Chapter 4

The Histogram of Gradient Orientations of Signal Plots

This Chapter introduces the EEG feature extraction procedure based on the Histogram of Gradient Orientations. This method is grounded on an extension and modification of the SIFT [77] Descriptor which is used computer Vision to extract and map local regions of an image. At the same time, this Chapter brings to completion the previous one, describing how to mine the information from the Plot and build a feature out of it.

4.1 Introduction

The work of Edelman, Intrator and Poggio [37] on how the visual cortex sense features was the inspiration to the development of an algorithm to identify and decode salient local information from image regions. The Scale Invariant Feature Transform (SIFT) method is composed of two parts, the SIFT Detector and the SIFT Descriptor. The first is the procedure to identify relevant areas of an image. The second is the procedure to describe and characterize a region of an image using the Histogram of Gradient Orientations [1]. The SIFT algorithm is biomimetically inspired in how the visual cortex detects shapes by analyzing orientations [37]. The patch description is also based on the Theory of Receptive Fields and other related ideas [72].

---

[1]It should not to be confused with HOG [33], the Histogram Of Gradients which is another method from Computer Vision based on similar ideas. Actually, the descriptor part of the SIFT Method has no specific name, but it is based on building a histogram of gradient orientations.

4.2 Feature Extraction: Histogram of Gradient Orientations

The basic procedure is composed of,

1. Keypoints **kp** are located on an image of a signal plot.

2. A region of an image, a patch, is established using keypoints as centers. Each patch has a horizontal S_t and vertical scale S_v, which determines the size in pixels $\mathbf{S}_x$ and $\mathbf{S}_y$, along the horizontal and vertical axis respectively.

3. From each patch, a descriptor **d** is derived which is used as a representation of the graphical information contained within the patch.

On the image generated by the procedure detailed in previous Chapter, a keypoint **kp** is placed on a pixel (x_{kp}, y_{kp}) over the image plot and a window around the keypoint is considered. A local image patch of size $\mathbf{S}_x \times \mathbf{S}_y$ pixels is constructed by dividing the window in 16 blocks. It is arranged in a 4×4 grid and the pixel **kp** is the patch center. Figure 4.1(a) shows a plot of a signal, a keypoint in red at the center and the surrounding patch.

Pixel intensity gradients can be obtained from an image by applying the Sobel filter [124] and using finite differences to obtain pixel differences on the x and y direction. Composing them as vectors, oriented gradients on each pixel can be calculated. Figure 4.1(b) and (c) show vector field of oriented gradients.

A local representation of the signal shape within the patch can be described by obtaining the gradient orientations on each of the 16 blocks and creating a histogram of gradients. In order to calculate the histogram, the interval $[0 - 360]$ of possible angles is divided in 8 bins, each one at 45 ees. Figure 4.1(d) shows a sample histogram obtained for eight orientations.

Hence, for each spacial bin $i, j = \{0, 1, 2, 3\}$, corresponding to the indexes of each block $B_{i,j}$, the orientations are accumulated in a 3-dimensional histogram h through the following equation:

$$h(\theta, i, j) = \sum_{\mathbf{p}} \omega_{\text{ang}}(\angle J(\mathbf{p}) - \theta)\, \omega_{ij}\, (\mathbf{p} - \mathbf{kp})\, \|J(\mathbf{p})\| \tag{4.1}$$

where **p** is a pixel from within the patch, θ is the angle bin with $\theta \in \{0, 45, 90, 135, 180, 225, 270, 315\}$, $\|J(\mathbf{p})\|$ is the norm of the gradient vector in the pixel **p**, computed using finite differences, and $\angle J(\mathbf{p})$ is the angle of the gradient vector. The scalar $\omega_{\text{ang}}(\cdot)$ and vector $\omega_{ij}(\cdot)$ functions

are linear interpolations used by [77] and [137] to provide a weighting contribution to eight adjacent bins. They are calculated as

$$\omega_{ij}(\mathbf{v}) = \omega\left(\frac{5\ v_x}{\Delta_s\ S_t} - x_i\right)\omega\left(\frac{5\ v_y}{\Delta_s\ S_v} - y_i\right) \tag{4.2}$$

$$\omega_{\text{ang}}(\alpha) = \sum_{r=-1}^{1} \omega\left(\frac{8\alpha}{2\pi} + 8r\right) \tag{4.3}$$

where x_i and y_i are the spatial bin centers located in $x_i, y_i \in \{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\}$. The function parameter $\mathbf{v} = (v_x, v_y)$ is a vector variable and α a scalar variable. On the other hand, r is an integer that can vary freely in the set $\{-1, 0, 1\}$ and allows the argument α to be unconstrained in terms of its values in radians. The interpolating function $\omega(\cdot)$ is defined as:

$$\omega(z) = \max(0, |z| - 1). \tag{4.4}$$

These binning functions conform a trilinear interpolation that has a combined effect of sharing the contribution of each oriented gradient between their eight adjacent bins in a tridimensional cube in the histogram space, and zero everywhere else. This procedure is important to avoid quantization issues that may appear with the histogram (i.e. elimination of intermediate values).

Lastly, on Equation 4.2 the fixed value of $\frac{5}{\Delta_s}$ allows a unit conversion from pixel to units-of-patch which goes from $-\frac{5}{2}$ to $+\frac{5}{2}$, by using the unit length of the patch Δ_s, which is described in the section 4.4. As the patch has 16 blocks and 8 bin angles are considered, a feature $\mathbf{d}$ called *descriptor* of 128 dimension is obtained. This technique is a modification of Lowe's SIFT Descriptor method.

