POLITECNICO DI TORINO

Master Degree in Telecommunication Engineering

Master Degree Thesis

Lateralized Readiness Potentials

Candidate:

Remo Arena(s219565)

Supervisors:

Gabriella Olmo

Roberto Garello

Vito De Feo

Academic year 2017/2018

Contents

Abstract

The goal of this thesis is to investigate and improve the computation of the *Lateralized Readiness Potential* by using tools like Machine Learning and Dimensionality Reduction side by side with the traditional tools usually employed in neurophysiology, such as filtering and trials averaging.
Within this work *Lateralized Readiness Potentials* (*LRP*) are investigated with the intention of understanding whether, in a near future, they could be used as a diagnostic tool in motor and consciousness disorders.
A typical scenario in which could be employed is the diagnosis of a a Locked-In Syndrome.
Locked-in syndrome (*LIS*) is a condition in which a patient has a complete paralysis of all voluntary muscles except for vertical eye movements and blinking. Locked-in syndrome may be confused with a loss of consciousness in patients, thus misleading to a diagnosis of Vegetative State and it may even resemble death.
Being the LRP, or more in general RP, associated to the intentionality of movement, it could unveil if there is still consciousness in what could be misunderstood as a muscle spasm.
During this work of thesis the use of several methods and techniques were investigated. All these methods will be shown in the following, with particular emphasis on the ones employed in the software developed and on the reason why they were preferred to the others.

Chapter 1

Introduction

1.1 Event-related potentials

Event-related potentials (ERPs) are small changes in the scalp-recorded electroencephalogram time-locked to the onset of an event such as a sensory stimulus or a motor act.
International Encyclopedia of the Social & Behavioral Sciences, 2001

The ERPs are electrical potentials generated by the brain as response to a specific event.
This event can be a stimulus presentation followed by sensory-related operations (such as estimation of color, shape, or category of the visual stimulus), a cognitive control operations (such as selection of appropriate response or suppression of prepared action), or an affective operations (such as associated with positive or negative emotions) or even memory-related operations (such as recalling an item or remembering a new item)[1].
The event can also be a motor response, and this is exactly the case of our study.

ERPs can be reliably measured using electroencephalography (EEG), which unfortunately reflects thousands of simultaneously ongoing brain processes. This means that the brain response to a single stimulus or event of interest is not usually visible in the EEG recording of a single trial.
In order to observe brain response to a stimulus, the experimenter must conduct several trials and average the results together, causing random brain activity to be averaged out and the relevant waveform to remain.Thisis the most common way to compute ERP (which is also known as averaged ERP or aERP)[2].

Therefore, by using event-related potentials, the neural correlates of cognitive processes is investigated with a non invasive procedure that has a high temporal resolution.
It is precisely the high temporal resolution the reason why have imaging techniques (e.g. fMRI) not made ERPs (or EEG in general) obsolete.

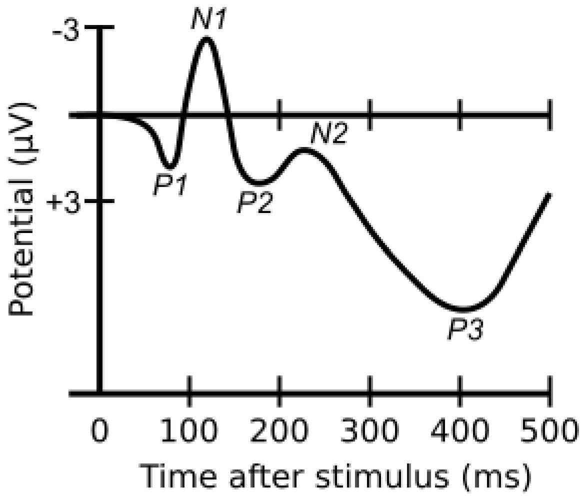


Figure 1.1: A waveform showing several ERP components, including the N100 (labeled N1) and P300 (labeled P3). [2].

Among all the event-related potentials that have been studied, this thesis will address a particular ERP known as *Bereitschaftspotential* or *Readiness Potential*, and more precisely its late component, the *Lateralized Readiness Potential*.

1.1.1 Bereitschaftspotential

The *Bereitschaftspotential* or *Readiness Potential*, also known as pre-motor potential is an event-related potential measured over the motor cortex and over the supplementary motor area of the brain preceding the occurrence of voluntary muscle movement and reflecting the motor planning of volitional movement.

Readiness potential was first recorded in 1964 by *Hans Helmut Kornhuber* and *Luder Deecke* and reported in many of their publications within a study of voluntary movement.

Since in the 1960's computer software did not exist to perform on-line back averaging, Kornhuber and Deecke (1964) recorded electroencephalogram and electromyogram (EMG) simultaneously, while the subjects were performing the same movements at a self-paced rate, and then they stored all the data on magnetic tape. Thus, by playing the tape backward, they performed an on-line averaging of the EEG segment that precedes the onset of the EMG.
In this way, Kornhuber and Deecke detected two components, one preceding the movement onset and one immediate following it: the *Bereitschaftspotential* (BP) or *Readiness Potential* (RP), and *Reafferente Potential* (which won't be treated in this work of thesis).
In further investigation, Kornhuber and Deecke, were able to separate two more signal components preceding the movement onset: the *Pre-Motion Positivity* (PMP) and *Motor Potential* (MP). All these potentials are collectively known as *Movement-related cortical potentials* (*MRCP*), in other words, potentials that occur in close temporal relation with movement or movement related activity (such as motor imaginery, or motor preparation).
According to Kornhuber and Deecke report[6] the BP is a slow cortical negativity that begins about $1,5$ s (average 800 ms) prior to voluntary finger movement and it is bilateral even with unilateral movements.
Furthermore in the last 150 ms before the movement onset, other two potentials with different topography and polarity, PMP and MP, superimposed, to the BP.
The pre-motion positivity PMP is also bilateral and widespread in the parietal and precentral leads of both side and in the midline with a maximum at the anterior parietal region. It occurs approximately $80 - 90$ ms prior to the

movement onset.

The Motor Potential (MP), instead, is the only unilateral potential that precedes unilateral voluntary movement. Its localization is in fact limited to the hand area of the motor cortex contralateral to the moving finger. It occurs approximately $50 - 60$ ms prior the movement onset.

Kornhuber and Deecke hypothesized that PMP might reflect cortical activity related to the initiation of movement, while the MP reflects the motor cortical activity immediately preceding the movement

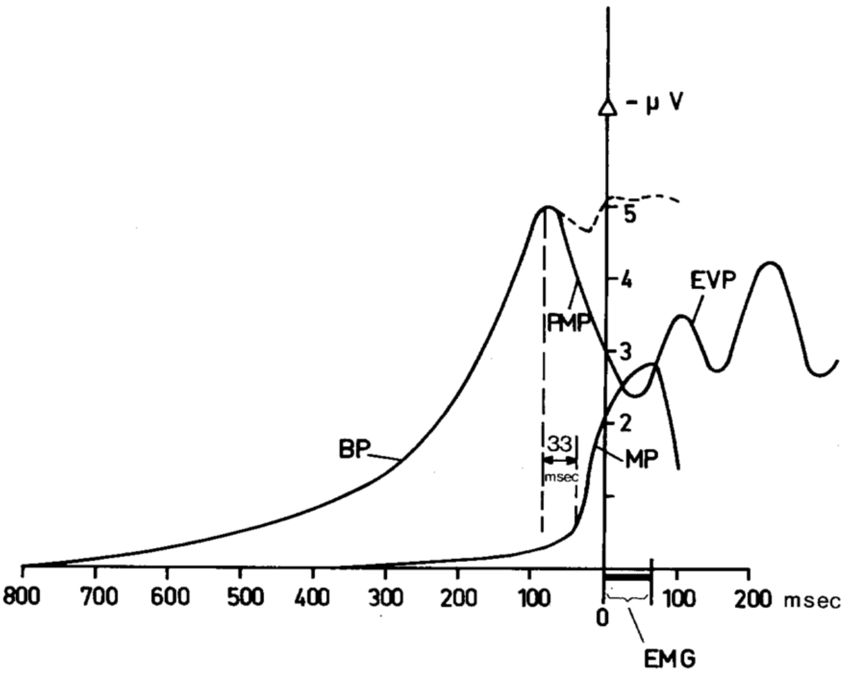


Figure 1.2: Typical potential course for ipsilateral precentral, mid- and lateral parietal leads showing Bereitschaftspotential (BP) and pre-motion positivity (PMP) prior to movement onset in the EMG and proprioceptive evoked potentials (EVP) after movement onset. Stippled, one of the possible variations of the contralateral precentral potential course showing superposition of PMP and an additional negativity immediately prior to movement onset (motor potential, MP). (L.Deecke et al. (1976)) [6].

