

# Histogram of Gradient Orientations of Signal Plots applied to P300 Detection

Rodrigo Ramele<sup>1,\*</sup>, Ana Julia Villar<sup>1</sup> and Juan Miguel Santos<sup>1</sup>

<sup>1</sup>*Centro de Inteligencia Computacional, Computer Engineering Department,  
Instituto Tecnológico de Buenos Aires, Buenos Aires, Argentina*

Correspondence\*:

Rodrigo Ramele, C1437FBH Lavarden 315, Ciudad Autónoma de Buenos Aires,  
Argentina  
rramele@itba.edu.ar

## 2 ABSTRACT

3 Word Count: ~~4841~~ 4952

4 The analysis of Electroencephalographic (EEG) signals is of ulterior importance for decoding  
5 patterns that could improve the implementation of Brain Computer Interfaces (BCI). These  
6 systems are meant to provide alternative pathways to transmit volitional information which could  
7 potentially enhance the quality of life of patients affected by neurodegenerative disorders and  
8 other mental illness. Of particular interests are those which are based on the recognition of  
9 Event-Related Potentials (ERP) because they can be elicited by external stimuli and used to  
10 implement spellers, to control external devices or even avatars in virtual reality environments.  
11 This work mimics what electroencephalographers have been doing clinically, visually inspecting  
12 and categorizing phenomena within the EEG by the extraction of features from images of signal  
13 plots. It also aims to provide a framework to analyze, characterize and classify EEG signals, with  
14 a focus on the P300, an ERP elicited by the oddball paradigm of rare events. The validity of the  
15 method is shown by offline processing a public dataset of Amyotrophic Lateral Sclerosis (ALS)  
16 patients and an own dataset of healthy subjects.

17 **Keywords:** **electroencephalography, histogram of gradient orientations, brain-computer interfaces, P300, SIFT, amyotrophic lateral**  
18 **sclerosis, naive-bayes near neighbours, waveforms**

## 1 INTRODUCTION

19 Although recent advances in neuroimaging techniques, particularly radio-nuclear and radiological  
20 scanning methods (?), have diminished the prospects of the traditional Electroencephalography (EEG),  
21 the advent and development of digitized devices has impelled for a revamping of this hundred years old  
22 technology. Their versatility, ease of use, temporal resolution, ease of development and production, and  
23 its proliferation as consumer devices, are pushing EEG to become the de-facto non invasive portable or  
24 ambulatory method to access and harness brain information (?).

25 A key contribution to this expansion has been the field of Brain Computer Interfaces (BCI) (?) which is  
26 the pursuit of the development of a new channel of communication particularly aimed to persons affected  
27 by neurodegenerative diseases.

28 One noteworthy aspect of this novel communication channel is the ability to transmit information from  
29 the Central Nervous System (CNS) to a computer device and from there use that information to control a  
30 wheelchair (?), as input to a speller application (?), in a Virtual Reality environment (?) or as aiding tool  
31 in a rehabilitation procedure (?). The holly grail of BCI is to implement a new complete and alternative  
32 pathway to restore lost locomotion (?).

33 EEG signals are remarkably complex and have been characterized as a multichannel non-stationary  
34 stochastic process. Additionally, they have high variability between different subjects and even between  
35 different moments for the same subject, requiring adaptive and co-adaptive calibration and learning  
36 procedures (?). Hence, this imposes an outstanding challenge that is necessary to overcome in order to  
37 extract information from raw EEG signals.

38 BCI has gained mainstream public awareness with worldwide challenge competitions like Cybathlon (??)  
39 and even been broadcasted during the inauguration ceremony of the 2014 Soccer World Cup. New  
40 developments have overcome the out-of-the-lab high-bar and they are starting to be used in real world  
41 environments (??). However, they still lack the necessary robustness, and its performance is well behind  
42 any other method of human computer interaction, including any kind of detection of residual muscular  
43 movement (?).

44 A few works have explored the idea of exploiting the signal waveform to analyze the EEG signal.  
45 In (?) an approach based on Slope Horizontal Chain Code is presented, whereas in (?) a similar  
46 procedure was implemented based on Mathematical Morphological Analysis. The seminal work of Bandt-  
47 Pompe Permutation Entropy (?) also explores succinctly this idea as a basis to establish the time series  
48 ordinal patterns. In the article (?), the authors introduce a method for classification of rhythmic EEG  
49 events like Visual Occipital Alpha Waves and Motor Imagery Rolandic Central  $\mu$  Rhythms using the  
50 histogram of gradient orientations of signal plots. Inspired in that work, we propose a novel application  
51 of the developed method to classify and describe transient events, particularly the P300 Event Related  
52 Potential. The proposed approach is based on the waveform analysis of the shape of the EEG signal, but  
53 using histogram of gradient orientations. The method is built by mimicking what ~~traditionally~~ regularly  
54 electroencephalographers have been performing for almost a century as it is described in (?): visually  
55 inspecting raw signal plots.

56 This paper reports a method to, (1) describe a procedure to capture the shape of a waveform of an ERP  
57 component, the P300, using histograms of gradient orientations extracted from images of signal plots, and  
58 (2) outline the way in which this procedure can be used to implement an offline P300-based BCI Speller  
59 application. Its validity is verified by offline processing two datasets, one of data from ALS patients and  
60 another one from data of healthy subjects.

61 This article unfolds as follows: Section 2.1 is dedicated to explain the Feature Extraction method based  
62 on Histogram of Gradient Orientations of the Signal Plot: Section 2.1.1 shows the preprocessing pipeline,  
63 Section 2.1.2 describes the image generation of the signal plot, Section 2.1.3 presents the feature extraction  
64 procedure while Section 2.1.4 introduces the Speller Matrix Letter Identification procedure. In Section 2.2,  
65 the experimental protocol is expounded. Section 3 shows the results of applying the proposed technique. In  
66 the final Section 4 we expose our remarks, conclusions and future work.

## 2 MATERIALS AND METHODS

67 The P300 (??) is a positive deflection of the EEG signal which occurs around 300 ms after the onset of a  
68 rare and deviant stimulus that the subject is expected to attend. It is produced under the oddball paradigm (?)

and it is consistent across different subjects. It has a lower amplitude ( $\pm 5\mu V$ ) compared to basal EEG activity, reaching a Signal to Noise Ratio (SNR) of around  $-15$  db estimated based on the amplitude of the P300 response signal divided by the standard deviation of the background EEG activity (?). This signal can be used to implement a speller application by means of a Speller Matrix (?). Fig. This matrix is composed of 6 rows and 6 columns of numbers and letters. The subject can focus on one character of the matrix. Figure 1 shows an example of the Speller Matrix used in the OpenVibe open source software (?), where the flashes of rows and columns provide the deviant stimulus required to elicit this physiological response. Each time a row or a column that contains the desired letter flashes, the corresponding synchronized EEG signal should also contain the P300 signature and by detecting it, the selected letter can be identified.

## 2.1 Feature Extraction from Signal Plots

In this section, the signal preprocessing, the method for generating images from signal plots, the feature extraction procedure and the Speller Matrix identification are described. Figure 2 shows a scheme of the entire process.