Figure 4.1 shows an example of a patch and a scheme of the histogram computation. In (a) a plot of the signal and the patch centered around the keypoint is shown. In (b) and (c), the vector field of oriented gradients is shown. Finally, in (c) an enlarged image plot is shown where the oriented gradient vector for each pixel can be seen. In Figure 4.2 the possible orientations on each patch are illustrated. Only the upper-left four blocks are visible. The first eight orientations of the first block, are labeled from 1 to 8 clockwise. The orientations of the second block $B_{1,2}$ are labeled from 9 to 16. This labeling continues left-to-right, up-down until the eight orientations for all the sixteen blocks are assigned. They form the corresponding **kp**-descriptor of 128 coordinates.
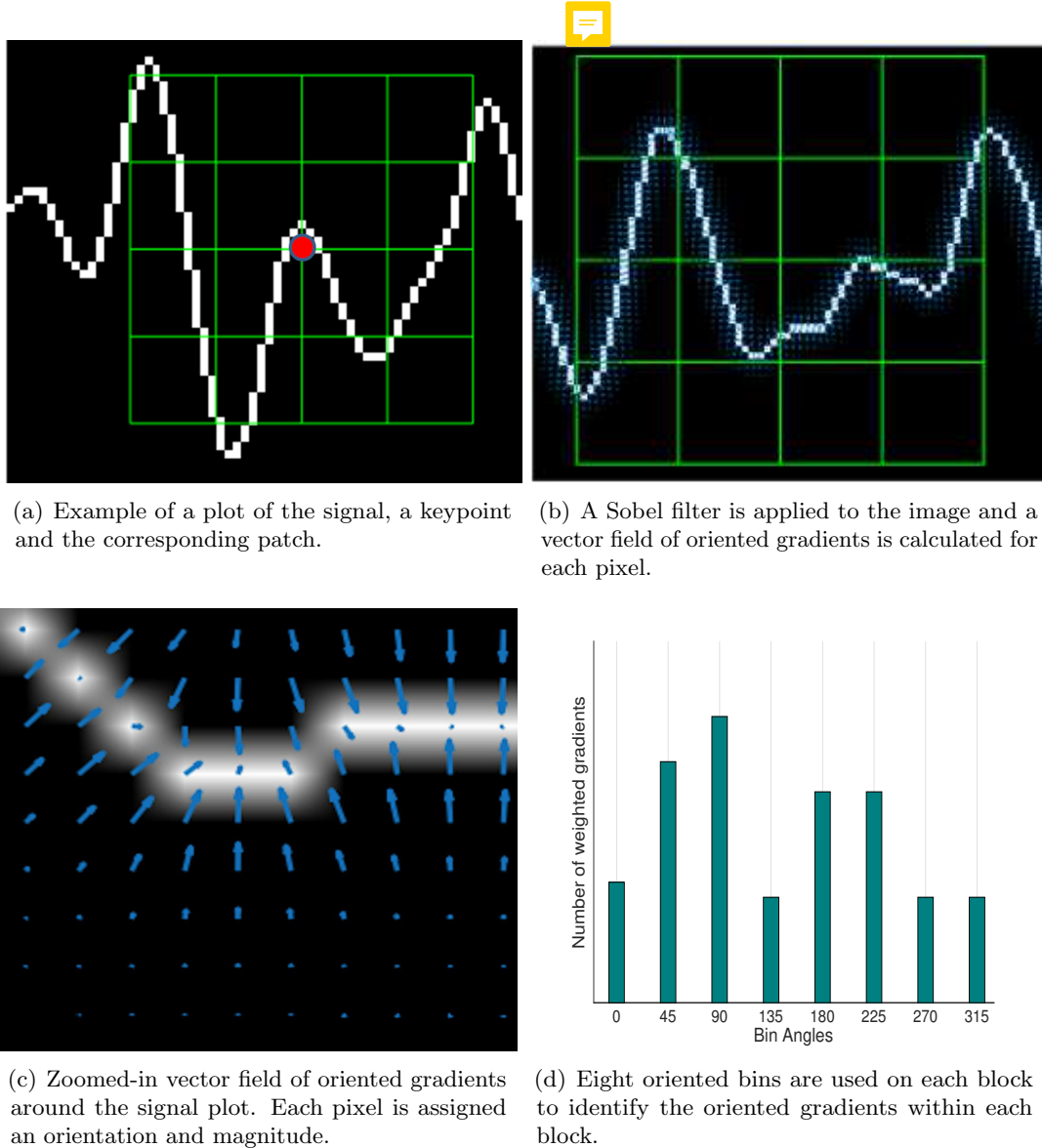
(a) Example of a plot of the signal, a keypoint and the corresponding patch.

(b) A Sobel filter is applied to the image and a vector field of oriented gradients is calculated for each pixel.

(c) Zoomed-in vector field of oriented gradients around the signal plot. Each pixel is assigned an orientation and magnitude.

(d) Eight oriented bins are used on each block to identify the oriented gradients within each block.

Figure 4.1: Patch and vector field of oriented gradients.
.

4.3 Keypoint Location

The keypoint **kp** location must be accurately specified in order to establish the region from the signal where the waveform is located.

For the horizontal position, the localization of the keypoint is based on a priori information, based on the characteristics of the event under study. For instance, ERPs have a specific timing that can be explored to elucidate in which position the expected signal pattern will be generated.
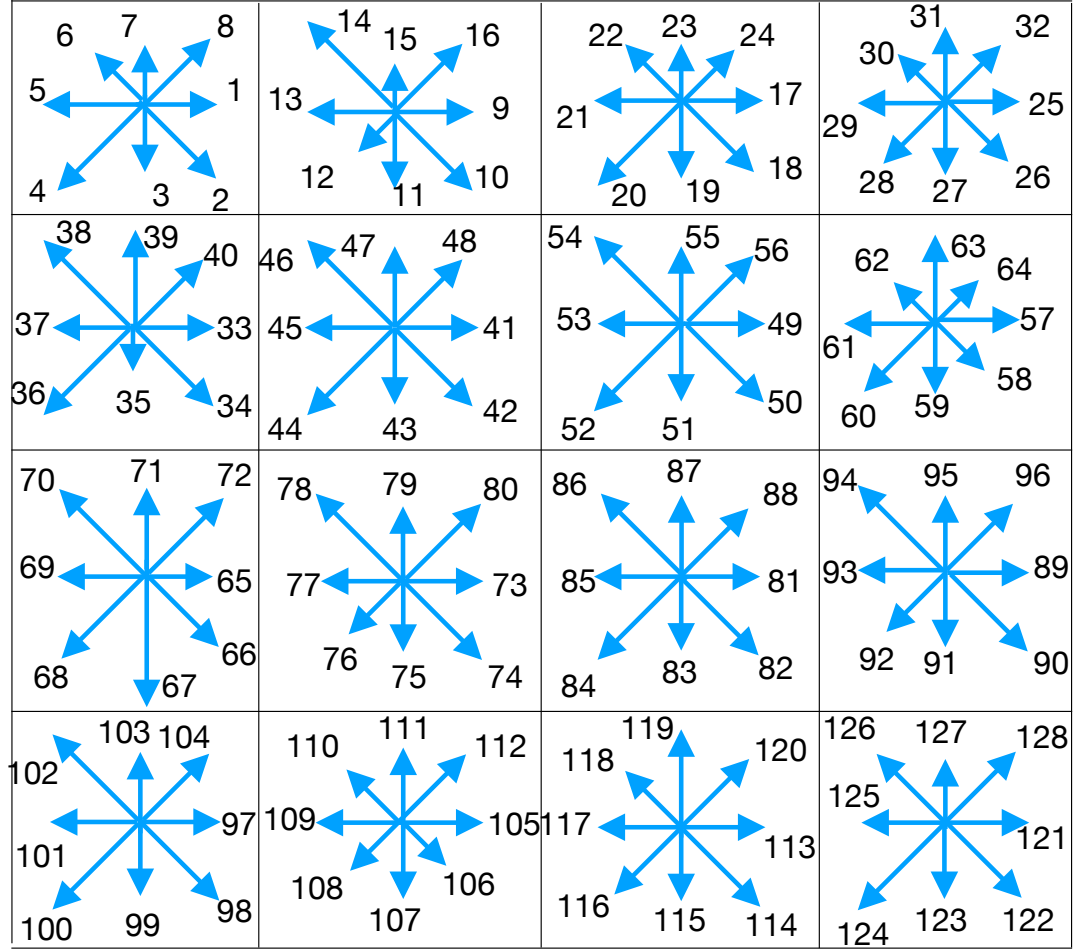
Figure 4.2: A scheme of the orientation's histogram computation. The first eight orientations of the first block, are labeled from 1 to 8 clockwise. The orientation of the second block $B_{1,2}$ is labeled from 9 to 16. This labeling continues left-to-right, up-down until the eight orientations for all the sixteen blocks are assigned. They form the corresponding descriptor of 128 coordinates. The length of each arrow represent the value of the histogram on each direction for each block.

Additionally, there can be more than just one keypoint and patch located over the signal plot. This is particular important for oscillatory processes where many waveforms are contained within the same signal segment. This needs to be addresses by defining a keypoint density parameter $\mathbf{kp_d}$.