Resuming, the potential starts with a slow negative deviation (upward, BP) which reverses to positivity (downward, PMP) about $90-80$ ms before EMG onset.
At the contralateral precentral electrode there is additional negativity (MP) around $60-50ms$ before the EMG onset when in the other leads the negative deflection either remains constant or diminishes. Notice that it is a common convention to plot ERP waveforms (or more in general EEG recordings) with negative voltages upward and positive voltages downward.

Bereitschaftspotential has a precise somatotopy, it is located over the parietal and precentral areas of both hemispheres and the midline. Frontally, it is usually positive, or absent but rarely negative; while it is bilaterally symmetrical in the parietal lead. It has been found out that the initial part of precentral BP is bilaterally symmetric but, after 400 ms prior to movement onset, it is typically slightly lateralized. This suggests that the contralateral motor cortex typically generates slightly more negativity respect the ipsilateral one. The lateralization of precentral BP becomes statistically significant around 150 ms prior to movement onset.
The amplitude of BP is directly related to BP onset time, in fact, the earlier BP begins, the larger it becomes and viceversa. Moreover, the amplitude and the time of BP can be influenced by several factors such as level of intention, preparatory state, learning and skill acquisition, force exerted, speed of movement and complexity of movement making the Bereitschaftspotential very fickle.
Shibasaki and Hallet in [7] cite the first part of BP as "early BP" , the second

part "late BP", and just BP to consider both early and late BP (Figure 1.3

The asymmetric distribution of the late BP related to unilateral hand movement was investigated, by Michael G. H. Coles [8], as the *Lateralized Readiness Potential* (LRP).
G.Coles derived the LRP by performing the subtraction between the potential recorded at C3 and at C4, for both the left-hand movement and the right-hand movement separately.

1.1.2 The Lateralized Readiness Potential (LRP)

As already mentioned above the Readiness Potential starts as bilateral over both hemispheres and become to lateralize before the movement onset, with higher amplitude over the contralateral hemisphere with respect to the movement. This lateralization becomes even more relevant for recording sites over the motor cortex.
The LRP is computed on the basis of ERPs recorded before and during the execution of a response over the left and right motor cortices. However, the exact positions of the recordings sites can vary slightly between experimental studies. Often it is chosen recording site pairs are C3' and C4' that are located 1 cm anterior of the C3 and C4 sites specified by the IS 10-20 system.

The most common method for deriving the LRP is the double subtraction method, and it is illustrated in Figure 1.4. The double subtraction method was introduced by De Jong et al., (1988)[9]. In literature other slightly different methods for deriving the LRP have also been investigated.

For instance, M.Coles (1989) [8] described an alternative way of comput-

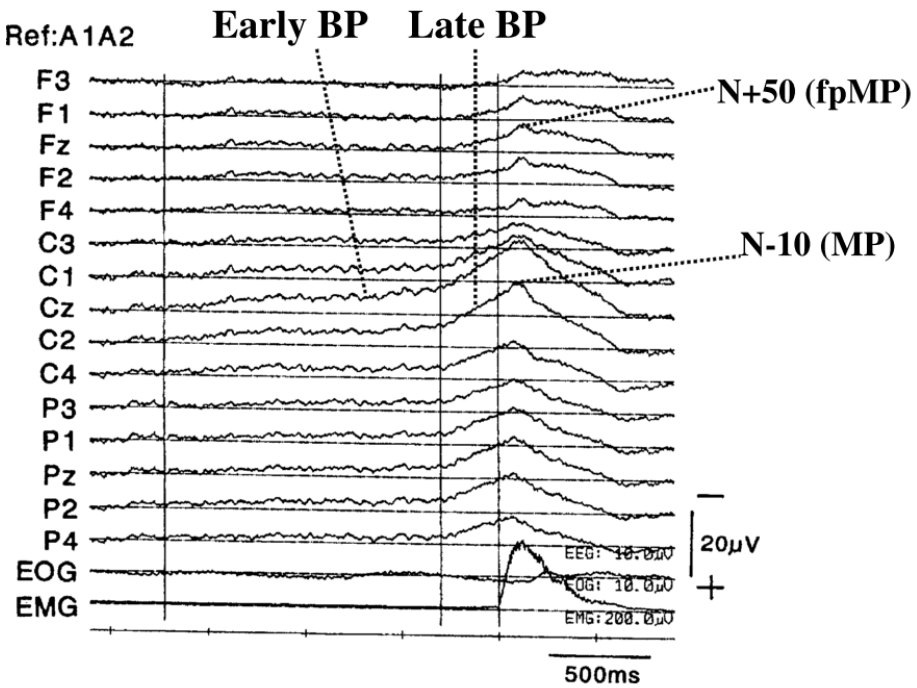


Figure 1.3: Waveforms and terminology of *movement-related cortical potentials (MRCPs)* from a single normal subject. Self-initiated left wrist extension. Average of 98 trials. Reference (Ref): linked ear electrodes ($A1$,$A2$). Early pre-movement negativity *(early BP)* starts 1.7 s before the onset of the averaged, rectified EMG of the left wrist extensor muscle, and is maximal at the midline central electrode (Cz) and widely and symmetrically distributed on both hemispheres. Later negative slope (*late BP*) starts 300ms before the EMG onset and is much larger over the right central region (contralateral to the movement). A negative peak localized at the contralateral central area ($C2$) is $N-10$ or MP. Another negative peak occurring shortly after $N-10$ is localized over the midline frontal region and corresponds to $N+50$ or the *frontal peak of motor potential (fpMP)*. Figure and caption adopted from [7].

ing LRP waveforms the averaging method.

1.1.3 The contingent negative variation

Contingent negative variation (CNV) is a low negative potential that develops in the interval between a "‘Warning"’ and a "‘Go"’ stimulus and shows anticipation for a forthcoming signal and preparation for execution of a response. In other words, CNV reflects preparation for signaled movements and is an index for expectation.
The earlier segment of the CNV has maximum amplitude over the frontal cortex and is generated in response to a "‘Warning"’ cue. The later or terminal CNV(tCNV) begins around 1.5 s before the "’Go"’ cue, it reflects preparation for motor response and has maximum amplitude over the motor cortex (M1).
CNV is a movement-preceding negativity (MPN), just as the Readiness Potential (RP). The RP reflects processes involved in the preparation of voluntary movements, and the CNV reflects processes involved in the preparation of signaled movements. In other words the RP and the CNV are both reflections of anticipatory behavior, at least as far as the motor system is involved. Deecke and Kornhuber pointed to the following differences between RP and CNV:

1. CNV is larger over the frontal areas while the BP over the parietal areas.

2. CNV is symmetrical while BP is lateralized over the precentral electrode.

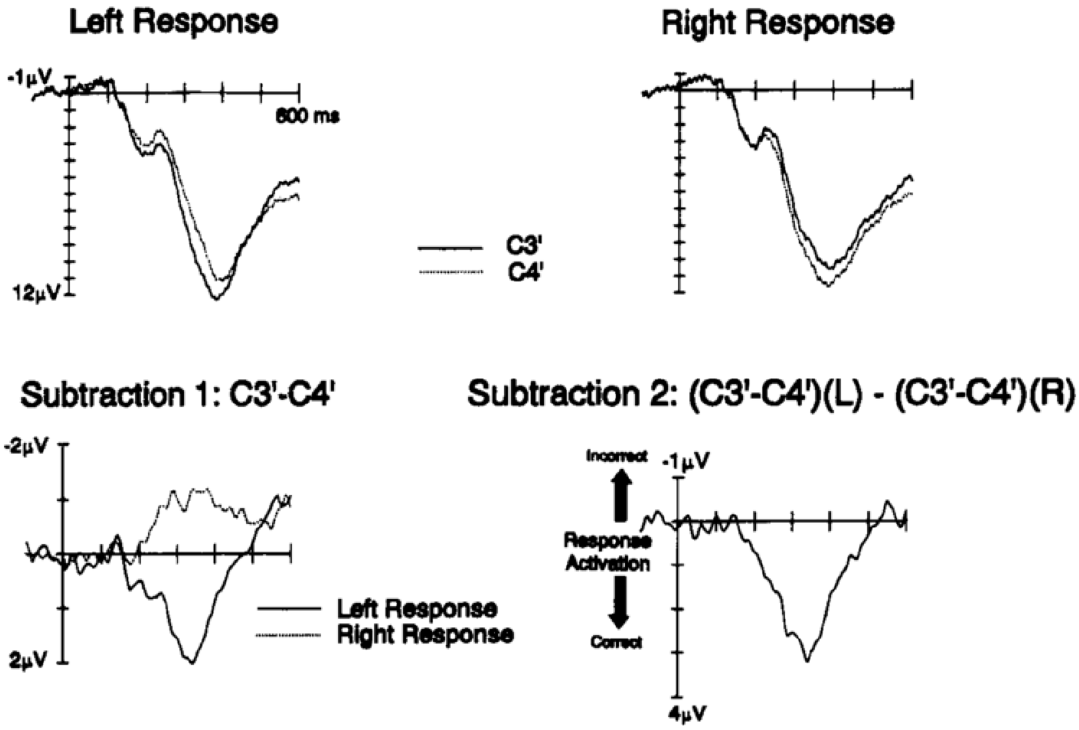


Figure 1.4: Computation of the Lateralized Readiness Potential(LRP) with the double subtraction method on the basis of event-related brain potential(ERP) waveforms elicited atelectrodes $C3'$ (left hemisphere) and $C4'$ (right hemisphere). Top: grand-averaged ERP waveforms elicited at $C3'$ (solid lines) and $C4'$ (dashed lines) in response to stimuli requiring a left-hand response (left side) and to stimuli requiring a right-hand response (right side). Bottom left: difference waveforms resulting from subtracting the ERPs obtained at $C4$' from the ERPs obtained at $C3'$ separately for left-hand responses (solid line) and right-hand responses (dashed line). Bottom right panel: LRP waveform resulting from subtracting the $C3' - C4'$ difference waveform for right-hand responses from the $C3' - C4'$ difference waveform for left-band responses.A downward-going (positive) deflection indicates an activation of the correct response; an upward-going (negative) deflection indicates an activation of the incorrect response. Figures and caption adopted from **M.Eimer, 1998, pag. 148** [10].