### 2.1.1 Preprocessing Pipeline

The data obtained by the capturing device is digitalized and a multichannel EEG signal is constructed, where rows are sample points and columns are channels (electrodes).

The 6 rows and 6 columns of the Speller Matrix are intensified providing the visual stimulus. The number of a row or column is a location. A sequence of twelve randomly permuted locations  $l$  conform an intensification sequence. The whole set of twelve intensifications is repeated  $k_a$  times.

- **Signal Enhancement:** The preprocessing This stage consists of the enhancement of the SNR of the P300 pattern above the level of basal EEG. The pipeline starts by applying a notch filter to the raw digital signal, a 4th degree 10 Hz lowpass Butterworth filter and finally a decimation with a Finite Impulse Response (FIR) filter of order 30 from the original sampling frequency down to 16 Hz (?).
- **Artifact Removal:** The multichannel EEG signal is processed on a channel by channel basis. For every complete sequence of 12 intensification intensifications of 6 rows and 6 columns, a basic artifact elimination procedure is implemented by removing the entire sequence when any signal deviates above/below  $\pm 70\mu V$ .
- **Segmentation:** For each of the 12 intensifications ,a window of one intensification sequence, a segment  $S_i^l$  of a window of  $t_{max}$  seconds of  $t_{max} = 1$  second of the multichannel signal is extracted, starting from the stimulus onset, corresponding to each row/column intensification  $l$  and to the intensification sequence  $i$ . As intensifications are permuted in a random order, the segments are rearranged corresponding to row flickering, labeled 1-6, whereas those corresponding to column flickering are labeled 7-12. Two of these segments should contain the P300 ERP signature time-locked to the flashing stimulus, one for the row, and one for the column.
- **Signal Averaging:** The P300 ERP is deeply buried under background basal EEG so the traditional standard approach to identify it is by point-to-point averaging the time-locked stacked signal segments. Hence the values which are not related to, and not time-locked to the onset of the stimulus are canceled out (?).

This last step determines the operation of any P300 Speller. In order to obtain an improved signal in terms of its SNR, repetitions of the sequence of row/column intensification are necessary. And, at the same time, as long as more repetitions are needed, the ability to transfer information faster is diminished, so there is a trade-off that must be acutely determined.

111      2.1.2 Ensemble Average

112      The procedure to obtain the point-to-point averaged signal goes as follows:

- 113      1. Highlight randomly the rows and columns from the matrix. There is one row and one column that  
114      should match the letter selected by the subject.
- 115      2. Repeat step 1  $k_a$  times, obtaining the ~~single trial segments~~  $S_1(n, c), \dots, S_{k_a}(n, c)$   $1 \leq l \leq 12$   
116      ~~segments~~  $S_1^l(n, c), \dots, S_{k_a}^l(n, c)$ , of the EEG signal where the variables  $n \in \{1, \dots, n_{max}\}$  and  
117       $c \in \{1, 2, \dots, Ch\}$   $1 \leq n \leq n_{max}$  and  $1 \leq c \leq C$  correspond to sample points and channel,  
118      respectively. The parameter ~~Ch~~  $C$  is the number of available EEG channels whereas  
119       $n_{max} = F_s \cdot t_{max}$   $n_{max} = F_s \cdot t_{max}$  is the segment length and  $F_s$  is the sampling frequency. The  
120      parameter  $k_a$  is the number of repetitions of intensifications and it is an input parameter of the  
121      algorithm.
- 122      3. Compute the Ensemble Average by

$$x^l(n, c) = \frac{1}{k_a} \sum_{i=1}^{k_a} S_i^l(n, c), \quad n \in \{1, n_{max}\}, c \in \{1, Ch\} \quad (1)$$

123      for ~~each row and column on the Speller Matrix~~.  $1 \leq n \leq n_{max}$  and for the channels  $1 \leq c \leq C$ .  
124      This provide an averaged signal  $x^l(n, c)$  for the twelve locations  $1 \leq l \leq 12$ .

125      2.1.2 Signal Plotting

126      Averaged signal segments are standardized and scaled ~~for  $1 \leq n \leq n_{max}$  and  $1 \leq c \leq C$~~  by

$$\tilde{x}^l(n, c) = \left[ \gamma \cdot \frac{(x(n, c) - \bar{x}(c))}{\hat{\sigma}(c)} \right] \frac{(x^l(n, c) - \bar{x}^l(c))}{\hat{\sigma}^l(c)}, \quad n \in \{1, n_{max}\}, c \in \{1, 2, Ch\} \quad (2)$$

where  $\gamma > 0$  is an input parameter of the algorithm and it is related to the image scale. In addition,  ~~$x(n, c)$~~   $\tilde{x}^l(n, c)$  is the point-to-point averaged multichannel EEG signal for the sample point  $n$  and for channel  $c$ . Lastly,

$$\bar{x}^l(c) = \frac{1}{n_{max}} \sum_{n=1}^{n_{max}} x^l(n, c)$$

and

$$\hat{\sigma}^l(c) = \left( \frac{1}{n_{max} - 1} \sum_{n=1}^{n_{max}} (x^l(n, c) - \bar{x}^l(c))^2 \right)^{\frac{1}{2}}$$

127      are the mean and estimated standard deviation of  ~~$x(n, c)$~~ ,  $n \in \{1, \dots, n_{max}\}$   $x^l(n, c)$ ,  $1 \leq n \leq n_{max}$ , for  
128      each channel  $c$ .

129      Consequently, ~~the image is constructed by placing the sample points for a pixel  $(z_1, z_2)$ , the image  $I^{(l,c)}$~~   
130      ~~is constructed~~ according to

$$I^{(l,c)}(z_1, z_2) = \begin{cases} 255 & \text{if } z_1 = \gamma n; z_2 = \tilde{x}^l(n, c) + z^l(c) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

131 where  $(z_1, z_2) \in \mathbb{N} \times \mathbb{N}$  iterate over the width (based on the length of the signal segment) and height (based  
 132 on the peak-to-peak amplitude) of the newly created image,  $n \in \{1, \dots, n_{max}\}$  and  $c \in \{1, 2, \dots, Ch\}$ .  
 133 The values  $z(c)$ ,  $c \in \{1, 2, \dots, Ch\}$  are the location on the image with  $1 \leq n \leq n_{max}$  and  $1 \leq c \leq C$ .  
 134 The value  $z^l(c)$  is the image vertical position where the signal's zero value has to be located situated in  
 135 order to fit the entire signal within the image for each channel  $c$ :

$$z^l(c) = \left\lfloor \frac{\max_n \tilde{x}(n, c) - \min_n \tilde{x}(n, c)}{2} \frac{\max_n \tilde{x}^l(n, c) - \min_n \tilde{x}^l(n, c)}{2} \right\rfloor - \left\lfloor \frac{\max_n \tilde{x}(n, c) + \min_n \tilde{x}(n, c)}{2} \frac{\max_n \tilde{x}^l(n, c) + \min_n \tilde{x}^l(n, c)}{2} \right\rfloor \quad (4)$$

136 where the minimization and maximization are carried out for  $n$  varying between  $1 \leq n \leq n_{max}$ , and  $\lfloor \cdot \rfloor$   
 137 denote the rounding to the smaller nearest integer of the number.