Regarding the Vertical Location, there are two options. The first one is along the signal, exactly on the sample points calculated by the Equation 3.15. The second is on a fixed
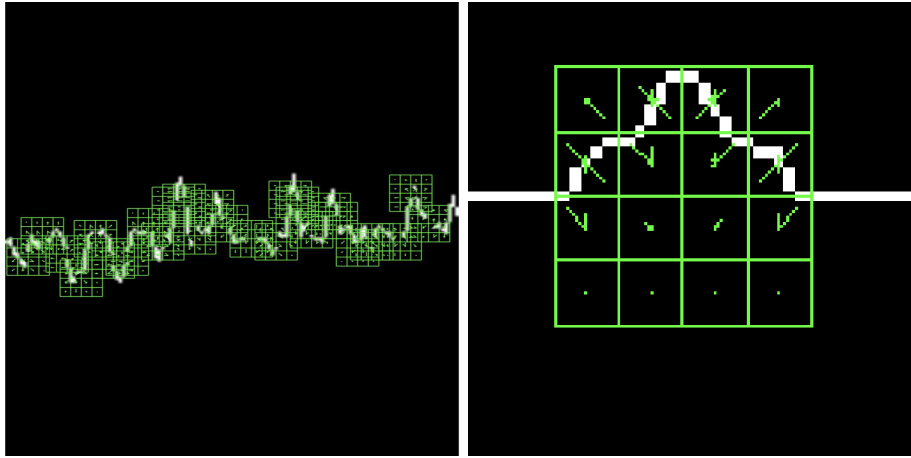
(a) Fixed size patches are located all along the EEG signal trace at a given keypoint density $\mathbf{kp_d}$.

(b) A patch is used to map an artificial signal using an autoscale plotting scheme, and mapping the entire waveform within the patch. The keypoint is located on the zero-level $z(c)$ value.
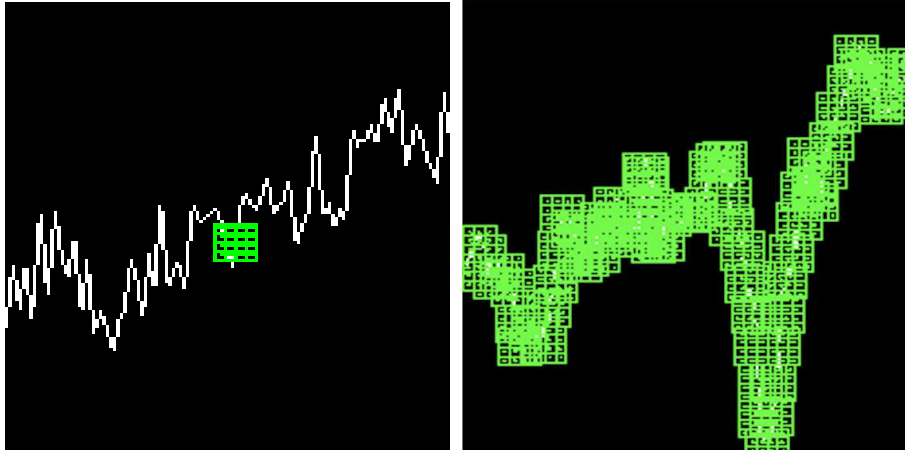
(c) Only one fixed size patches are used on a very specific position of the signal trace.

(d) An entire signal is mapped with fixed size patches with a very high keypoint density $\mathbf{kp_d}$.

Figure 4.3: Four different alternatives of keypoint location and patch geometry.

.

position over the zero-level as described by 3.11. Figure 4.3(a) show the former while on (b) the latter can be spotted.

4.4   Patch Geometry

The Horizontal Patch Scale S_t determines the size of the patch on the image horizontal axis, and it is related to the span λ of the waveform to analyze according to

$$S_t = \frac{\lambda \; F_s \; \gamma_t}{\Delta_s} \tag{4.5}$$

where F_s is the sample frequency, γ_t is the time scale factor and Δ_s is the unit length of the patch which determines the pixel conversion factor. This value depend   the actual implementation of the Histogram of Gradient Orientations of the SIFT method. In this case, its value is $\Delta_s = \sqrt{2} \, 3 \, 5$, where 3 is the fixed magnification factor, and 5 correspond to the number of blocks in which the patch is divided, plus half the size of the block on each direction. Check Appendix A for details on the SIFT method implementation.

On the other hand, on the vertical axis, the vertical patch scale depends on the peak-to-peak amplitude $\Delta\mu V$, and the amplitude scale factor γ, as

$$S_v = \frac{\Delta\mu V \; \gamma}{\Delta_s}. \tag{4.6}$$

The vertical scale can be dynamically adjusted according to the peak-to-peak amplitude of each segment, or it can be set fixed. This is more appropiate if the underlying signal is bounded which is the case if the standardized procedure described in 3.3 is applied.

Figure 4.4 shows the different parameters of the patch and how they are related to the underlying signal. Once these parameters are set, the size in pixels of the patch can be obtained in both dimensions. Hence, the horizontal patch size in pixels is

$$\mathbf{S}_x = \lfloor \Delta_s \; S_t \rfloor + 1 \tag{4.7}$$

and the vertical patch size in pixels can be calculated from

$$\mathbf{S}_y = \lfloor \Delta_s \; S_v \rfloor + 1 \tag{4.8}$$

where Δ_s being the unit length of th   ch. The parameters S_t and S_v are the horizontal and vertical patch scale. This region is arranged in a 4×4 grid and the pixel $\mathbf{kp}$ is the patch center. For instance, for a given set of values of $S_v = 1$ and $S_t = 1$, the patch is a squared region on the image of size 22 pixels.
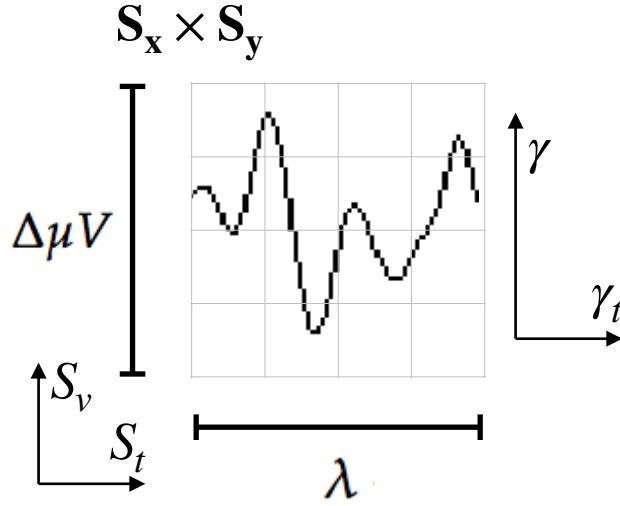
Figure 4.4: The scale of local patch is selected in order to capture the whole waveform, which can be scaled in the time γ_t and amplitude γ direction. This determines appropriate horizontal S_t and vertical S_v patch scales. The size of the patch is $\mathbf{S_y} \times \mathbf{S_x}$ pixels. The vertical size consists of 4 blocks of size $\mathbf{S_y}$ pixels which should be high enough as to contain the signal $\Delta\mu V$, the peak-to-peak amplitude of the signal component. The horizontal size includes 4 blocks, up to $\mathbf{S_x}$ pixels, and should cover the entire duration in seconds of the signal waveform, λ.

The patch size cannot be bigger than the image itself, whose width is W_x and its height is H_y . This is reflected by the following two inequalities that restrict the size of the patch according to

$$\frac{W_x - 1}{\Delta_s} \geq S_t, \tag{4.9}$$

on the horizontal axis, and on the vertical axis,

$$\frac{H_y - 1}{\Delta_s} \geq S_v. \tag{4.10}$$

4.4.1 Oscillatory Processes

For these patterns, the strategy is to locate keypoints, and their patches, all along the signal trace, filling the entire signal segments with all the possible patches. In this case, the keypoint density $\mathbf{kp_d}$ determines the step at which a keypoint is located along the trace of the signal, sample point after sample point. Care must be taken close to the margins,

where there should be a gap to avoid locating incomplete patches. This can be observed in Figure 4.3(a).