3. RP is smaller and increases gradually while the CNV increases more suddenly.
4. Speed instructions enhance the amplitude of the CNV. Comparing CNV recordings prior to fast and slow responses from the same series of trials, the largest amplitudes are found prior to fast responses while for the BP is exactly the opposite.
5. If muscular effort is required for a response to the RS, amplitudes of the CNV late wave increase, compared to a condition in which this effort is not needed. Kutas and Donchin (1977) have described a similar result for the RP.
6. The RP is smaller than the CNV.

The reason why CNV is treated in this thesis, despite the fact that this study is not directly concerned about this ERP, is that both in previous research and in this study it has been often questioned the the hypothetical interaction or correlation between Readiness Potential and Contingent Negative Variation.

1.2 Outline of electroencephalography

Electroencephalography (EEG) is an electrophysiological monitoring method to record electrical activity of the brain. It is typically noninvasive, with the electrodes placed along the scalp, although invasive electrodes are sometimes used such as in electrocorticography.
EEG measures voltage fluctuations resulting from ionic current within the

neurons of the brain[3]. In the Figure below, the block diagram for an EEG Acquisition system is shown:

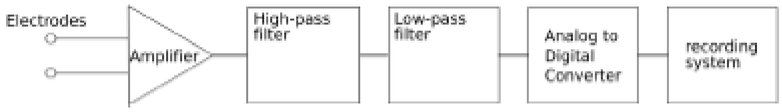


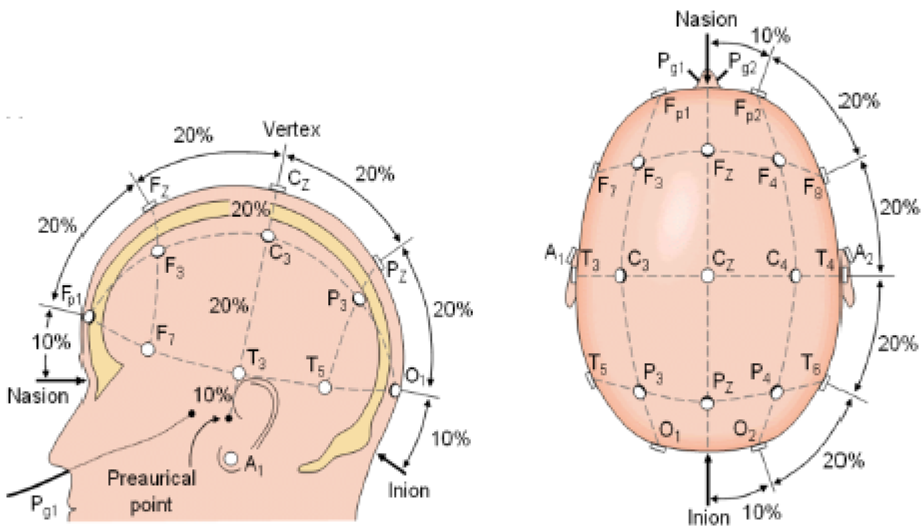
Figure 1.5: EEG acquisition chain block diagram. Figure and caption adopted from K.Blinowska et al., 2012, pag. 107. [4].

The first elements in the diagrams are the electrodes (placed on patients'scalp) and the differential amplifier. A key figure for the quality of the measurement is the ratio between the electrodes impedance and the amplifier input impedance. In order to attain a good quality of the EEG acquisition, the amplifiers must have a very high input impedance, up to the order of $10^{12}\Omega$, while the resistance of the electrodes must be kept under $5k\Omega$.

A High-pass filter is employed after the amplifier in order to eliminate the the baseline (direct component) and the low frequency artifacts. Then, before the EEG signal is digitally converted, a low-pass anti-aliasing filter is employed.

The sampling frequency mights range from 100 Hz for spontaneous EEG and several hundred Hz for ERP, up to several kHz for recording intracranial activity. [4]

It is crucial to know the space location of electrodes exactly, in order to allow a right interpretation of a single recording and the comparison of results obtained between different subjects. The traditional *IS 10-20* electrode system makes use of 19 EEG electrodes placed over specific anatomic landmarks, in a way that 10-20 of the distance between them is the electrode spatial interval. Two more electrodes are placed on earlobes (A1 and A2).



(a) side view

(b) top view

Figure 1.6: 10-20 system view.

Being the EEG is a measure of potential difference, in the referential (or unipolar) setup it is measured relative to the same electrode for all derivations. There is no universal consent regarding the best position of the reference electrode. Since currents coming from the bio-electrical activity of muscles, heart, or brain, propagate all over the human body, the reference electrode has to be placed in proximity of the brain: on the earlobe, nose, mastoid, chin, neck, or scalp center. In the bipolar setup (montage) each channel registers the potential difference between two particular scalp electrodes. The "common average reference" montage is obtained by subtracting from each channel the average activity from all the remaining derivations.

1.2.1 10-20 International System of Electrode Placement

The 10-20 system or International 10-20 system is an internationally recognized method to describe and apply the location of scalp electrodes.
This method was developed to maintain standardized testing methods ensuring that a subject's study outcomes (clinical or research) could be compiled, reproduced, and effectively analyzed and compared using the scientific method. The system is based on the relationship between the location of an electrode and the underlying area of the brain, specifically the cerebral cortex.[13]

Each site has a letter to identify the lobe and a number to identify the hemisphere location.

Pre-frontal (Fp), Frontal (F), Temporal (T), Parietal (P), Occipital (O),

- The letter 'O' identifies the occipital lobe.
- The letter 'F' identifies the frontal lobe.
- The letters 'Fp' identify the pre-frontal lobe.
- The letter 'P' identifies the Parietal lobe.
- The letter 'T' identifies the Temporal lobe.
- No central lobe exists, the 'C' letter is used for identification purposes only.

- The 'z' (zero) refers to an electrode placed on the mid line.
- Even numbers (2; 4; 6; 8) refer to electrode positions on the right hemisphere.
- Odd numbers (1; 3; 5; 7) refer to electrode positions on the left hemisphere

Four anatomical landmarks are used for the essential positioning of the electrodes: first, the nasion which is the point between the forehead and the nose; second, the inion which is the lowest point of the skull from the back of the head and is normally indicated by a prominent bump; the pre-auricular points anterior to the ear. Extra positions can be added by utilizing the spaces in between the existing IS 10/20 system. The IS 10/10 system is shown in figures below.

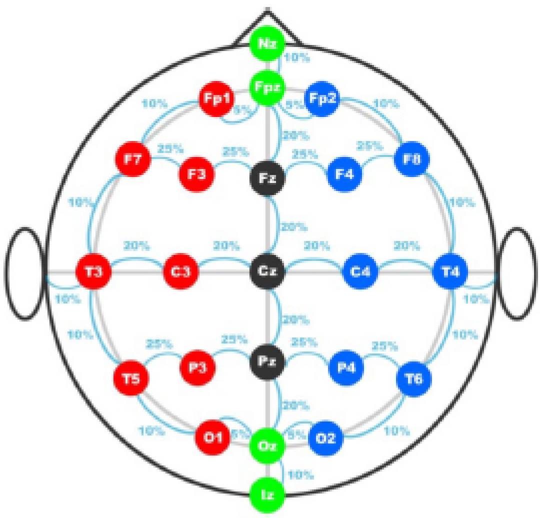


Figure 1.7: 10/20 International System of Electrode Placement (IS 10/20)

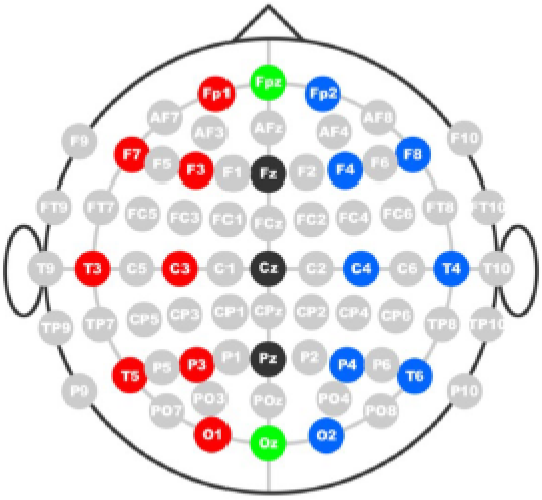


Figure 1.8: 10/10 International System of Electrode Placement (IS 10/10)

1.3 Noise in EEG recordings

EEG recording, like any other neurophysiological signal, is highly corrupted by many forms and sources of noises that are significantly stronger than the signal itselfs.

