., **25**, *1*, 57-68 (2008).

Lazăr A. M. *et al.*, *Interfața Creier-Calculator Paradigme Posibile.* Edit. CE, Iaşi, 2009, 20-23.

Lotte F. *et al.*, *A Review of Classification Algorithms for EEG-Based Brain–Computer Interfaces.* J. of Neur. Engng., **4**, *2*, R1-R13 (2007).

Mark R. G. *et al.*, *PhysioNet: Physiologic Signals, Time Series and Related Open Source Software for Basic*. Clinical, and Applied Citation Accessed Citable Link Detailed Terms PhysioNet: Physiol. Sign. , Time Series a. Related Open Source Software for Basic, Clin., 2014.

Naeem M. *et al.*, *Seperability of Four-Class Motor Imagery Data Using Independent Components Analysis.* J. Neur. Engng., **3**, *3*, 208-16 (2006).

Nicolas-Alonso L. F., Gomez-Gil J., *Brain Computer Interfaces, a Review.* Sensors (Basel)., **12,** *2*, 1211-1279 (2012).

Pfurtscheller G., Silva L.F.H., *Event-Related EEG/MEG Synchronization and Desynchronization: Basic Principles*. Clin. Neurophysiol., **110**, *11*, 1842-1857 (1999).

Proakis J., Manolakis D., *Digital Signal Processing – Principles, Algorithms and Applications.* Fourth Ed., 2007.

Pineda J., *The Functional Significance of Mu Rhythms: Translating 'Seeing' and 'Hearing' into 'Doing*. Brain Res. Rev., **50**, *1*, 57-68 (2007).

Schmid H., *How to Use the FFT and Matlab's pwelch Function for Signal and Noise Simulations and Measurements.* Inst. of Microelectron., August, 2012.

Sleight J, Pillai P., Mohan S., *Classification of Executed and Imagined Motor Movement EEG Signals*. Computer Sci. a. Engng., Univ. of Michigan, Ann Arbor, December, 2009, 1-10.

Stavrinou M. *et al.*, *Evaluation of Cortical Connectivity during Real and Imagined Rhythmic Finger Tapping*. Brain Topography, **19**, *3*, 137-145 (2007).

DETECȚIA ŞI CLASIFICAREA RITMULUI MU FOLOSIND BAZA DE DATE „EEG MOTOR MOVEMENT/IMAGERY”

(Rezumat)

Interfața creier–calculator (Brain–Computer Interface – BCI) oferă un nou mod de comunicare între oameni şi computere, folosind ca semnal de intrare activitatea cerebrală. Electroencefalografia (EEG) este metoda cea mai utilizată pentru monitorizarea activității creierului folosită de sistemele BCI. Semnalele achiziționate trebuie să fie interpretate corect prin extragerea trăsăturilor şi clasificare. Clasificarea a fost realizată prin 3 clasificatori: discriminant liniar (LDA), discriminant pătratic (QDA) şi distanța Mahalanobis (MD).