4.4.2 Transient Events

For transient events, descriptors are treated as representatives of the single transient waveforms. This lead to usually just one keypoint that is located in a meaningful position along the horizontal axis. Additionally, for autoscale plotting, the zero level can be used to localize keypoints.

4.5 BCI Algorithm

Now that all the ingredients has been exposed, the general layout of the BCI algorithm can be described. Going back to BCI model referenced in 2.1, the following sections describe the Preprocessing, Calibration and Classification step.

Terminology Clarification

Definition 4.5.1. Keypoint: *A keypoint is a specific pixel on the image. The keypoint is the center of a patch and it is used to outline a region of interest.*

Definition 4.5.2. Patch: *An image region centered around a keypoint. It is divided in a rectangular 4x4 grid.*

Definition 4.5.3. Descriptor: *A 128-dimensional feature vector. Contains the histogram of 8 angular directions per each block of the 16 blocks of the patch.*

Definition 4.5.4. Waveform: *A signal shape, a transient component or an oscillatory wave, with a potential cognitive implication.*

Definition 4.5.5. Signal *The EEG Signal. This Thesis refers to a multidimensional signal or to a single-channel signal. When the clarification is of relevance, it is provided.*

Definition 4.5.6. Image *The canvas with sample points created from equation 3.15.*

Definition 4.5.7. Plot *The trace of the EEG signal on the Image.*

Definition 4.5.8. HIST *This the acronym that hereafter is being used to describe the feature extraction method proposed in this Thesis.*

4.5.1 Preprocessing

Signal preprocessing can be applied to $x(n,c)$ before applying the plotting scheme. Preprocessing depends on the cognitive paradigm under study.

4.5.2 Calibration

The calibration step of BCI entails designing an experimental protocol or procedure to allow the computer to learn the signals that identify a cognitive pattern. Once those patterns are learned, the BCI system tries to identify new unlabeled signals comparing them against those learned patterns. By performing the feature extraction method outlined in this Chapter, descriptors can be extracted from images and stored in dictionary templates. These templates are generated on a channel-by-channel basis. Each dictionary T^{bpc} is used to represent one type of signal to discriminate. The data structure proposed to store descriptors is a KD-tree [77].

Moreover, the calibration can be also used to identify the bpc, the Best Performing Channel. This value is used to evaluate the spatial performance based on the electrode where the best classification accuracy is obtained [24].

4.5.3 Classification

A discriminative [142] semi-supervised classification method based on Naive Bayes Nearest Neighbor (NBNN) [13] is applied to classify EEG signals using the features provided by the calculated descriptors. One problem that frequently arises when using local features is how to go back from the classification of those local characteristics to the image where those descriptors come from. The NBNN technique overcomes this problem by comparing each image against a whole label class which is characterized by the set of descriptors q that are closest to each one of the descriptors of the query image $\mathbf{d}_i$. This algorithm is very easy to implement, and is based on the following Equation:

$$\hat{L} = \arg\min_L \sum_{\mathbf{d_i}^{(bpc)}} \sum_{q \in N_T(\mathbf{d_i}^{(bpc)})} \left\| q - \mathbf{d_i}^{(bpc)} \right\|^2 \qquad (4.11)$$

where $N_T(\mathbf{d}^{(bpc)})$ is defined as $N_T(\mathbf{d}^{(bpc)}) = \{\mathbf{d} \in T^{bpc}$ / and is the k-nearest neighbor of $\mathbf{d}^{(bpc)}\}$ for the best performing channel. Hence, the estimated class label $\hat{L}$ of a query image is calculated as the class label L that minimize the summation of the distance between each descriptor $\mathbf{d_i}^{(bpc)}$ that belongs to the query image and its corresponding near neighbor

distances and descriptors

$N_T(\mathbf{d}^{(bpc)})$ descriptor for each class. Figure 4.5 show a scheme of how this classification method works.
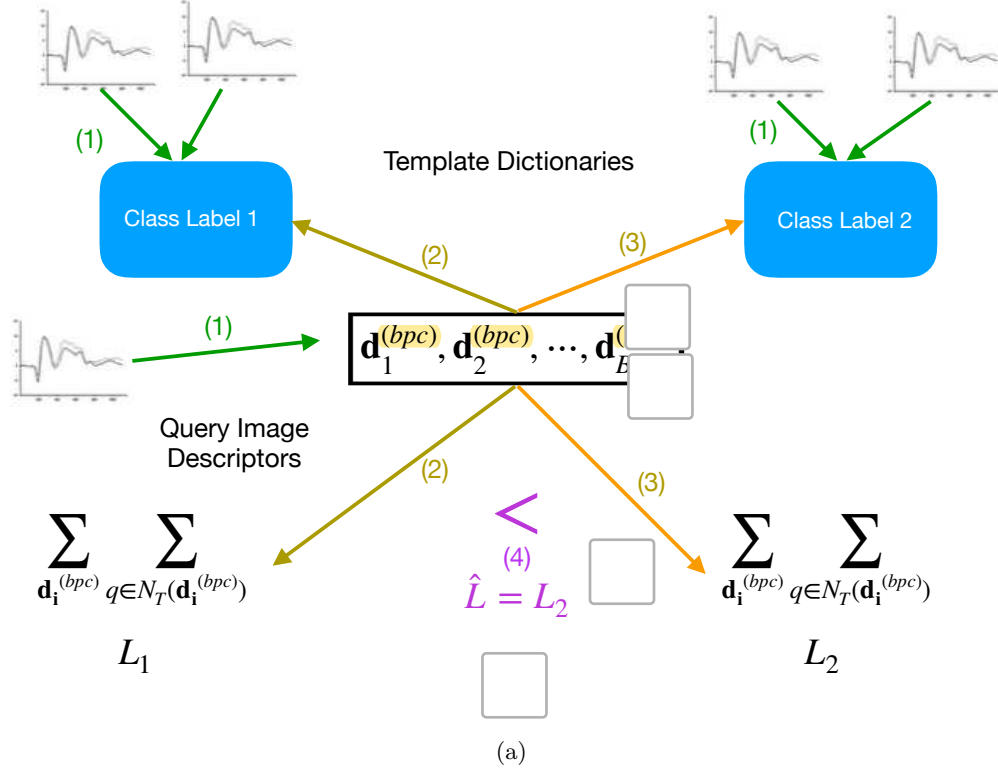

Figure 4.5: Cl___ication Algorithm: **Step 1** Two Dictionaries are created from templates descriptors ob___ed in a calibration session for two different classes, labeled 1 and 2. On the other hand, a set of query descriptors are extracted from a new image that needs to be categorized.**Step 2** Distances from every descriptor $\mathbf{d}_i$ are calculated a___st the closest one from the dictionary of class 1. Distances summations are accumulated.**Step 3** Distance values from every descriptor $\mathbf{d}_i$ are calculated against the closest one from the dictionary of class 2, the other class. Distances summations are accumulated. **Step 4** The two summation values for each class label are compared against each other. The summation that achieved the lesser value is the one that more closely resembles the set of templates, thus is the one predicted by the classification algorithm.

4.5.4   Algorithm

In brief, based on segmented signals from at least two labeled classes, a set of images is first generated. For each image, keypoints are localized and descriptors are extracted during the training or calibration step of a BCI procedure and they are grouped in template dictionaries for each one of the classes. Additionally, the spatial performance is evaluated and the *bpc*

value is computed.

Hence, given a new unlabeled signal segment, an image plot is generated as well, their keypoints localized and their descriptors extracted. They are fed into Equation 4.11 in order to determine the class which minimizes the summation and thus provide the information bit to the BCI controller.

4.A Model Summary

This section provides a mapping cheat-sheet to convert and obtain the parameters of the algorithm for a given set of signal parameters.