There are some strategies to deal with noise in EEG recordings, some of which have been employed within this work of thesis and are described here in the following.

1.3.1 Sources of noise

Sources of noise in electroencephalographic signals are also known as artifacts. As well as the noise, artifacts are considered unwanted signals that degrade the quality of the recording. The main sources that leads to artifacts can be classified into:

1. External Artifacts (due to acquisition system or to external electromagnetic interference i.e. arising from extra-cerebral sites).

2. Internal Artifacts (or biological artifacts)

1.3.1.1 External artifacts

External artifacts are due both to acquisition system or to electric interference cause by external electrical equipment.
In the following the major cause of external artifact will be resumed.

1.3.1.1.1 Power line artifacts These artifacts are due to the power line and it is a rapid continuous (in time) activity whose spectrum is centered at 50 Hz (60 Hz for Anglo-saxon and for some eastern countries).
Power line artifact are not due to the recording equipment, because they usually employ a notch filter to eliminate it, but electric cable inside the walls produce a constant electric field.

1.3.1.1.2 Mobile phones artifacts These artifacts are due to the presence of a mobile phone in the recording room. They appear as spike waves discharges at 30 Hz and may resemble an electroencephalographic seizure.

1.3.1.2 Internal Artifacts

1.3.1.2.1 Eye movement Eye movement artifacts are quite often the largest artifact in EEG recording. They are generated by vertical and horizontal eye movements.
The main source of the artifacts is the potential of the eyeball. The eyeball

acts as an electric dipole with the positive pole oriented anteriorly.

Eye blink results in reflexive upward vertical eye movement that produces positive deflection at frontal areas with maximum at Fp1, Fp2 electrodes.

Eyes closing is associated with a similar artifact, while eyes opening results in downward vertical eye movement and negative deflection at Fp1, Fp2 electrodes.

Horizontal eye movements (also called saccades) produce opposite changes of potentials at F7, F8 electrodes. Figure 8.18 represents a sequence of horizontal eye movement (saccade) and eye blink.

1.3.1.2.2 Muscle Artifact Muscle artifacts arise from electrical activity of muscles. In particular, frontalis and temporalis muscles are the most common source of myogenic activity respectively in frontal electrodes (mostly Fp1 and Fp2) and in temporal electrodes (mostly T3, T4).

Usually, it is not difficult to separate muscle activity from beta cortical activity. Indeed, at the spectra the range of muscle artifact is usually broader than the range of beta activity. Because of that, at recordings muscle activity looks like a thicker line when compared with genuine EEG. Single muscle discharges may look like epileptic spikes, but muscle "‘spikes"’ are shorter in duration and are limited to only one electrode.

1.3.1.2.3 ECG Artifact In individuals with short necks and large hearts electrical fields may be detected by ears or other basal electrodes. It is difficult to confuse ECG artifacts with epileptic spikes because these artifacts are regular and usually seen with the same polarity in many electrodes. Simultaneous ECG recording usually helps to differentiate these artifacts (labeled as ECG artifacts), but an experienced electroencephalographer can easily do it without such recording.

1.3.1.2.4 Cardio-Ballistic Artifact Another common type of non-brain-related potential changes is called cardioballistic artifact. This type of artifact is caused by a periodic (with a period of heart beating) movement of electrode located just above a blood vessel of the head. Pulsation of the vessel moves the electrode which induces a periodic artifact. The cardio-ballistic artifact is usually observed under one electrode. This is the reason why it can be better seen when a local average montage is applied. This artifact can be easily detected on the map of EEG spectra as a local peak in about 1 Hz frequency.

1.3.1.3 Elimination of noise sources

Some source of noise can be easily avoided. This is the case of external, environmental sources of noise such as AC power line electronic equipment (displays, mobile phones , routers). The first thing that can be done to avoid this sources of noise is removing an unnecessary source of electromagnetic noise from the recording room.

In principle it would be better to insulate the recording room by use of a Faraday cage, but EM insulation requires either advance planning or costly work. Another source of noise are movement artifacts due to muscle contraction. EMG noise can be avoided or at least mitigated by asking the patient to find a comfortable position and relax before the recording session. EOG noise generated by eye movements or blinks is another impairment to the quality of the EEG recordings. In order to avoid artifacts due to eye movement it patient could be encouraged to hold gaze in the same location, e.g. in our experiment the participants were asked to stare a timer. For what concerns blinking, asking the participant not to blink could be very challenging for him and moreover since both blinking and spontaneous eye movement are unconscious behaviour and therefore withholding either of them requires voluntary attention that might interact with the task performance introducing a further EEG signal component.

1.3.1.4 Rejection of noisy data

Whenever noise in the recorded data is easily recognizable, the easiest way to get rid of it is to eliminate the portion of data where the noise is easily detectable. Rejection of noisy data can be straightforwardly applied to the study of ERP because, usually in order to compute the potential a task is performed several times in what is called epoch. Therefore in this case it is easier to isolate and discard epoch which are visibly affected by noise. Rejection can be performed relying on visual inspection (which is obviously not feasible with large datasets) or by using different techniques, some of them will be resumed here in the following. Channels with significant level of noise are often characterized with high power at high frequency or they present spikes at power line frequencies (50 or 60 Hz). Noisy channels also show a higher variability in the signal across time with respect to other channels. The easiest methods to automatically identify noisy recording is the detection of extreme values either by setting a threshold or considering the statistics or again abnormal frequency spectrum.

1.3.1.4.1 Removal of noise There are many technique used to remove noise from EEG data. In the following will be considered mainly the techniques employed for this work of thesis.

Filtering One of the easiest way remove noise is by filtering it out. Obviously to let the filtering possible, noise needs to fall outside the frequency spectrum of the signal of interest, meaning that it has to fall below or above the spectrum of the electroencephalographic signal. For instance, muscle contraction typically leads to strong signal component above 100 Hz, which is for sure outside the spectrum of interest and therefore can be eliminated without risking to filter part of the useful signal.
Filtering is also used to cope with the aliasing due to the sampling of the data.

Signal averaging Another important method to increase the signal to noise ratio and improving the quality of the noisy electroencephalographic signal is averaging several measure of the same signal. The averaging procedure is based on the assumption that the noise is independent throughout the measurements and, above all, the signal is constant over all the trial. Therefore signal averaging can be used only in the case in which we are able to collect pattern in the data that reoccur under certain condition. This hypothesis is implicit in the study of event related potentials (ERPs). The averaging of the ERP are based on the following assumptions:

1. The brain solves the same task in a very similar way and therefore the signal is (almost) the same in all trials.

2. Ongoing brain activity (not task related) is independent in each trial

3. Noise is a zero-mean random process independent among trials

Under this assumption the average ERP is computed as follows:

$$< s_i(t) >= \frac{1}{N}\sum_{i=1}^{N} s_i(t) \qquad i = 1, \ldots, N. \tag{1.1}$$

where $s_i(t)$ is the measured signal in the i-th trial at time t.
Note that to simplify the notation it is chosen to use $s(t)$ even for the discrete data $s(1) \ldots s(n)$ where n is the number of samples. Considering an additive noise having zero mean ($< \epsilon(t) >= 0 \forall$ t), then it is possible to represent the data according with the following model (M. Ihrke et al. (2000) [11])

$$s_i(t) = u(t) + \epsilon_i(t) \tag{1.2}$$

where $u(t)$ represents the recovered signal from s_i. This is the so-called Signal-Plus-Noise (SPN) model. Therefore, ideally, averaging the signal over a sufficient number if trials eliminates the noise leaving the constant signal intact. Although this model is employed in most of the research on the event-related potential it has often been questioned, mainly for the hypothesis of stationarity. In fact repetition of a task may be accompanied by different neural activity, either because of setting-dependent (e.g. slightly different displays in the same experimental condition) or subject-dependent variations (e.g. growing tiredness or learning).

Chapter 2

Materials and method

2.1 Materials

2.1.1 Recordings

The EEG datasets used for this work of thesis have been acquired during a previous study about Disorder Of Consciousness, focusing in particular on the intentionality of the movement based on the Libet's paradigm.
The subjects sat in a chair looking at a LCD monitor placed approximately 1 m in front of them. Subjects were instructed to observe a timer and perform a brisk forefinger flexion (of the dominant hand) every 10 s starting from 5s. The 40 finger flexions were perfomed for a session of 400 s.