138 In order to complete the plot  $I^{(l,c)}$  from the pixels, the Bresenham (?) algorithm is used to interpolate  
 139 straight lines between each pair of consecutive pixels.

#### 140 2.1.3 Feature Extraction: Histogram of Gradient Orientations

141 On the generated image  $I^{(l,c)}$  For each generated image  $I^{(l,c)}$ , a keypoint  $\mathbf{kp}_p$  is placed on a pixel  $(x_{kp}, y_{kp})$   
 142  $(x_{pk}, y_{pk})$  over the image plot and a window around the keypoint is considered. A local image patch of size  
 143  $S_p \times S_p$  pixels is constructed by dividing the window in 16 blocks of size  $3s$  each one, where  $s$   
 144 is the scale of the local patch and it is an input parameter of the algorithm. It is arranged in a  $4 \times 4$  grid and  
 145 the pixel  $\mathbf{kp}_p$  is the patch center, thus  $S_p = 12s$   $X_p = 12s$  pixels.

146 A local representation of the signal shape within the patch can be described by obtaining the gradient  
 147 orientations on each of the 16 blocks  $B_{i,j}$  with  $0 \leq i, j \leq 3$  and creating a histogram of gradients. This  
 148 technique is based on Lowe's SIFT (?) method, and it is biomimetically inspired in how the visual cortex  
 149 detects shapes by analyzing orientations (?). In order to calculate the histogram, the interval  $[0, 360]$   
 150  $[0, 360]$  of possible angles is divided in 8 bins, each one at 45 degrees.

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

$$h(\theta, i, j) = 3s \sum_{\mathbf{p} \in I^{(l,c)}} w_{ang}(\angle J(\mathbf{p}) - \theta) w_{ij} \left( \frac{\mathbf{p} - \mathbf{kp}}{3s} \frac{\mathbf{p} - \mathbf{pk}}{3s} \right) |J(\mathbf{p})| \quad (5)$$

154 where  $\mathbf{p}$  is a pixel from the image  $I^{(l,c)}$ ,  $\theta$  is the angle bin with  $\theta \in$   
 155  $\{0, 45, 90, 135, 180, 225, 270, 315\}$ ,  $|J(\mathbf{p})|$  is the norm of the gradient vector in the pixel  $\mathbf{p}$  and it is  
 156 computed using finite differences and  $\angle J(\mathbf{p})$  is the angle of the gradient vector. The scalar  $w_{ang}(\cdot)$  and  
 157 vector  $w_{ij}(\cdot)$  functions are linear interpolations used by ? and ? to provide a weighting contribution to  
 158 eight adjacent bins. They are calculated as

$$w_{ij}(\mathbf{v}) = w(v_x - x_i)w(v_y - y_{ij}) \quad (6)$$

$$w_{\text{ang}}(\alpha) = \sum_k w\left(\frac{8\alpha}{2\pi} + 8r\right)$$

159 with  $0 \leq i, j \leq 3$  and

$$w_{\text{ang}}(\alpha) = \sum_{r=-1}^1 w\left(\frac{8\alpha}{2\pi} + 8r\right) \quad (7)$$

160 where  $x_i$  and  $y_i$  are the spatial bin centers located in  $x_i, y_i = \{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\}$ ,  $x_i, y_i \in \{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\}$ ,  
 161  $\mathbf{v} = (v_x, v_y)$  is a dummy vector variable and  $\alpha$  a dummy scalar variable. On the other hand,  $r$  is an integer  
 162 that can vary freely between  $[-1, 1]$  which allows the argument  $\alpha$  to be unconstrained in terms of its values  
 163 in radians. The interpolating function  $w(\cdot)$  is defined as  $w(z) = \max(0, |z| - 1)$ .

$$w(z) = \max(0, |z| - 1)$$

164 These binning functions conform a trilinear interpolation that has a combined effect of sharing the  
 165 contribution of each oriented gradient between their eight adjacent bins in a tridimensional cube in the  
 166 histogram space, and zero everywhere else.

167 Lastly, the fixed value of 3 is a magnification factor which corresponds to the number of pixels per each  
 168 block when  $s = 1$ . As the patch has 16 blocks and 8 bin angles are considered, for each location  $l$  and  
 169 channel  $c$  a feature called descriptor  $\mathbf{d}^{(l,c)}$  of 128 dimension is obtained.

170 Fig. 3 shows an example of a patch and a scheme of the histogram computation. In (A) a plot of  
 171 the signal and the patch centered around the keypoint is shown. In (B) the possible orientations on each  
 172 patch are illustrated. Only the upper-left four blocks are visible. The first eight orientations of the first  
 173 block, are labeled from 1 to 8 clockwise. The orientations of the second block  $B_{1,2}$  are labeled from 9 to  
 174 16. This labeling continues left-to-right, up-down until the eight orientations for all the sixteen blocks are  
 175 assigned. They form the corresponding kp-descriptor descriptor  $\mathbf{d}$  of 128 coordinates. Finally, in (C) an  
 176 enlarged image plot is shown where the oriented gradient vector for each pixel can be seen.

#### 177 2.1.4 Speller Matrix letter Identification

178 The aim is to identify the selected letter from the matrix. Previously, during the training phase, two  
 179 descriptors are extracted from averaged signal segments which correspond to the letter where the user was  
 180 supposed to be focusing onto. These descriptors are the P300 templates which are grouped in a template  
 181 set called  $T$ . This set is constructed using the steps described in Section 2.1.1 and the steps A and B of the  
 182 following algorithm.

##### 183 2.1.4.1 P300 ERP Extraction

184 Segments corresponding to rows-row flickering are labeled 1-6, whereas those corresponding to columns  
 185 column flickering are labeled 7-12. The whole extraction process has the following steps:

186 First highlight randomly the

- **Step A:** First highlight rows and columns from the matrix in a random permutation order and obtain the Ensemble Average as detailed in steps 1, 2 and 3 in Section 2.1.1.
- **Step AB:** Plot the signals  $x(n, e)$ ,  $n \in \{1, \dots, n_{max}\}$ ,  $e \in \{1, \dots, Ch\}$   $\tilde{x}^l(n, c)$ ,  $1 \leq n \leq n_{max}$ ,  $1 \leq c \leq C$ , according Section 2.1.2 in order to generate the images  $I_1^{row}, \dots, I_6^{row}$  and  $I_7^{col}, \dots, I_{12}^{col}$   $I^{(l,c)}$  for rows and columns, respectively  $1 \leq l \leq 12$ .
- **Step BC:** Obtain the descriptors  $d_1^{row}, \dots, d_6^{row}$  and  $d_7^{col}, \dots, d_{12}^{col}$   $\mathbf{d}^{(l,c)}$  for rows and columns, respectively from  $I_1^{row}, \dots, I_6^{row}$  and  $I_7^{col}, \dots, I_{12}^{col}$  from  $I^{(l,c)}$  in accordance to the method described in Section 2.1.3.