The input signal parameters are N,F_s, λ and the peak-to-peak amplitude $\Delta\mu V$ of the waveform to study. The unit length of the patch is $\Delta_s = \sqrt{2}\ 15$ and as has been earlier mentioned, depends on the particular SIFT implementation.

Output parameters are: γ γ_t H_y W_x S_t S_v $\mathbf{S}_y$ $\mathbf{S}_x$ $\mathbf{kp}$

Mappings:

Amplitude scale factor

$$\gamma = \frac{H_y}{\Delta\mu V}$$

Time scale factor

$$\gamma_t = \frac{W_x}{F_s\ w}$$

Restriction on the waveform time scale

$$\frac{W_x - 1}{\sqrt{2}\ 15} \geq S_t$$

Restriction on the waveform amplitude scale

$$\frac{H_y - 1}{\sqrt{2}\ 15} \geq S_v$$

Horizontal Patch scale

$$S_t = \frac{\lambda \; F_s \; \gamma_t}{\Delta_s}$$

Vertical Patch scale

$$S_v = \frac{\Delta \mu V \; \gamma}{\Delta_s}$$

Time to sample point conversion

$$n = \lfloor F_s \; \Delta_t \rfloor \; \gamma_t$$

Horizontal Patch size in pixels

$$\mathbf{S}_x = \lfloor \Delta_s \; S_t \rfloor + 1$$

Vertical Patch Size in pixels

$$\mathbf{S}_y = \lfloor \Delta_s \; S_v \rfloor + 1$$

Span of a Patch

$$\Delta_t = \frac{S_t \; \Delta_s}{F_s \; \gamma_t}$$

Pixel Resolution on the horizontal axis

$$1P_x \equiv \frac{1}{F_s \; \gamma_t}[\text{s}]$$

Pixel Resolution on the vertical axis

$$1P_y \equiv \frac{1}{\gamma}[\mu V]$$

Chapter 5

Alpha Wave: spotting wiggles

> The electroencephalogram represents a continuous curve with continuous oscillations in which... one can distinguish larger first order waves with an average duration of 90 milliseconds...
>
> ———————————————
>
> Berger

This Chapter describes the experiments performed over a well-established but still mysterious EEG cognitive signal: The Berger Rhythm or Visual Occipital Alpha Waves. An own dataset of resting subjects with and without alpha blocking is produced and the details of their generation are outlined. Additionally, an experiment on a public dataset is also delineated. Conclusions and discussion are described in the last section.

5.1 Introduction

Alpha Waves were the first signals ever spotted from the Electroencephalography. They are regularly characterized as 10Hz, or more broadly between the frequency band of 8-12 Hz. They are physiologically consistent across subjects, though it has been reported inter- and intra- variations with functional cognitive implications [51]. Moreover, they are associated with synchronous inhibitory processes and attention shifting [109]. They tend to be more prominent while subject's eyes are closed and appear stronger in occipital regions, around O_1 and O_2 [142, 122]. These waves are also called Prominent Posterior Alpha or Posterior Dominant Rhythm due to their pervasiveness in EEG [113, 51].

Figure 5.1 shows two records of 8-channels signals. Figure 5.1(a) contains the registered alpha waves of a subject with their eyes open while the (b) contains the same information

with their eyes closed. The characteristic pattern of wiggles can be spotted in the latter, while their absence entails the blocking of alpha waves in the former [9].
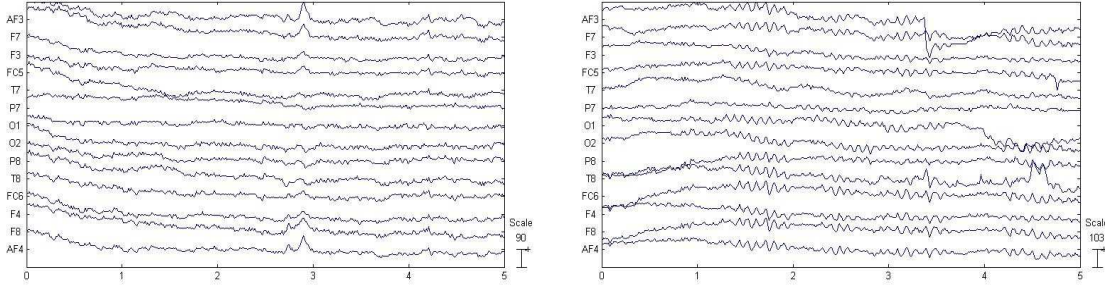
(a) EEG signals of a relaxed healthy subject with their eyes open.

(b) EEG signals of the same relaxed subject with their eyes closed. Alpha Waves wiggles can be spotted since the first second.

Figure 5.1: Five seconds of EEG signals obtained from the Emotiv EPOC device. Fourteen channels are shown on the vertical axis, while x-axis shows time in seconds.

This important rhythm is an oscillatory process. As such, it is understood and studied in the frequency-domain. Figure 5.2 shows the results of applying the Fast Fourier Transform to two different segments of 10s length. For each segment, the power spectral density is calculated and their values are shown on the vertical axis. Frequency values are shown on the horizontal axis. On Figure (a) no particular frequency component can be spotted. However, on Figure (b) the prevalence of the 10-Hz alpha wave component can be observed.

5.2   Materials and Methods

These experiments consist in performing a binary classification of EEG signal segments between the two defined classes. Class 2 is assigned to segments containing significant alpha waves (i.e. eyes closed), whereas class 1 identifies those where these signals are blocked (i.e. eyes open).

5.2.1   Dataset I - Emotiv EPOC alpha waves own dataset

The first dataset is gathered using the EEG EPOC Emotiv Headset. Although this is a commercial-grade device, it provides an acceptable price-performance ratio and it has been used to investigate basic EEG processes [35, 34]. In order to obtain the multichannel raw EEG signal, a C++ SDK library provided by the manufacturer is used and an in-house

(a) Subject was sit, relaxed in front of the Computer Monitor with his eyes open.

(b) Subject was sit, relaxed in front of the Computer Monitor with his eyes closed. A strong 10Hz component can be observed.
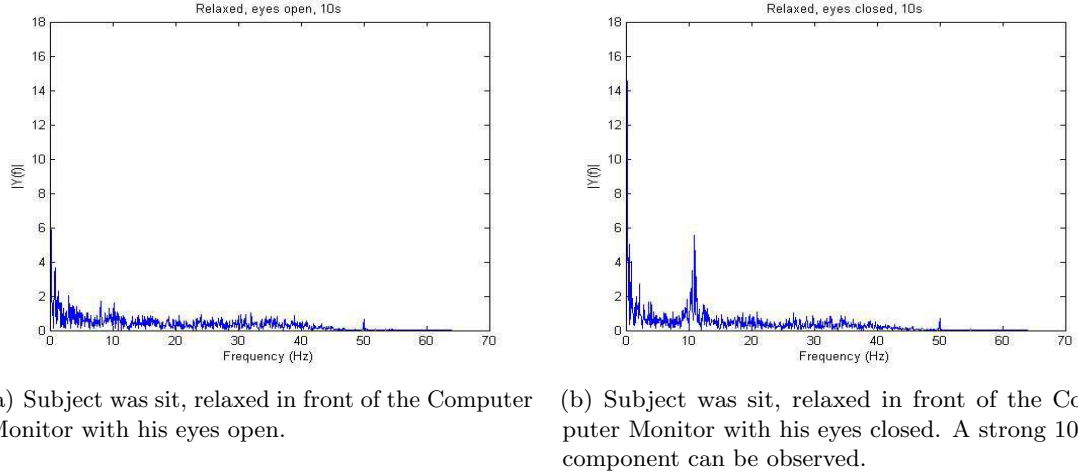
Figure 5.2: Spectrum components of a 10s signal segment of a subject with their eyes open. Horizontal axis shows different frequencies while the vertical axis represents the power spectral density. In both diagrams a 50Hz line component can be visualized.

software program is developed. This device has 14 channels, and a sampling rate of 128 Hz[122]. Available channels are AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4. Ten healthy subjects between ages 20-50 are recruited and they accept to wear the device and to participate in the experiments.