EEG was recorded by 7 monopolar scalp Ag/Ag/AgCl electrodes according to the standard 10-20 system referenced to the ground over A1 and A2. The precise electrode placements used in our study is shown in Figure 2.1 and Figure 2.2.

Since the LRP is recorded over the frontal, prefrontal and parietal cortex, in correspondence of SMA and M1 areas, 7 channels have been used: Cz, C3, C4, for the central lobe; Fcz, Fc3, Fc4 for the frontal lobe; Pz for the parietal lobe; Oz for the occipital lobe.

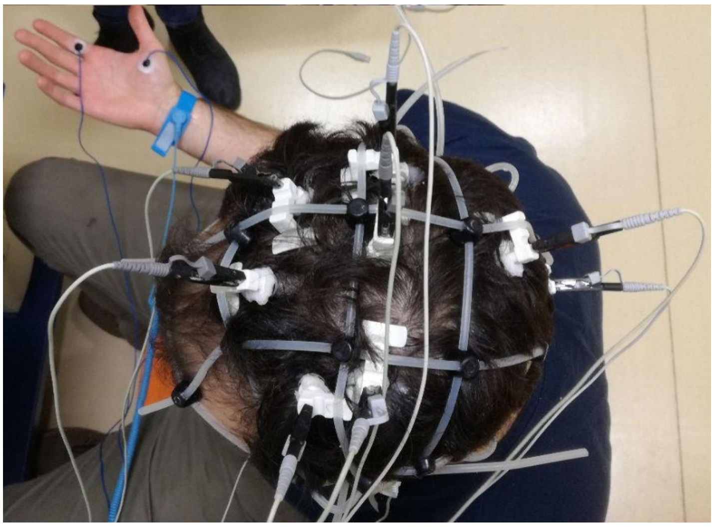


Figure 2.1: EEG electrodes

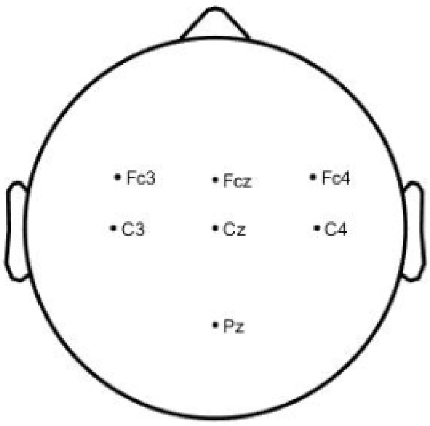


Figure 2.2: EEG electrodes locations

Figure 2.3: EMG electrodes

Figure 2.4: Experimental design.

The EEG was amplified , bandpass filtered between 0.015 Hz and 50 Hz and then sampled at 512 Hz with a time constant of 0.1 s.
Only the 15 healthy subject EEG datasets was considered: 4 males, 11 females, right-handed, aged between 21 and 26.
During the datasets acquisition it was not possible to acquire the electroocu-

logram (EOG), therefore any artifact removal algorithm based on the regression of the EOG channel has been adopted and consequently the recordings may be impaired by ocular artifacts.

Surface EMG was recorded from the right hand palm by means of two electrodes, one above the thenar eminence and the other one in correspondence to the first joint of the index finger. sEMG was amplified , bandpass filtered between 5 Hz and 500 Hz and then sampled at 512 Hz with a time constant of 0.3 s. Both EEG and sEMG were notch filtered at 50 Hz to remove the power line noise.

The impedance of each electrode was kept above $5K\Omega$.

2.1.2 Instrumentation

The EEG measurements were acquired using GalileoNT (EEG NT), B8300033000, EBNeuro([14]).

Figure 2.5: GalileoNT

The parameters related to EMG was set as follows:

- Range: $65mV$,
- Sample frequency: $512Hz$,

- Pass-band filter cut-off frequencies: $[500.0...5.0]Hz$,
- Notch filter cut-off frequency: $50Hz$.

The parameters related to EEG was set as follows:

- Range: $4mV$,
- Sample frequency: $512Hz$,
- Pass-band filter cut-o frequencies: [50.000...0.015] Hz,
- Notch filter cut-off frequency: $50Hz$

The electrodes impedance was kept below $5k\Omega$.

2.1.3 Software

2.1.3.0.1 EEGLAB EEGLAB is an interactive Matlab toolbox, for processing continuous and event-related EEG, MEG and other electrophysiological data incorporating independent component analysis (ICA), time/frequency analysis, artifact rejection, event-related statistics, and several useful modes of visualization of the averaged and single-trial data[15].

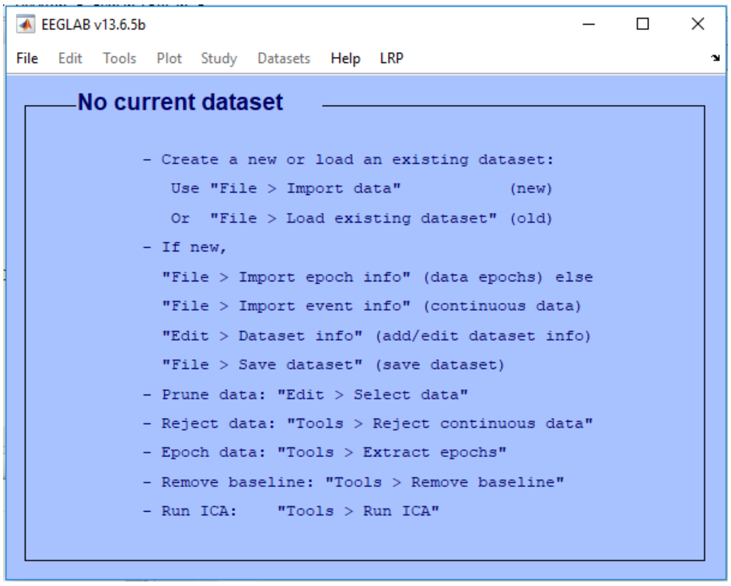


Figure 2.6: EEGLAB Graphic User Interface

Furthermore, EEGLAB provides an interactive graphic user interface (GUI), for visualizing event-related brain dynamics and also a structured programming environment for storing and manipulating event-related EEG data. EEGLAB is distributed under the free GNU GPL license.

2.1.3.0.2 LRPLAB During this work of thesis a plug-in for EEGLAB has been developed. This plug-in allowed us to integrate all the algorithms employed with EEGLAB.

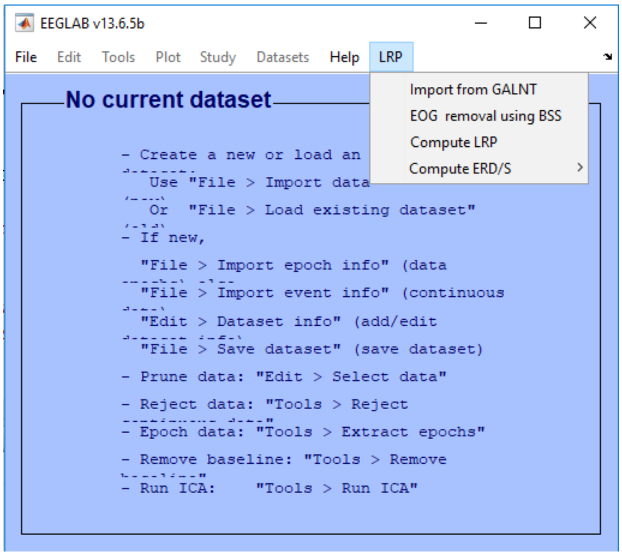


Figure 2.7: LRP drop down menu

In particular in the drop down menu the element "'Import from GALNT"' is used to import a dataset in the format produced by Galileo NT and make it compatible with EEGLAB. The the element "'Compute LRP"' is used to actually compute the Lateralized Readiness Potential.

2.2 Methods

2.2.1 Pre-processing

In the following sections all the pre-processing phase will be treated in detail for both the electroencephalographic and electromyographic signal.

2.2.1.1 EMG Onset Detection

Electromyographic signal is used in our experiment in order to align all the EEG traces for each trial to a specific event.

In this case the specific event could be of two kinds:

1. EMG Peaks
2. EMG Onset

Aligning the EEG traces on the peaks of the EMG would be the easiest and immediate procedure to adopt, but since we are interested in a event-related potential that take place before the movement execution it is much more suited to compute the onset of the movement and therefore the onset of the EMG.

Nevertheless, the peaks are still detected and they will be used in all the cases in which the movement in most trials are not brisk, and this could be the case of patients with movements disorders.

2.2.1.1.1 Acquisition As reported in the section Instrumentation, a preliminary filtering stage is performed during the acquisition phase. In our case for both EMG and EEG a pass band filter was employed. The cut-off frequencies for the EMG filter are 5.0 and 500.0 Hz, while for the EEG signal the frequencies are 0.015 and 50000 Hz. For both signal a notch filter with cut-off frequency 50 Hz has been employed in order to remove the power line interference component.