### 2.1.4.2 Calibration

A trial, as defined by the BCI2000 platform (?), is every attempt to select just one letter from the speller. A set of trials is used for calibration and once the calibration is complete it can be used to identify new letters from new trials.

During the calibration phase, two descriptors  $\mathbf{d}^{(l,c)}$  are extracted for each available channel, corresponding to the locations  $l$  of a selection of one previously instructed letter from the set of calibration trials. These descriptors are the P300 templates, grouped together in a template set called  $T^c$ . The set is constructed using the steps described in Section 2.1.1 and the steps A, B and C of the P300 ERP extraction process.

Additionally, the best performing channel,  $bpc$  is identified based on the the channel where the best Character Recognition Rate is obtained.

### 2.1.4.3 Letter identification

In order to identify the selected letter, the template set  $T^{bpc}$  is used as a database. Thus, new descriptors are computed and they are compared against the descriptors belonging to the calibration template set  $T^{bpc}$ .

- **Step CD:** Match to the Template  $T$  calibration template  $T^{bpc}$  by computing

$$\hat{row} = \arg \min_{u \in \{1, \dots, 6\}} \sum_{l \in \{1, \dots, 6\}} \sum_{q \in NN_T(d_u^{row})} \left\| q - \underline{d_u^{row}} \mathbf{d}^{(l,bpc)} \right\| \quad (8)$$

and

$$\hat{col} = \arg \min_{u \in \{7, \dots, 12\}} \sum_{l \in \{7, \dots, 12\}} \sum_{q \in NN_T(d_u^{col})} \left\| q - \underline{d_u^{col}} \mathbf{d}^{(l,bpc)} \right\| \quad (9)$$

where  $NN_T(d_u^l)$ ,  $l \in \{row, col\}$  is the set of the  $k$  nearest neighbors to  $d_u^l$  and  $q$  is a template descriptor that belongs to it.  $N_T(\mathbf{d}^{(l,bpc)})$  is defined as  $N_T(\mathbf{d}^{(l,bpc)}) = \{\mathbf{d} \in T^{bpc} / \mathbf{d}$  is the  $k$ -nearest neighbor of  $\mathbf{d}^{(l,bpc)}\}$  for the best performing channel. This set is obtained by sorting all the elements in  $T$  based on the  $T^{bpc}$  based on distances between them and  $d_u^l \mathbf{d}^{(l,bpc)}$ , choosing the  $k$  smaller elements with smaller values, with  $k$  a parameter of the algorithm. This procedure is a modification of based on the k-NBNN algorithm (?).

By computing the aforementioned equations, the letter of the matrix can be determined from the intersection of the row  $\hat{row}$  and column  $\hat{col}$ . Figure 2 shows a scheme of this process.

---

**220 2.2 Experimental Protocol**

221 To verify the validity of the proposed framework and method, the public dataset 008-2014 (?) published  
222 on the BNCI-Horizon website (?) by IRCCS Fondazione Santa Lucia, is used. Additionally, an own dataset  
223 with the same experimental conditions is generated. Both of them are utilized to perform an offline BCI  
224 Simulation to decode the spelled words from the provided signals.

225 The algorithm is implemented using VLFeat (?) Computer Vision libraries on MATLAB V2014a  
226 (Mathworks Inc., Natick, MA, USA). Furthermore, in order to enhance the impact of our paper and for a  
sake of reproducibility, the code of the algorithm has been made available at: <https://bitbucket.org/itba/hist>.  
227  
228

229 In the following sections the characteristics of the datasets and parameters of the identification algorithm  
230 are described.

**231 2.2.1 P300 ALS Public Dataset**

232 The experimental protocol used to generate this dataset is explained in (?) but can be summarized as  
233 follows: 8 subjects with confirmed diagnoses but on different stages of ALS disease, were recruited and  
234 accepted to perform the experiments. The Visual P300 detection task designed for this experiment consisted  
235 of spelling 7 words of 5 letters each, using the traditional P300 Speller Matrix (?). The flashing of rows  
236 and columns provide the deviant stimulus required to elicit this physiological response. The first 3 words  
237 are used for training calibration and the remaining 4 words, for testing with visual feedback. A trial, as  
defined by the BCI2000 platform (?), is every attempt to select a letter from the speller. It is composed  
238 of signal segments corresponding to  $k_a = 10$  repetitions of flashes of 6 rows and  $k_a = 10$  repetitions of  
239 flashes of 6 columns of the matrix, yielding 120 repetitions. Flashing of a row or a column is performed for  
240 0.125 s, following by a resting period (i.e. inter-stimulus interval) of the same length. After 120 repetitions  
241 an inter-trial pause is included before resuming with the following letter.

243 The recorded dataset was sampled at 256 Hz and it consisted of a scalp multichannel EEG signal for  
244 electrode channels Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8, identified according to the 10-20 International  
245 System, for each one of the 8 subjects. The recording device was a research-oriented digital EEG device  
246 (g.Mobilab, g.Tec, Austria) and the data acquisition and stimuli delivery were handled by the BCI2000  
247 open source software (?).

248 In order to assess and verify the identification of the P300 response, subjects are instructed to perform a  
249 copy-spelling task. They have to fix their attention to successive letters for copying a previously determined  
250 set of words, in contrast to a free-running operation of the speller where each user decides on its own what  
251 letter to choose.

**252 2.2.2 P300 for healthy subjects**

253 We replicate the same experiment on healthy subjects (?) using a wireless digital EEG device (g.Nutilus,  
254 g.Tec, Austria). The experimental conditions are the same as those used for the previous dataset, as detailed  
255 in section 2.2.1. The produced dataset is available in a public online repository (?).

256 Participants are recruited voluntarily and the experiment is conducted anonymously in accordance with  
257 the Declaration of Helsinki published by the World Health Organization. No monetary compensation  
258 is handed out and all participants agree and sign a written informed consent. This study is approved  
259 by the *Departamento de Investigación y Doctorado, Instituto Tecnológico de Buenos Aires (ITBA)*. All  
260 healthy subjects have normal or corrected-to-normal vision and no history of neurological disorders. The

261 experiment is performed with 8 subjects, 6 males, 2 females, 6 right-handed, 2 left-handed, average age  
 262 29.00 years, standard deviation 11.56 years, range 20-56 years.

263 EEG data is collected in a single recording session. Participants are seated in a comfortable chair, with  
 264 their vision aligned to a computer screen located one meter in front of them. The handling and processing  
 265 of the data and stimuli is conducted by the OpenVibe platform (?).

266 Gel-based active electrodes (g.LADYbird, g.Tec, Austria) are used on the same ~~locations~~positions Fz,  
 267 Cz, Pz, Oz, P3,P4, PO7 and PO8. Reference is set to the right ear lobe and ground is preset as the AFz  
 268 position. Sampling frequency is slightly different, and is set to 250 Hz, which is the closest possible to the  
 269 one used with the other dataset.