A 30 minutes procedure is required to adjust the headset to each user, in order to decrease the impedance on each electrode (bellow 5kΩ). This is achieved by physically adjusting the headset position over the scalp, and by embedding each electrode electrode pad in saline solution. A software program developed by the manufacturer is used to obtain the measured impedance. Once the set up is finished, each subject is instructed to sit in a relaxed position. Subsequently, she/he is commanded to watch the screen for 15 seconds, trying to avoid, as much as possible, to abruptly move its body or head. During that time, a single-trial of 10 seconds-length window of EEG signals data is transferred to a PC and logged into standard binary files. After a 5 minutes pause, the subject is asked to close the eyes avoiding any movement while keeping the same pose for another batch of 15 seconds. Again, 10 seconds of EEG information are transferred and logged into the PC. This finally produces a dataset of 10 subjects, 2 classes per subject, composed of 14 channels, 10-seconds length or 1280 samples per window. These windows are further divided into 10 segments per class and per subject.
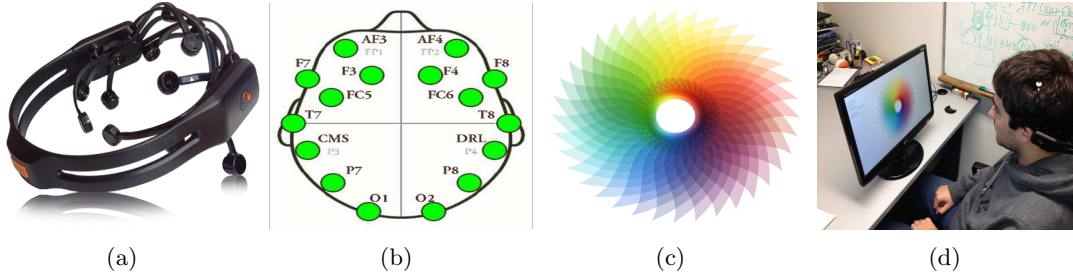
(a)   (b)   (c)   (d)

Figure 5.3: (a)EPOC Emotiv consumer-grade 14-channels wireless EEG.(b) This device has a fixed set of positions according to the 10-20 International System.(c) While resting and sitting confortable in a chair, subjects had to fixate their sight to the center of this image which was being displayed on a computer monitor, 1 meter away from the subject.(d) Subject performing the experiment to produce the dataset described in 5.2.1

.

5.2.2 Dataset II - AlphaNet Dataset

Additionally, the performance of HIST was tested against the public dataset EEG Motor Movement/Imagery Dataset of the PhysioNet effort published and mantained by the U.S. National Health Service [111, 44].

Baseline records and motor/imagery tasks were performed by 109 healthy volunteers, using the BCI2000 system [111] at a sampling frequency $F_s = 160$. At the same time 64-channel EEG records were registered where each subject completed 14 tasks, called Runs. The first two are the baseline calibration tasks, of relevance for this Chapter. These Runs 1 and 2 are each one-minute records of resting subjects with eyes open and eyes closed. From these records, 60 segments of 1-second length are further extracted per subject. Class labeling is the same as Dataset I. The experiment was performed on 25 of the 109 subjects.

5.2.3 Parameters

Images and plots are generated according to Section 3.7, with an autoscale plotting scheme. Keypoint localization is determined according to Section 4.3 by following the trace of the signal, with a $\mathbf{kp_d} = 1$. Descriptors are extracted from all the generated images, from both classes, and they are used to classify images in a 10-fold cross-validation procedure. The classification method described in Section 4.5.3 is used to perform a binary classification. The parameter γ is set to 2, as well as γ_t.

5.3   Results

Dataset I was controlled and verified by processing it with a 8-12Hz band-pass filter, and calculating the average power spectral density across subjects for all channels. It can be observed on Figure 5.4(a) that values corresponding to class 2 (eyes closed) are always higher than the values obtained for class 1 (eyes open). On the other hand, for the sake of illustration, a scatter plot of the obtained segment's power spectral density for O1 vs O2 for Subject 2 is shown, where a separation of classes can be devised. This proves that there is discriminative information in the frequency-domain.

(a) Power spectral density averaged across 10 subjects for each channel. Values for Class 2 (red) are always higher than values for Class 1 (blue).

(b) Scatterplot of power spectral density for the channel O1 vs O2. There is a separation of classes between the red (class 2) and blue(Class 1) dots.
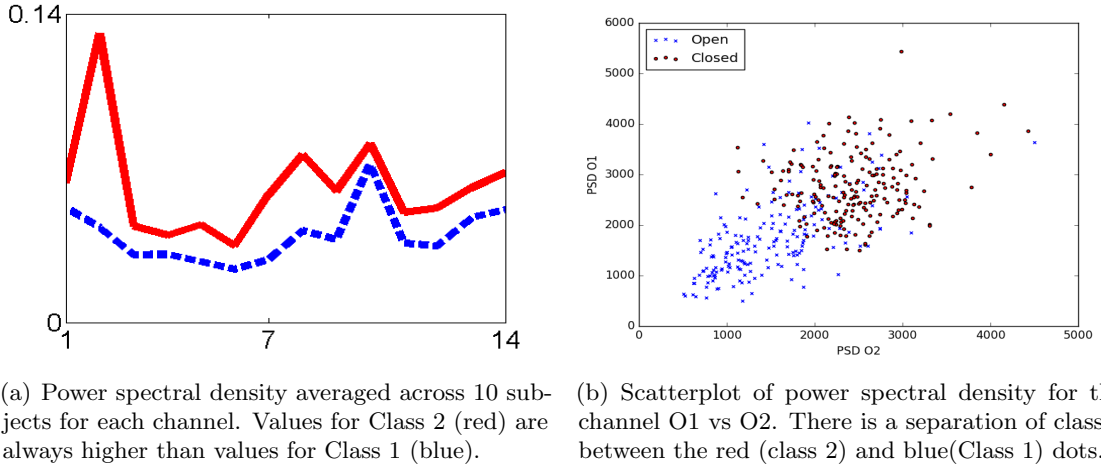
Figure 5.4: Frequency analysis of Dataset I

Results of applying the 10-fold cross validation procedure on the entire set of labeled descriptors is shown on Figure 5.5. Template Dictionaries for Class 1 and Class 2 are formed using training descriptors for all the subjects at the same time. Hence, the testing step of the cross-validation procedure is implemented subject by subject, but performing a transfer learning between them. The classification is subject-independent, and an average above 70% is obtained in occipital channels.