2.2.1.1.2 Resampling Resampling, more precisely downsampling or decimation, is done in order to reduce the sampling frequency from 512 Hz to 128 Hz.

This operation is done as a first approach to cope with *aliasing*. Another advantage is to a have a minimum reduction of the background noise [12]

2.2.1.1.3 Detrending

Linear detrending has been used to remove linear trend in each epoch.

Detrending involves the computation of the straight lint that better interpolate the signal and the subtraction from the signal itself.[12] In this way the continuous component of the EMG signal is removed without using any other high pass filter that could create artifact in the signal.

2.2.1.1.4 Hilbert Transform

The Hilbert transform is a specific linear operator that takes a function, u(t) of a real variable and produces another function of a real variable $H(u)(t)$. This linear operator is given by convolution with the function $\frac{1}{\pi * t}$ [16]

2.2.1.1.5 Smoothing

The EEG signal, divided in epoch of 10 seconds, is then smoothed by using two moving average filters, one moving from the beginning of the epoch to the end, and the other one on the opposite direction. The two signal obtained are then averaged together.

2.2.1.1.6 Linear Regression

In order to compute the slope of the electromyographic signal Linear regression is employed.

Linear regression models the relation between a dependent, or response, variable y and one or more independent, or predictor, variables $x_1, x_2, ...x_n$. Simple linear regression considers only one independent variable using the relation

$$y = mx + q + \epsilon \tag{2.1}$$

where q is the y-intercept, m is the slope (or regression coefficient), and

ϵ is the error term.

2.2.1.1.7 Movement Onset Detection Movement onset and end are detected by means of a threshold detection algorithm. First of all, the signal is normalized with respect to the the standard deviation:

$$y(t) = \frac{x(t) - \mu}{\sigma} \tag{2.2}$$

where $x(t)$ is the signal, μ is the signal mean, and σ is the standard deviation. The threshold is computed as the 96-percentile and the signal onset is chosen as the first sample to exceed the this threshold. In the same way, the movement end is computed as the sample that exceed a 80-percentile threshold.

2.2.1.2 EEG Preprocessing

2.2.1.2.1 Resampling As for the electromyographic signal, the electroencephalographic signal is downsampled to $128Hz$.

2.2.1.2.2 Detrending Linear detrending has been used to remove linear trend in each electroencephalographic epoch.

2.2.1.2.3 Filtering Two low-pass filter are employed in cascade. The first filter has the following characteristics:

- Cut-off frequency $12.5Hz$
- Stop frequency $13.5Hz$

- Kaiser window with order 204

The second low pass filter instead has:

- Cut-off frequency $1.5Hz$
- Stop frequency $2Hz$
- Kaiser window with order 408

Both filters' group delay are computed and compensated by means of a circular shift in each epochs.

2.2.1.2.4 Averaging In order to increase the signal to noise ratio and improving the quality of the noisy electroencephalographic signal is averaged over all the epochs. As already stated in the introduction of this thesis, the averaging procedure is performed according to the following assumptions:

1. The brain respond in the very similar way when performing the same task, in this particular case it is expected that the motor response to the finger flexion is the same during all the epoch.
2. Ongoing brain activity (not related to the finger movement) is independent in each trial
3. Noise is a zero-mean random process independent among trials

Under this assumption the average ERP is computed as follows:

$$< s_i(t) >= \frac{1}{N}\sum_{i=1}^{N} s_i(t) \qquad i = 1, \ldots, N. \tag{2.3}$$

where $s_i(t)$ is the measured signal in the i-th epoch at time t.

Therefore, ideally, averaging the signal over a sufficient number if trials eliminates the noise leaving the constant signal intact. It has been observed during the EEG recordings that, the finger movement was not always performed correctly, either for some background trembling or for a movement executed in more than one step.

2.2.2 Epoch discarding

In order to further improve the Signal-to-Noise ratio, epochs in which the finger movement is not executed properly are discarded. This operation is mainly done because, computing the onset of the EMG is very important since the allignment of the EEG epochs is performed according to the onset itself. It has been observed that when the movement is not executed properly:

1. It becomes very hard to determine the movement onset, either for the threshold detection algorithm, for the way it has been implemented, and even for a visual inspection.

2. The task is not compliant with the experiment, differing in the way it has been performed and thus accompanied by different brain activity.

Frequently, in other neurophysiology study these type of epochs are visually inspected to be discarded, as well as epochs affected by artifact. In this work of thesis, it has been decided to automate this procedure by implementing a classification algorithm able to, with a certain precision, identify the epoch where movement is not executed properly

2.2.3 Classification

A classification problem occurs when an object needs to be assigned into a predefined group or class based on a number of observed attributes related to that object. The individual observations are analyzed into a set of quantifiable properties known as explanatory variables, or features. These properties may variously be categorical, ordinal, integer-valued or real-valued.
During classification given objects are assigned to prescribed classes. A classifier is a mathematical function implemented by a classification algorithm, that maps input data to a category which performs classification[17].
Said it otherwise, classification is the problem of identifying to which of a set of categories(sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known.
In the terminology of machine learning, classification is considered an instance of supervised learning i.e. learning where a training set of correctly identified observations is available.

An algorithm that implements classification, especially in a concrete implementation, is known as a classifier. The term "classifier" sometimes also refers to the mathematical function, implemented by a classification algorithm, that maps input data to a category [18]

In this thesis two algorithm were employed in order to classify EMG traces and discriminate between properly executed movement from those performed badly, e.g. the case in which the finger flexion is not brisk or executed in two steps or those in which there is a background trembling.

These two algorithms are:

1. Linear Discriminant Analysis

2. Support Vector Machine

This step is important in order to discard all the trials in which the finger flexion is not properly executed, and in this way try to improve the signal to noise ratio of the average ERP.

2.2.3.1 Support Vector Machine

A Support Vector Machine (SVM) is a Machine Learning algorithm which learn from a training set and tries to generalize and make correct prediction of novel data. For the training stage we have a set of m input vectors x_i, each with a number of components called features.

These input vectors are paired with m corresponding labels, denoted by yi. The training data can be seen as labeled points in an input hyperplane and the learning task for a two classes of well separated datapoint aims to find a directional hyperplane that divides the hyperspace in such a way that all the datapoints of each class lays on their half hyperspace.

The hyperplane found by SVM is the one that maximizes the distance from the two classes of labeled points on each side.

The closest points to the hyperplan are the ones that most influences the position of the hyperplane itself and therefore are known as support vectors.

In the best situation the training set is made up of well separated clusters of point each of them representing a class of input data.

In real world application, the clusters might be highly intermeshed with overlapping points: in this case the dataset is not linearly separated. The separating hyperplane is given as $w\Delta x + b = 0$ (where $\cdot$ denotes the scalar product). b is the bias or offset of the hyperplane from the origin in input space, x are points located within the hyperplane and the normal to the hyperplane, the weights w, determine its orientation. If we consider a binary classification task with datapoints $x_i(i = 1, ..., m)$ having corresponding labels $y_i = + - 1$ the decision function is:

$$f(x) = sign(w * x + b) = y$$

The hyperplanes passing through $w\Delta x + b = 1$ and $w\Delta x + b = -1$ are known as canonical hyperplanes, and the region between these canonical hyperplanes is called the margin band. Geometrically, the distance between these two hyperplanes is $\frac{2}{||w||}$, and it is auspicable that such hyperplane are as far away as possible. In order to maximize this distance we have to minimize $||w||$ under the constraint $y_i(w\Delta x_i + b) = 1 \forall i$.

As a constrained optimization problem, the above formulation can be solved through the method of Lagrange multipliers.

As a consequence of this geometric description, the max-margin hyperplane is completely determined by the support vectors.

After the training stage our SVM has to predict a class y_i to any other x_i given as input.

2.2.3.2 Linear Discriminant Analysis

Linear Discriminant Analysis is a generalization of Fisher's linear discriminant analysis and it is a method used to characterize or separate two or more classes of objects[19].

LDA is employed in the case where the within-class frequencies are unequal and their performances has been examined on randomly generated test data. This algorithm, in fact, allows to maximize the variance between classes, and to minimize the variance within classes, in any particular data set, guaranteeing maximal separability among classes[20]. LDA uses a hyperplane to separate different classes, this hyperplane is attained by finding the projection that maximizes the distance between the class means and minimizes the classes variance. Due to the low computational cost, LDA classifiers have been widely used in real-time Brain Computer Interface (BCI) systems[28].

Summarizing the general steps for performing a linear discriminant analysis are [29]:

1. Compute the d-dimensional mean vectors for the different classes from the dataset.

2. Compute the scatter matrices (in-between-class and within-class scatter matrix).

3. Compute the eigenvectors $(e_1, e_2, ..., e_d)$ and corresponding eigenvalues $(\lambda_1, \lambda_2, ..., \lambda_d)$ for the scatter matrices.

4. Sort the eigenvectors by decreasing eigenvalues and choose k eigenvec-

tors with the largest eigenvalues to form a dxk dimensional matrix WW (where every column represents an eigenvector).