### 270 2.2.3 Parameters

271 The patch size is  ~~$S_P = 12s \times 12s$~~   $X_P = 12s \times 12s$  pixels, where  $s$  is the scale of the local patch and it  
 272 is an input parameter of the algorithm. The P300 event can have a span of 400 ms and its amplitude can  
 273 reach  $10\mu V$  (?). Hence it is necessary to utilize a ~~size patch~~ $S_P$  signal segment of size  $t_{max} = 1$  second  
 274 and a size patch  $X_P$  that could capture an entire transient event. With this purpose in consideration, the  $s$   
 275 value election is essential.

276 We propose the Equations 10 and 11 to compute the scale value in horizontal and vertical directions,  
 277 respectively.

$$s_x = \frac{\lambda \cdot F_s}{12} \cdot \gamma \frac{\lambda \cdot F_s}{12} \quad (10)$$

$$s_y = \frac{\Delta\mu V}{12} \cdot \gamma \frac{\Delta\mu V}{12} \quad (11)$$

278 where  $\lambda$  is the length in seconds covered by the patch,  ~~$F_s$~~  $F_s$  is the sampling frequency of the EEG  
 279 signal (downsampled to 16 Hz) and  $\Delta\mu V$  corresponds to the amplitude in microvolts that can be covered  
 280 by the height of the patch. The geometric structure of the patch forces a squared configuration, then  
 281 we discerned that by using  $s = s_x = s_y = 3$  and  $\gamma = 4$ , the local patch and the descriptor can  
 282 identify events of  $9\mu V$  of amplitude, with a span of  $\lambda = 0.56$  seconds. This also determines that 1  
 283 pixel represents  $\frac{1}{\gamma} = \frac{1}{4}\mu V$  on the vertical direction and  $\frac{1}{F_s \cdot \gamma} = \frac{1}{64} \cdot \frac{1}{F_s} = \frac{1}{64}$  seconds on the horizontal  
 284 direction. ~~Descriptors kp~~ The keypoints  $p_k$  are located at  $(x_{kp}, y_{kp}) = (0.55F_s \cdot \gamma, z(c)) = (35, z(c))$   
 285  $(x_{pk}, y_{pk}) = (0.55F_s \cdot \gamma, z'(c)) = (35, z'(c))$  for the corresponding channel  $c$  (see Eq. and location  $l$  (see  
 286 Equation 4)). In this way the whole transient event is captured. Figure 4 shows a patch of a signal plot  
 287 covering the complete amplitude (vertical direction) and the complete span of the signal event (horizontal  
 288 direction).

289 Lastly, the number of channels  ~~$Ch_C$~~  is equal to 8 for both datasets, and the number of  
 290 intensification sequences  $k_a$  is ~~statically assigned fixed~~ to 10. The parameter  $k$  used to construct the  
 291 set  ~~$NN_T(d_u^l)$~~ ,  $l \in \{row, col\}$   $N_T(d^{(l,c)})$  is assigned to  $k = 7$ , which was found empirically to achieve  
 292 better results. In addition, the norm used on Equations 8 and 9 is the cosine norm, and descriptors are  
 293 normalized to  $[-1, 1]$ .

### 3 RESULTS

Table ??-1 shows the results of applying the [proposed](#) algorithm to the subjects of the public dataset of ALS patients. The percentage of correctly spelled letters is calculated while performing an offline BCI Simulation. From the seven words for each subject, the first three are used [as training for calibration](#), and the remaining four [are used](#) for testing. The best performing channel [bpc](#) is informed as well. The target ratio is 1 : 36; hence chance level is 2.8%. It can be observed that the best performance of the letter identification method is reached in [various channels](#) a dissimilar channel depending on the subject [been studied](#) [being studied](#). This table shows for comparison the obtained performance rates using single-channel signals with the SVM (?) classifier. This method is configured to use a linear kernel. The best performing channel, where the best letter identification rate was achieved, is also depicted.

The Information Transfer Rate (ITR), or Bit Transfer Rate (BTR), in the case of reactive BCIs (?) depends on the amount of signal averaging required to transmit a valid and robust selection. Fig. Figure 5 shows the performance curves for varying intensification sequences [for the subjects included in the dataset of ALS patients](#). It can be noticed that the percentage of correctly identified letters depends on the number of intensification sequences [k<sub>a</sub>](#) that are used to obtain the averaged signal. Moreover, when the number of intensification sequences tend to 1, which corresponds to [single-trial letter identification](#) [single-intensification character recognition](#), the performance is reduced. As mentioned before, the SNR of the [single-trial](#) P300 [obtained from only one segment of the intensification sequence](#) is very low and the shape of its P300 component is not very well defined.

In Table ??-2 the results obtained for 8 healthy subjects are shown. The obtained performance were slightly inferior than those obtained for ALS patients but well above chance level. It can be observed that the performance is above chance level and higher than those achieved by the other method.

Tables 3 and 4 are presented in order to compare the performance of the Histogram of Gradient Orientations (HIST) method versus a multichannel version of the SWLDA and SVM classification algorithms for both datasets. The feature was formed by concatenating all the channels (?). SWLDA is the methodology proposed by the ALS dataset's publisher. Since authors ? did not report the Character Recognition Rate obtained for this dataset, we replicate their procedure and include the performance obtained with the SWLDA algorithm at letter identification. It was verified that HIST method has an improved performance at letter identification than other methods that process the signals on a channel by channel strategy, and it even has a comparable performance against other methods like SWLDA or SVM, which uses a multichannel feature.

The P300 ERP consists of two overlapping components: the P3a and P3b, the former with frontocentral distribution while the later stronger on centroparietal region (?). Hence, the standard practice is to find the stronger response on the central channel Cz (?). However, ? show that the response may also arise in occipital regions. We found that by analyzing only the waveforms, occipital channels PO8 and PO7 show higher performances for some subjects.

As subjects have varying *latencies* and *amplitudes* of their P300 components, they also have a varying stability of the *shape* of the generated ERP (?). Figure 6 shows the 10 sample P300 templates patches for patients 8 and 3 from the dataset of ALS patients. It can be discerned that in coincidence with the performance results, the P300 signature is more clear and consistent for subject 8 (A) while for subject 3 (B) the characteristic pattern is more difficult to perceive.

334 Additionally, the stability of the P300 component waveform has been extensively studied in patients  
335 with ALS (?????) where it was found that these patients have a stable P300 component, which were also  
336 sustained across different sessions. In line with these results we do not find evidence of a difference in  
337 terms of the performance obtained for the group of patients with ALS and the healthy group of volunteers.  
338 Particularly, the best performance is obtained for a subject from the ALS dataset for which, based on visual  
339 observation, the shape of they P300 component is consistently identified.

340 It is important to remark that when applied to binary images obtained from signal plots, the feature  
341 extraction method described in Section 2.1.3 generates sparse descriptors. Under this subspace we found  
342 that using the cosine metric yielded a significant performance improvement. On the other hand, the unary  
343 classification scheme based on the NBNN algorithm proved very beneficial for the P300 Speller Matrix.  
344 This is due to the fact that this approach solves the unbalance dataset problem which is inherent to the  
345 oddball paradigm (?).