For the Dataset II, training and testing steps of the cross-validation procedure are implemented subject by subject. An accuracy median higher than 70% for 25 subjects, also on occipital channels O1, Oz, O2 and Iz (numbered 61 to 64) is obtained while discriminating Runs 1 and 2 (Baseline eyes open vs Baseline eyes closed). This information can be devised on Figure 5.6(a) where boxplots of the averaged classification accuracies for all the subjects are represented. On the other hand, Figure 5.6(b) shows the 10-fold validated accuracy for
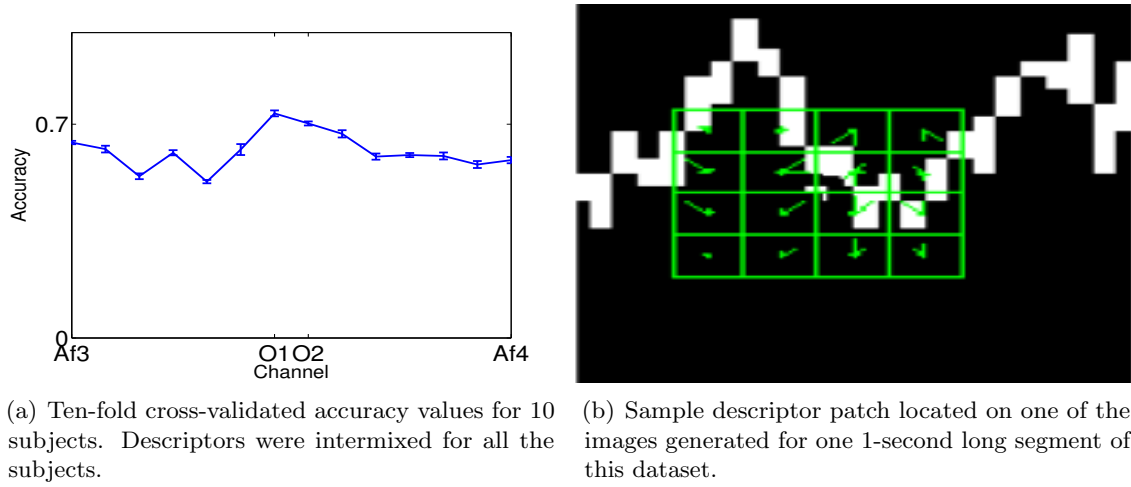
(a) Ten-fold cross-validated accuracy values for 10 subjects. Descriptors were intermixed for all the subjects.

(b) Sample descriptor patch located on one of the images generated for one 1-second long segment of this dataset.

Figure 5.5: The classification accuracy is maximum on occipital channels O1 and O2. The horizontal patch scale S_t and the vertical patch scale S_v are set to 1, whereas γ and γ_t are set to 2, which corresponds to a variation of $\Delta\mu V = 12$ microvolts in the signal amplitude during $\lambda = 0.09$ seconds.

(a) Ten-fold cross-validated accuracies for O1, Oz, O2 and Iz channels for 25 subjects of the Alphanet Dataset. Medians are above 75%.

(b) Ten-fold cross-validated accuracies values for subject number 12, using Runs 1 and 2 of the Alphanet Dataset.
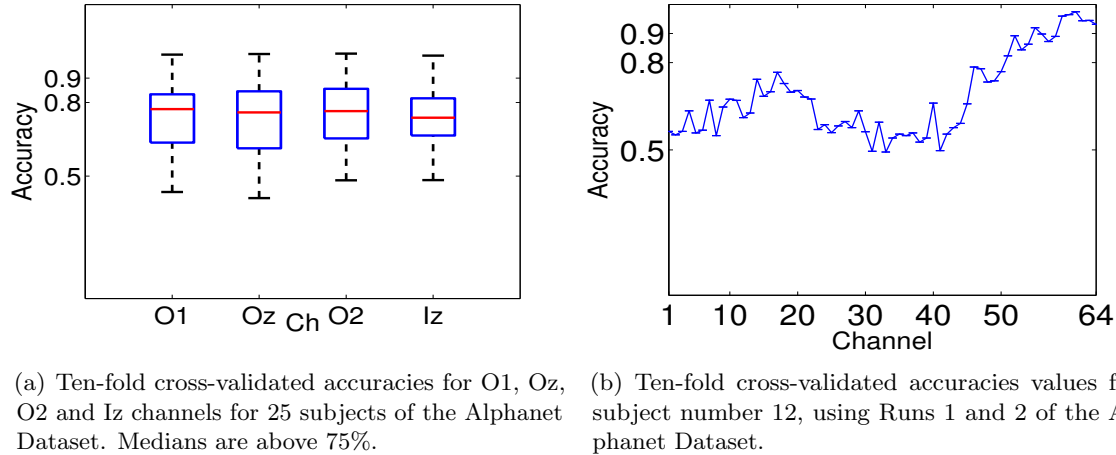
Figure 5.6: Classification Accuracy for segments of 1s ($N = 160$) of EEG, between Class 1 and Class 2.

one random subject. A higher accuracy in the classification of the signals can also be seen over occipital channels.

5.4   Conclusion

It is known and it was verified here that the discriminative information in EEG alpha waves is mostly contained in the frequency-domain. In spite of this, there is enough information

encoded in the alpha waves wiggles to classify signal segments solely on the *features* captured by the HIST method, proposed in this Thesis.

It was also verified that the presence of oscillatory alpha waves is higher around occipital regions and that an automated procedure which analyze visually the image structure can detect them. This important oscillatory rhythm has many connections with shifting of attention and with volitional changes and is of quite relevance in BCI research. Particularly, the BCI paradigm of Visual Spatial Covert Attention is a further area of research for this method due to the fact that it is entirely based on analyzing alpha waves. Moreover, the posterior rythm has many implications outside the field of BCI and is very important to assess healthy EEG patterns. The exact meaning of alpha waves is still debated[1] and this basic procedure can open the possibility to explore it under a different perspective and verify if they can be effectively tied to some unexpected form of volitional control which may be effective for BCI understanding and improvement.

Chapter 6

Motor Imagery: the hunt for a greek letter

> ...utilizing the brain signals in
> man-computer dialogue.
>
> —————————————
>
> Vidal

6.1 Introduction

Motor Imagery is an EEG or ECoG based BCI paradigm originated on changes of SMR, sensorimotor rhythms, that are altered when a person engages in motor behavior, but it can also be elicited when a person imagines to perform any movement. Particularly, the Rolandic wicket rhythm, the μ rhythm, is of the same frequency (e.g. 8-12 Hz) of visual occipital alpha waves, but from a spatially different location (posterior frontal and anterior parietal areas)[142]. Although SMR patterns presents a high inter- and intra-subjects variability regarding the signal features required to identify them, an Event Related Desynchronization/Synchronization of μ rhythm is in general consistent across subjects, regardless of the specificity of the imagined movement (i.e. what is being imagined to move).

6.2 Materials and Methods

In order to verify if this ERD/ERS could be detected by the method presented in this Thesis, i.e. by automatically extracting the information from the signal plots, a BCI Simulation is performed against the public Motor Imagery Dataset 002-2014, published by BNCI-Horizon 2020 website and initiative [120]. This dataset is composed of 8 runs for 14 participants,

aged between 30 and 30 years, five females, with a sampling frequency of $F_s = 512$. Nine out of these 14 participant reported never been exposed to a BCI device. One session per participant is recorded on a single day and one session consisted of eight consecutive runs with short pauses between them. The first 5 runs are used for training without feedback, and the remaining 3 runs are used to test the results. The original online experiment was performed with 20 trials on each run, 10 corresponding to imagining moving the right hand and the other 10 to feet movement. Figure 6.1 schematize the protocol and the structure of this published dataset. This BCI simulation experiment is divided in two. In the first simulation, baseline signals, corresponding to the 1st second of each trial are compared against right hand motor imagery, which is 4.5 seconds ahead of the beginning of each trial. Signal segments of 1-second length are processed for 10 trials for each of 5 runs and their descriptors are extracted for both classes. The second BCI simulation is similar but only extracting trials corresponding to feet movement imagery.

BCI Simulation or Cross Validation?