5. Use this $d*k$ eigenvector matrix to transform the samples onto the new subspace. This can be summarized by the matrix multiplication: $YY = XX \times WW$ (where XX is a $n*d-dimensional$ matrix representing the n samples, and yy are the transformed $n \times k - dimensional$ samples in the new subspace).

2.2.3.3 LDA vs SVM

In all the tests that have been conducted within this study, the LDA classifier outperformed SVM in target detection accuracy and robustness. Moreover, compared to the SVM, LDA classifiers require far less computation in the training process, making it more suitable for further developments of the study involving real-time systems. Furthermore, it has been observed during this study that SVM for the classification of the correct performance of the finger movements has been too severe, leading too a high rate of discarded epoch and a consequent decrease of the Signal-to-noise-ratio due to a lower number of signal averaged. For all these reasons, LDA has been preferred to SVM and therefore finger movements are classified by means of LDA. In particular the LDA classifier was instructed by means of a training set made up of 200 epochs. In the figure above there are six examples of electromyographic signal corresponding to 6 epochs, three of which are considered good execution of the finger flexion (left) and three bad execution (right).

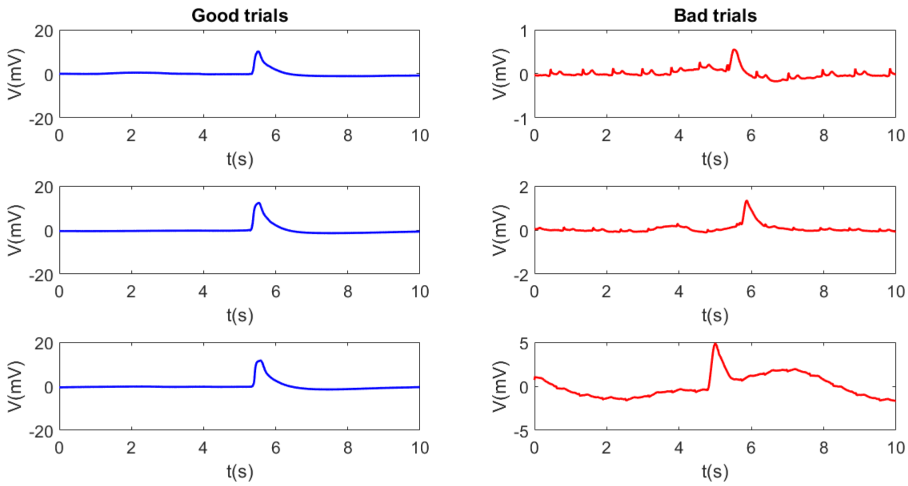


Figure 2.8: Electromyographic epochs

2.2.4 Dimensionality Reduction

Real world data, such as financial data, medical data (e.g DNA sequences , fMRI scans or EEG recordings) usually has a very high dimensionality. In order to cope with such kind off data, dimensionality reduction is needed. Dimensionality reduction is the process of reducing the number of random variables under consideration, by obtaining a set of principal variables [22]. For this work of thesis, the Principal Component Analysis is employed to analyze the epoch time series and represent it in a more meaningful and compact manner. In our classification problem there are too many factors (128 samples per second) to be used as features of the classifier, and moreover these factors are also correlated to each other and therefore redundant. This is a typical case in which dimensionality reduction can come into play.

2.2.4.1 Principal Component Analysis

Principal Component Analysis, PCA, (also known as *Karhunen-Loève* transform) is a statistical technique that linearly transforms an original set of variables into a substantially smaller set of uncorrelated variables that represents most of the information in the original set of variables. If the variables are correlated, then they can be linearly transformed into a smaller set of variables so that the resulting set represent almost the totality [23].
The basic idea behind Principal Component Analysis is to reduce the dimensionality of a data set while retaining as much information as possible.
This is achieved by transforming the original dataset to a new set of variables, namely the principal components (PCs), which are, as already stated, uncorrelated. Moreover these principal component are ordered so that the first few retain most of the variation present in the original dataset, and therefore most of the information is these few variable rather than being sparse in the whole collection of variables.[25]
Principal component analysis therefore is used to enhance the understanding of the structure of a dataset and it is employed in application such as dimensionality reduction, feature extraction,data compression and data visualization.

Given a set of variables on p dimensions, PCA uses a linear transformation to find direction of maximum variance, so that the information contained in the original set of variables can be summarized through a smaller of number of uncorrelated variables retaining most of the information.

Said it otherwise, PCA searches for a few uncorrelated linear combination of the original variables that captures most of the information in the original values.

Suppose that **x** is a vector of p random variables,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \tag{2.4}$$

and that the variances of the p random variables is of interest.

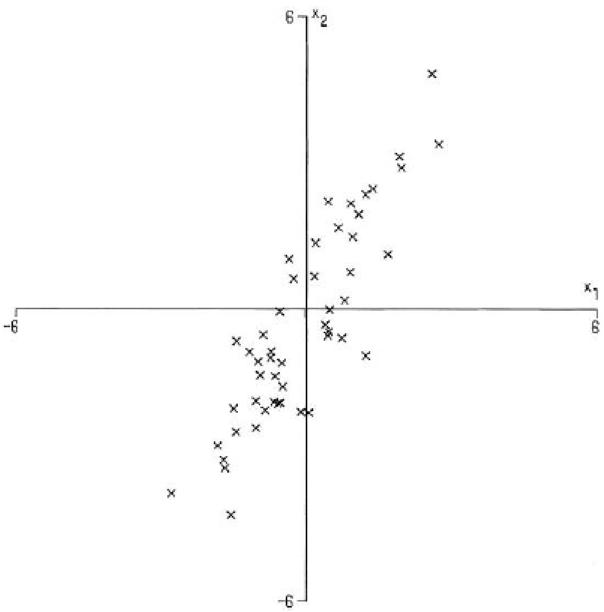


Figure 2.9: Plot of 50 observations on two variables x1,x2. Figure adopted from [25].

The first step is to look for a linear function $\alpha_1^T x$ of the elements of x

having maximum variance, where α_1 is a vector of p constants $\alpha_{11}, \alpha_{12}, ..., \alpha_{1p}$, so that

$$\alpha_1^T x = \alpha_{11}x_1, \alpha_{12}x_2, ..., \alpha_{1p}x_p = \sum_{j=1}^{p} \alpha_{1j}x_j \tag{2.5}$$

Then the algorithm look for a linear function $\alpha_2^T x$ uncorrelated with $\alpha_1^T x$ having maximum variance, and so on, so that at the k_{th} stage a linear function $\alpha_k^T x$ is found that has maximum variance subject to being uncorrelated with $\alpha_1^T x$,$\alpha_2^T x$, ... $\alpha_k^T x$.
The k_{th} derived variable, $\alpha_k^T x$ is the k_{th} PC. Up to p PCs could be found, but it is hoped, in general, that most of the variation in $\mathbf{x}$ will be accounted for by m PCs, where $m << p$.

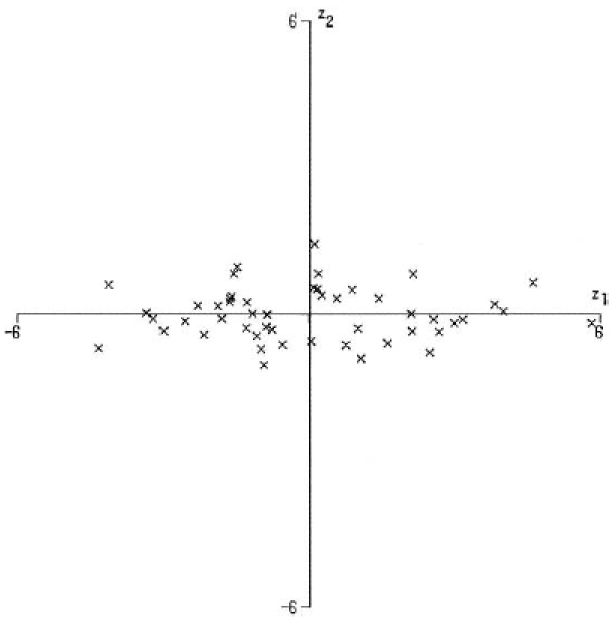


Figure 2.10: Plot of the 50 observations with respect to their PCs z_1, z_2. Figure adopted from [25].

Figure 2.9 gives a plot of 50 observations on two highly orrelated variables

x_1, x_2. There is considerable variation in both variables, though rather more in the direction of x_2 than x_1. If we transform to PCs z_1, z_2, we obtain the second plot (Figure 2.10).

It is clear that there is greater variation in the direction of z_1 than in either of the original variables, but very little variation in the direction of z_2. More generally, if a set of $p(> 2)$ variables has substantial correlations among them, then the first few PCs will account for most of the variation in the original variables. On the other hand, the last few PCs identify directions in which there is very little variation; that is, they identify near-constant linear relationships among the original variables[25].