## 4 DISCUSSION

346 Among other applications of Brain Computer Interfaces, the goal of the discipline is to provide  
347 communication assistance to people affected by neuro-degenerative diseases, who are the most likely  
348 population to benefit from BCI systems and EEG processing and analysis.

349 In this work, a method to detect transient P300 components from EEG signals based on their waveform  
350 characterization in digital time-space, is presented. Additionally, its validity is evaluated using a public  
351 dataset of ALS patients and an own dataset of healthy subjects.

352 ~~This method has the~~ It was verified that this method has an improved performance at letter identification  
353 than other methods that process the signals on a channel by channel strategy, and it even has a  
354 comparable performance against other methods like SWLDA or SVM, which uses a multichannel feature.  
355 Furthermore, this method has the advantage that shapes of waveforms can be analyzed in an objective  
356 way. We observed that the shape of the P300 component is more stable in occipital channels, where  
357 the performance for identifying letters is higher. We additionally verified that ALS P300 signatures are  
358 stable in comparison to those of healthy subjects. ~~Further work should be conducted over larger samples  
359 to cross-check the validity of these results.~~

360 We believe that the use of descriptors based on histogram of gradient orientation, presented in this work,  
361 can also be utilized for deriving a shape metric in the space of the P300 signals which can complement  
362 other metrics based on time-domain as those defined by ?. It is important to notice that the analysis of  
363 waveform shapes is usually performed in a qualitative approach based on visual inspection (?), ~~and~~  
364 a complementary methodology which offer a quantitative metric will be beneficial to these routinely  
365 analysis of the waveform of ERPs.

366 The goal of this work is to answer the question if a P300 component could be solely determined by  
367 inspecting automatically their waveforms. We conclude affirmatively, though two very important issues  
368 still remain:

369 First, the stability of the P300 in terms of its shape is crucial: the averaging procedure, montages, the  
370 signal to noise ratio and spatial filters all of them are non-physiological factors that affect the stability of  
371 the shape of the P300 ERP. We tested a preliminary approach to assess if the morphological shape of the  
372 P300 ~~of the averaged signal~~ can be stabilized by applying different ~~latency shifts to segments alignments~~  
373 ~~of the stacked segments (see Figure 2)~~ and we verified that there is a better performance when a correct

374 single-trial segment alignment is applied. We also applied Dynamic Time Warping (DTW) (?) to automate  
375 the alignment procedure but we were unable to find a substantial improvement. Further work to study the  
376 stability of the P300 signature component needs to be addressed.

377 The second problem is the amplitude variation of the P300. We propose a solution by standardizing the  
378 signal, shown in Eq:Equation 2. It has the effect of normalizing the peak-to-peak amplitude, moderating its  
379 variation. It has also the advantage of reducing noise that was not reduced by the averaging procedure. It  
380 is important to remark that the averaged signal variance depends on the number of single-trials-segments  
381 used to compute it (?). The standardizing process converts the signal to unit signal variance which makes it  
382 independent of the number  $k_a$  of signals averaged. Although this is initially an advantageous approach, the  
383 standardizing process reduces the amplitude of any significant P300 complex diminishing its automatic  
384 interpretation capability.

385 In our opinion, the best benefit of the presented method is that a closer collaboration with physicians can  
386 be fostered, since this procedure intent to imitate human visual observation. Automatic classification of  
387 patterns in EEG that are specifically identified by their shapes like K-Complex, Vertex Waves, Positive  
388 Occipital Sharp Transient (?) are a prospect future work to be considered. We are currently working  
389 in unpublished material analyzing KComplex-K-Complex components that could eventually provide  
390 assistance to physicians to locate these EEG patterns, specially in long recording periods, frequent in  
391 sleep research. Additionally, it can be used for artifact removal which is performed on many occasions  
392 by visually inspecting signals. This is due to the fact that the descriptors are a direct representation of  
393 the shape of signal waveforms. In line with these applications, it can be used to build a database (?) of  
394 quantitative representations of waveforms and improve atlases (?), which are currently based on qualitative  
395 descriptions of signal shapes.

## CONFLICT OF INTEREST STATEMENT

396 The authors declare that the research was conducted in the absence of any commercial or financial  
397 relationships that could be construed as a potential conflict of interest.

## AUTHOR CONTRIBUTIONS

398 This work is part of the PhD thesis of RR which is directed by JS and codirected by AV.

## FUNDING

399 This project was supported by the ITBACyT-15 funding program issued by ITBA University from Buenos  
400 Aires, Argentina.

**Table 1.** Percentage of correctly predicted letters while performing an offline BCI Simulation Character recognition rates for the best performing channel for each subject of the public dataset of ALS patients using the Histogram of Gradient (HIST) calculated from single-channel plots. The spelled words Performance rates using single-channel signals with the SVM classifier are *GATTO*, *MENTE*, *VIOLA* and *REBUS* shown for comparison. The best performing channel *bpc* for each method is visualized

Participant	BPC <i>bpc</i>	Performance HIST	<i>bpc</i>	Single Channel SVM
1	Cz	35%	Cz	15%
2	Fz	85%	PO8	25%
3	Cz	25%	Fz	5%
4	PO8	55%	Oz	5%
5	PO7	40%	P3	25%
6	PO7	60%	PO8	20%
7	PO8	80%	Fz	30%
8	PO7	95%	PO7	85%

**Table 2.** Percentage of correctly predicted letters while performing an offline BCI Simulation Character recognition rates for the own dataset of healthy subjects using the Histogram of Gradient (HIST) calculated from single-channel plots. Performance rates using single-channel signals with the SVM classifier are shown for comparison. The best performing channel *bpc* for each subject of the own dataset method is visualized. The spelled words are *MANSO*, *CINCO*, *JUEGO* and *QUESO*.