The task of decoding information from brain signals inherits practices from Machine Learning (ML). Cross-validation is used in ML to reduce overfitting bias and to increase the independence on the dataset that is used as calibration (see Section 4.5.2). However, the brain data used in BCI are extracted from a person who is performing a task and whose signals are changing while trying to adapt to this operation. Hence, mixing the dataset, shuffling the sessions and trials is at least a challenging assumption. BCI Simulation, on the other hand, is not very well defined in BCI research, but their practice, without naming it, has been the regular approach for BCI Competitions. It consists in reproducing the operational sequence that was utilized to generate the dataset. Hence, the experiment is replicated offline using the training information to train or calibrate a classifier, and to classify the testing signals as if they were generated at that same moment.

Regardless, of any definition, the online validation with feedback of any BCI system is the unquestioned gold standard of the discipline [142].

6.3 Results

Binary classification accuracies are calculated based on the output of the BCI simulation on the remaining 3 runs for each participant, in a single-trial approach: for each sampled segment of 1-second length, classification based on the classification algorithm described in 4.5.3 is applied and a match or mismatch is obtained. Results are shown in Figure 6.2
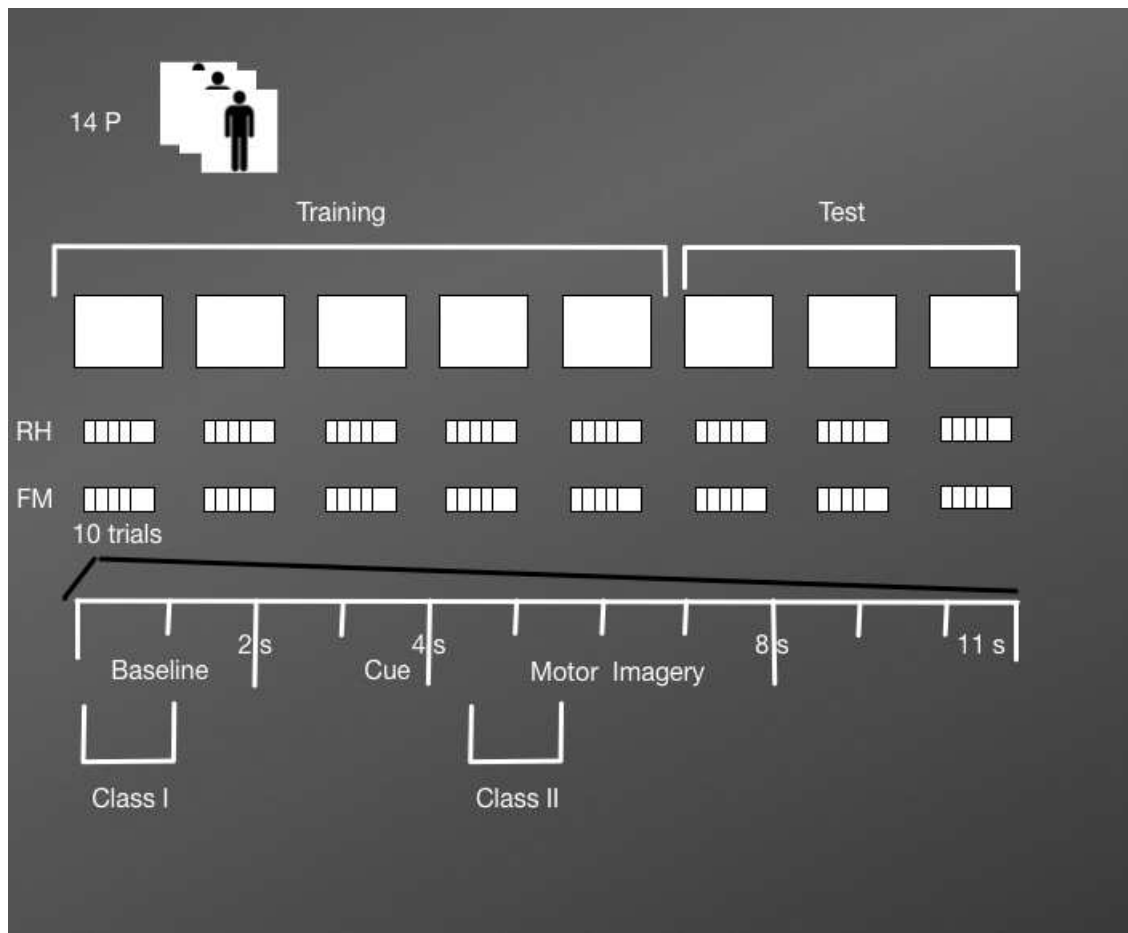
Figure 6.1: Fourteen voluntary participants performed 5 sessions of training and 3 sessions of testing. On each session each subject had 10 trials to perform Right Hand Motor Imagery (RH) and 10 trials for Feet Movement (FM). At the same time, each trial has a 2-seconds baseline and a 4-seconds section to perform the imagery task. For each BCI Simulation, class 1 is defined from the EEG segments obtained from the baseline section, while class 2 is based on extracting segments from the imagery section of the EEG signal.

where for right-hand detection RH, average accuracies of around 70% are obtained for the channel C3, the best-performing channel bpc, coincidentally with the contralateral structure of the imagined movement. On the other hand, the binary classification accuracy for feet imagery detection FM, achieves in all the channels accuracies of just above chance level.
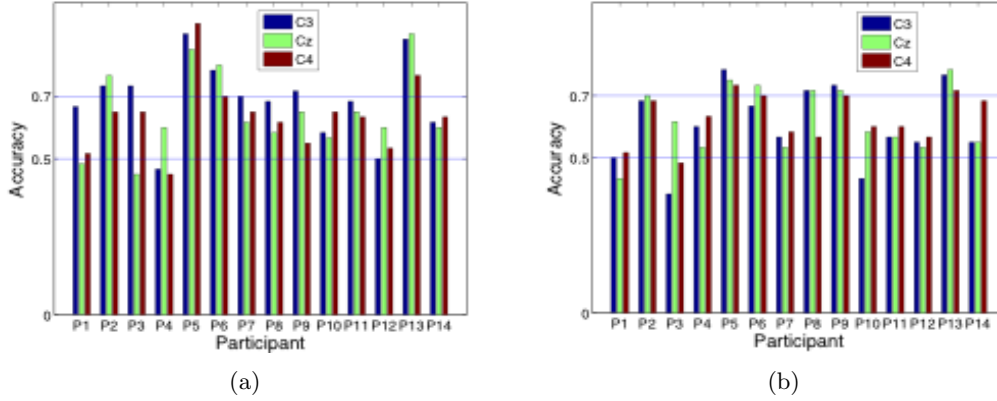

Figure 6.2: Classification Accuracy for discriminating segments of 1s ($N = 512$) of EEG for Motor Imagery detection BCI simulation. (a) Accuracy values for channels C3, Cz and C4 for the 14 participants of the described MI dataset discriminating between baseline and right-hand imagery. (b)The same procedure for feet imagery. Accuracy levels averaged to 70% are obtained only for right-hand movement on the contralateral channel C3. The horizontal S_t and vertical S_v patch scale are adjusted to 6

.

6.4 Conclusion

Offline BCI Simulation of single trial asynchronous triggering for right hand MI based on signal plots is implemented with a level of success of 70% for 7 out of 14 Participants. Single trial asynchronous triggering of BCI can be implemented with this paradigm, particularly for right-hand motor imagery. The name μ rhythm was precisely coined because the shape of the waves have some resemblance to the greek letter [29]. Additionally, in line with previous chapter results, though the differentiation information is contained in the frequency domain, the method based on the Histogram of Gradient Orientation detected differences in the shape of the signals. Coincidentally with results obtained from Alpha Waves, there is information that is mapped in the structure of the waveform, at least for frequencies on the 10 Hz range, which characterize both types of waves.