In this work of thesis Principal Component Analysis has been employed to reduce the dimensionality of an epoch time series. This operation is done to eliminate redundancy in the signal and let the classifier use only few principal components instead of trying to classify the epoch basing on the entire time series, that in this case is made up of 1280 samples (128 samples/s x 10 s).
In order to compute the Principal Components the electromyographic signal is first of all normalized. This normalization is done by subtraction the signal average and by dividing by the signal standard deviation :

$$y_i = \frac{x_i - \mu}{\sigma} \forall i \in [1...1280] \tag{2.6}$$

where i is the $i - th$ sample of the time series. Then the the principal components are computed, and the first five are used as features for the classifier.

2.2.5 Feature Extraction

Apart from the EMG time series principal component, other features were considered and employed to classify the finger movements and will be resumed in the following.

2.2.5.0.1 Peakness The so called peakness is computed as:

$$\frac{max(x(t))}{t_{end} - t_{onset}} witht \in (0, 10) \tag{2.7}$$

where x(t) is the electromyographic signal within one epoch, t_{onset} in the time instant in which the movement onset is identified, and in the same way, t_{end} is the time instant in which the movement end is identified.
This feature gives the classifier the information on how the movement is executed rapidly, and it is employed in the classification procedure because it has been observed that good EMG epochs where always flat with a high and narrow peak.

2.2.5.0.2 Peak to average power ratio The so called Peak to average power ratio is computed as the ration between the power of the maximum of the electromyographic signal and the its mean power:

$$\frac{P_{max}(t)}{P_{mean}(t_{off}} = \frac{max(x(t))^2}{x(\bar{t})^2} \tag{2.8}$$

This feature gives the classifier the information of how the power is distributed in the epoch. In the movement executed properly the power of the peak is much more higher than the mean power (which is usually too high in the epochs where the movement is not properly performed).

2.2.5.0.3 Second Maximum The second maximum is computed as the maximum of the electromyographic signal outside the interval $(t_{onset}; t_{end})$

$$max(x(t))t \in (t_{onset}, t_{end}) \tag{2.9}$$

The second maximum has to be as small as possible in order to have a movement properly executed. This feature has been discarded in the last analysis, because it is redundant since the peak to average power ratio is employed and it has been proven (K-Fold Validation) that the last outperformed the first.

Chapter 3

Results and conclusions

In this chapter the results attained and the conclusions that have been drawn will be presented.

3.1 EMG classification

The electromyographic epochs have been classified in order to discard those epochs in which the movement are not properly executed. In order to classify the EMG epochs two different classification algorithm were employed:

1. Support Vector Machine
2. Linear Discriminant Analysis

Both of them relied on the same features to perform the classification:

- The first 5 principal component
- The Peakness

- The Peak to average power ratio

These features were chosen among several other by evaluating the classifier performance using different combination of them. The classifier performance were evaluated by mean of the Cross-validation.
Cross-validation is a model assessment technique used to evaluate a machine learning algorithms performance in performing classification.
This is done by partitioning a dataset and using a subset to train the algorithm and the remaining data for testing.
Each round of cross-validation involves randomly partitioning the original dataset into a training set and a testing set[26].
The Cross-validation technique employed to validate the EMG epoch classification is the Leave-one-out.
Leave-one-out cross-validation technique is a specific case of the k-fold validation in which the original dataset is randomly partitioned into k equal sized subsamples. Of the k subsamples, a single subsample is retained as the validation data for testing the model, and the remaining k - 1 subsamples are used as training data.
The cross-validation process is then repeated k times, with each of the k subsamples used exactly once as the validation data. The k results can then be averaged to produce a single estimation[27].
In leave-one-out technique k is equal to 1, therefore in this case 199 epochs are used as training set a 1 as the test set. This procedure is done as many times as the number of epochs.
Obviously the 200 samples of the training set was chosen to represent all the possible cases in order to instruct as good as possible the classifiers.

The results attained with the two classifier are resume below by their confusion matrix:

$$\begin{bmatrix} 91 & 9 \\ 8 & 92 \end{bmatrix} \tag{3.1}$$

$$\begin{bmatrix} 99 & 1 \\ 1 & 99 \end{bmatrix} \tag{3.2}$$

where the first matrix is the confusion matrix of the Support Vector Machine classifier, while the second matrix is the Linear Discriminant Analysis classifier one.
A confusion matrix, also called table of confusion, resumes the performance of a supervised learning algorithm.
In particular

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix} \tag{3.3}$$

TP is the number of true positive outcome, FP the false positive, FN the false negative and finally TN the true negative.
As it can be seen, the Linear Discriminant Analysis classifier performed much more better with respect to the Support Vector Machine, moreover as both

the theory and the experiment confirms, LDA is much more faster than SVM. In this case of offline processing of the EEG signal, performance in terms of processing time are not really of interest but in real time application, as this study is intended to go toward, this is a key feature.
For this reasons LDA has been preferred over the SVM.

3.2 Conclusions

Within this work of thesis several improvements were brought to the computation of the Lateralized Readiness Potential. All these improvements however were focused on focused on the increasing of the Signal to noise ratio in the conventional manner to compute a LRP, namely the averaging procedure.
Being the intention of this study the investigation of whether LRP could be used as a diagnostic tool in motor and consciousness disorders the averaging technique is not suitable for further developments.
In fact, asking a patient in one of these conditions is certainly unfeasible . Moreover, having adopted the above illustrated experimental protocol, it could happen that also the CNV potential is elicited.
The running time, shown by the timer on monitor's screen, could be act as an imperative and preparative stimulus for the subject, that could pre-planned the flexion of the right forefinger.
Therefore, further developments should be focused on an online computation of the Lateralized Readiness Potential, and on its classification to distinguish if from non voluntary movements, for instance spasms.

Bibliography

[1] Juri D. Kropotov, *Functional Neuromarkers for Psychiatry*, 2016

[2] $https://en.wikipedia.org/wiki/Event-related\_potential$

[3] Niedermeyer E.; da Silva F.L. (2004). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* Lippincott Williams Wilkins.

[4] Kataryzna J Blinowska and Jaroslaw Zygierewicz. *Practical Biomedical Signal Analysis Using MATLAB.* Ed. by CRC Press. illustrate. 2012

[5] Oriano Mecarelli, *Manuale teorico pratico di Elettroencefalografia*,2009,Wolters Kluwer Health Italia.

[6] Luder Deecke, Berta Grozinger, and H. H. Kornhuber,*Voluntary finger movement in man: cerebral potentials and theory*,In: Biological Cybernetics (1976).

[7] Hiroshi Shibasaki and Mark Hallett. *What is the Bereitschaftspotential?* In: Clinical Neurophysiology 117.11 (2006), pp. 2341-2356.

[8] M G Coles, *Modern mind-brain reading: psychophysiology, physiology, and cognition*, 1989.

[9] R De Jong et al. *Use of partial stimulus information in response processing*. In: Journal of Experimental Psychology: Human Perception and Performance 14.4 (1988), pp. 682-692.

[10] Martin Eimer. *The lateralized readiness potential as an on-line*. In: (1998), pp. 146-156.

[11] Matthias Ihrke, Hecke Schrobsdorff, and J Michael Herrmann. *Denoising and Averaging Techniques for Electrophysiological Data*. In: (2000), pp. 1-23.

[12] Oriano Mecarelli,*Manuale teorico pratico di Elettroencefalografia*,2009

[13] *https://en.wikipedia.org/wiki/10%E2%80%9320_system_(EEG)*

[14] *http://win.ebneuro.biz/NeuroProdottiEegBeFamily.asp*

[15] *https://sccn.ucsd.edu/eeglab/*

[16] *https://en.wikipedia.org/wiki/Hilbert_transform*

[17] Bamparopoulos Giorgos, *Statistical Classification: A review of some techniques*, Aristotle University of Thessaloniki, Slides

[18] *https://en.wikipedia.org/wiki/Statistical_classification*

[19] *https://en.wikipedia.org/wiki/Linear_discriminant_analysis*

[20] S. Balakrishnama, A. Ganapathiraju, Jan 1998, *LINEAR DISCRIMINANT ANALYSIS - A BRIEF TUTORIAL*, Department of Electrical and Computer Engineering Mississippi State University

[21] Tiam Lan, *Feature extraction feature selection and dimensionality reduction techniques for brain computer interface* , 2011

[22] *https : //www.geeksforgeeks.org/dimensionality − reduction*

[23] George H. Dunteman, *Principal Components Analysis (Quantitative Applications in the Social Sciences)*,1989

[24] Shravani Sur and V. K. Sinha, *Event-related potential: An overview*,2009

[25] Joliffe I.T., *Principal Component Analysis*, Springer 2nd edition, 2002.

[26] *https : //it.mathworks.com/discovery/cross − validation.html*

[27] *https : //en.wikipedia.org/wiki/Cross − validation$_($statistics)*

[28] Tian Lan, *Feature extraction feature selection and dimensionality reduction techniques for brain computer interface*,Oregon Health and Science University.

[29] *https : //sebastianraschka.com/Articles/2014$_p$ython$_l$da.html*