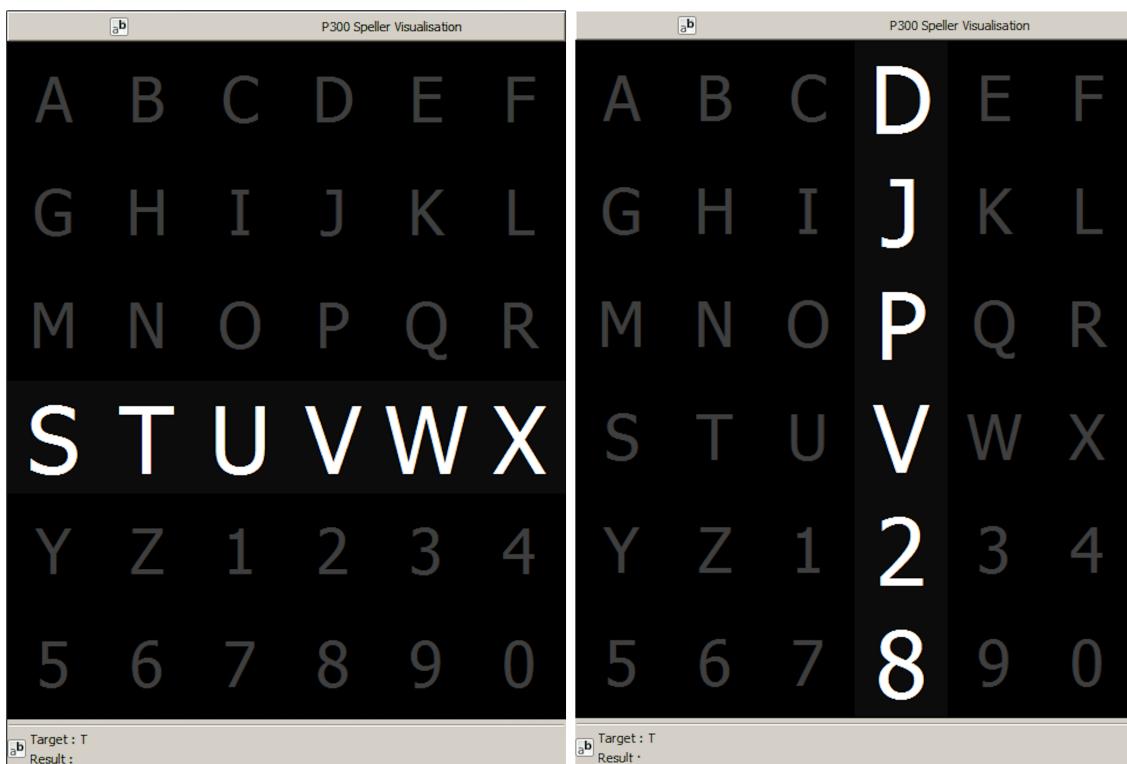
Participant	BPC <i>bpc</i>	Performance HIST	<i>bpc</i>	Single Channel SVM
1	Oz	40%	Cz	10%
2	PO7	30%	Cz	5%
3	P4	40%	P3	10%
4	P4	45%	P4	35%
5	P4	60%	P3	10%
6	Pz	50%	P4	25%
7	PO7	70%	P3	30%
8	P4	50%	PO7	10%

**Table 3.** Character recognition rates for the public dataset of ALS patients using the Histogram of Gradient (HIST) calculated from plots of single channels with the best performing channel *bpc* informed as well. Performance rates obtained by SWLDA and SVM classification algorithms with a multichannel concatenated feature.

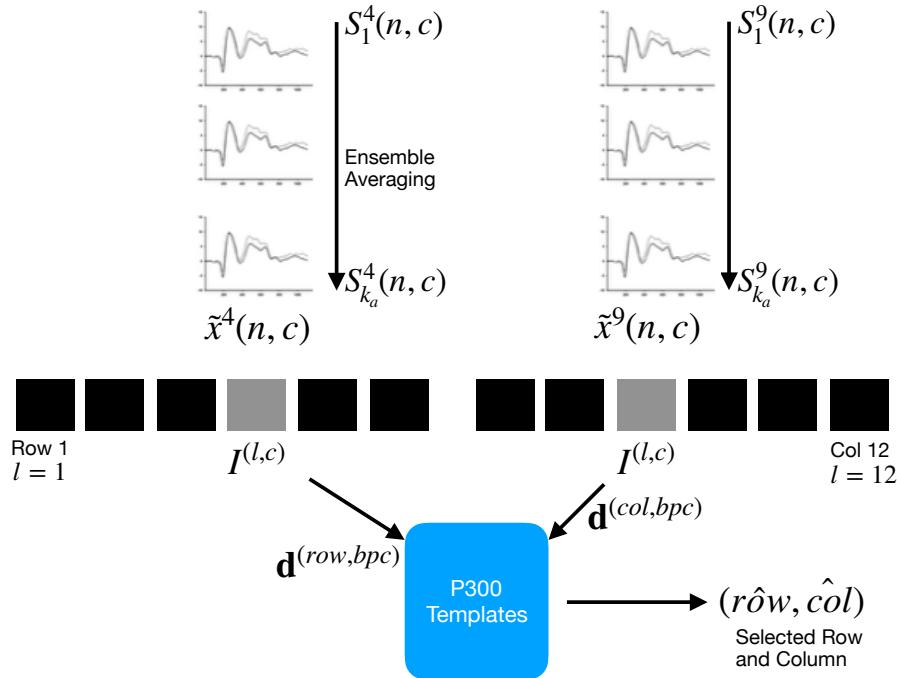
Participant	<i>bpc</i> for HIST	HIST	Multichannel SWLDA	Multichannel SVM
1	Cz	35%	45%	40%
2	Fz	85%	30%	50%
3	Cz	25%	65%	55%
4	PO8	55%	40%	50%
5	PO7	40%	35%	45%
6	PO7	60%	35%	70%
7	PO8	80%	60%	35%
8	PO7	95%	90%	95%

**Table 4.** Character recognition rates for the own dataset of healthy subjects using the Histogram of Gradient (HIST) calculated from plots of single channels with the best performing channel *bpc* informed as well. Performance rates obtained by SWLDA and SVM classification algorithms with a multichannel concatenated feature.

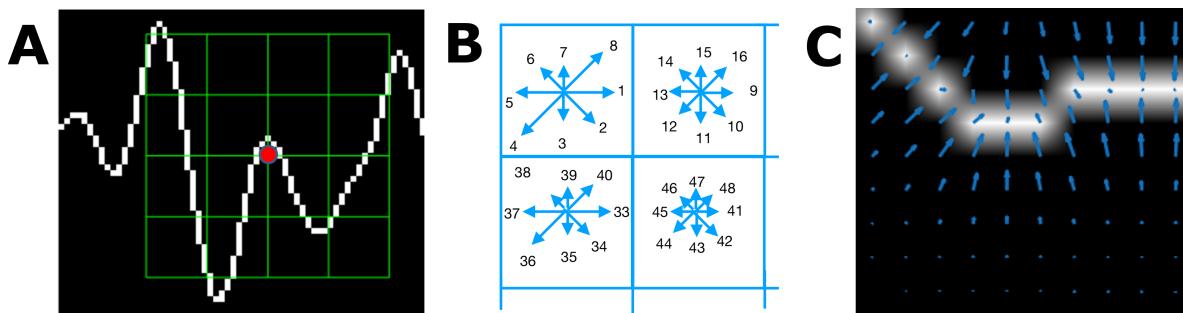
Participant	<i>bpc</i> for HIST	HIST	Multichannel SWLDA	Multichannel SVM
1	Oz	40%	65%	40%
2	PO7	30%	15%	10%
3	P4	40%	50%	25%
4	P4	45%	40%	20%
5	P4	60%	30%	20%
6	Pz	50%	35%	30%
7	PO7	70%	25%	30%
8	P4	50%	35%	20%



**Figure 1.** Example of the  $6 \times 6$  Speller Matrix used in the study [obtained from the OpenVibe software](#). Rows and columns flash **intermittently** in random permutations.

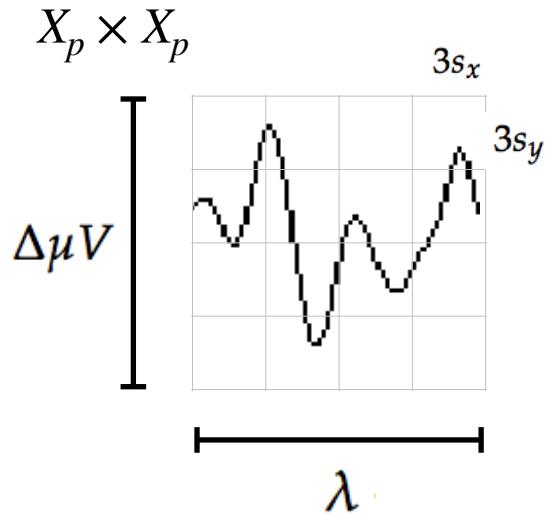


**Figure 2.** For each column and row, an averaged, standardized and scaled signal  $\tilde{x}^l(n, c)$  is obtained from the segments  $S_i^l$  corresponding to the  $k_a$  intensification sequences with  $1 \leq i \leq k_a$  and location  $l$  varying between 1 and 12. From the averaged signal, the image  $I^{(l,c)}$  of the signal plot is generated and each descriptor is computed. By comparing each descriptor against the set of templates, the P300 ERP can be detected, and finally the desired letter from the matrix can be inferred.

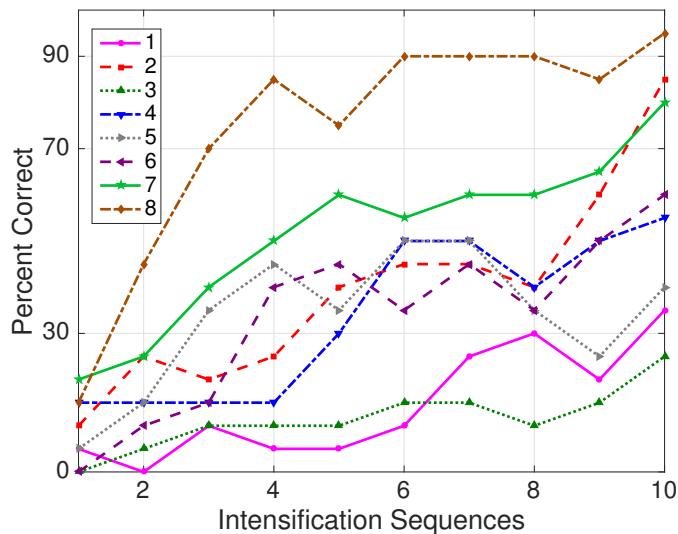


**Figure 3.** (A) Example of a plot of the signal, a keypoint and the corresponding patch. (B) A scheme of the orientation's histogram computation. Only the upper-left four blocks are visible. The first eight orientations of the first block, are labeled from 1 to 8 clockwise. The orientation of the second 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 kp-descriptor descriptor of 128 coordinates. The length of each arrow represent the value of the histogram on each direction for each block. (C) Vector field of oriented gradients. Each pixel is assigned an orientation and magnitude calculated using finite differences.

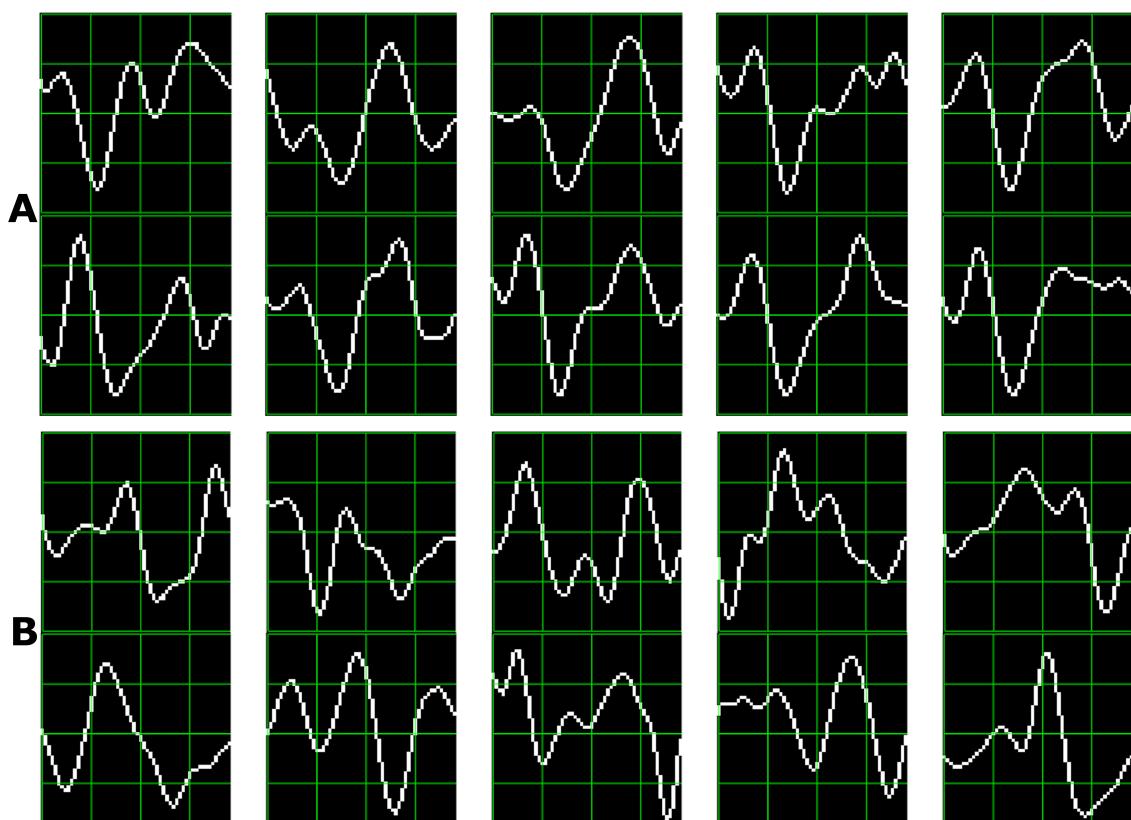
~~Single trial segments  $S_i$  are averaged for the 6 rows and 6 columns. From the averaged signal, the image of the signal plot is generated and each descriptor is computed. By comparing each descriptor against the set of templates, the P300 ERP can be detected, and finally the desired letter from the matrix can be inferred.~~



**Figure 4.** The scale of local patch is selected in order to capture the whole transient event. The size of the patch is  $X_p \times X_p$  pixels. The vertical size consists of 4 blocks of size  $3s_y$  pixels which is high enough as to contain the signal  $\Delta\mu V$ , the peak-to-peak amplitude of the transient event. The horizontal size includes 4 blocks of  $3s_x$  and covers the entire duration in seconds of the transient signal event,  $\lambda$ .



**Figure 5.** Performance curves for the eight subjects included in the dataset of ALS patients. Three out of eight subjects achieved the necessary performance to implement a valid P300 speller.



**Figure 6.** [Ten sample P300 template patches for subjects 8 \(A\) and 3 \(B\) of the ALS Dataset.](#) As traditional done in neuroscience research, [downward](#) [Downward deflection](#) is positive polarity.