Histogram of Gradient Orientations of Signal Plots applied to P300 Detection

Rodrigo Ramele [1,*], **Ana Julia Villar** [1] **and Juan Miguel Santos** [1]

[1]*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

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

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

A key contribution to this expansion has been the field of Brain Computer Interfaces (BCI) (**?**) which is the pursuit of the development of a new channel of communication particularly aimed to persons affected 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. In (**?**)
45   an approach based on Slope Horizontal Chain Code is presented, whereas in (**?**) a similar procedure
46   was implemented based on Mathematical Morphological Analysis. The seminal work of Bandt-Pompe
47   Permutation Entropy (**?**) also explores succinctly this idea as a basis to establish the time series ordinal
48   patterns. In the article (**?**), the authors introduce a method for classification of rhythmic EEG events like
49   Visual Occipital Alpha Waves and Motor Imagery Rolandic Central μ Rhythms using the histogram of
50   gradient orientations of signal plots. Inspired in that work, we propose a novel application of the developed
51   method to classify and describe transient events, particularly the P300 Event Related Potential. The
52   proposed approach is based on the waveform analysis of the shape of the EEG signal, but using histogram
53   of gradient orientations. The method is built by mimicking what traditionally electroencephalographers
54   have been performing for almost a century as it is described in (**?**): visually inspecting raw signal plots.

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

60   This article unfolds as follows: Section 2.1 is dedicated to explain the Feature Extraction method based
61   on Histogram of Gradient Orientations of the Signal Plot: Section 2.1.1 shows the preprocessing pipeline,
62   Section 2.1.3 describes the image generation of the signal plot, Section 2.1.4 presents the feature extraction
63   procedure while Section 2.1.5 introduces the Speller Matrix Letter Identification procedure. In Section 2.2,
64   the experimental protocol is expounded. Section 3 shows the results of applying the proposed technique. In
65   the final Section 4 we expose our remarks, conclusions and future work.

2 MATERIALS AND METHODS

66   The P300 (**??**) is a positive deflection of the EEG signal which occurs around 300 ms after the onset of a
67   rare and deviant stimulus that the subject is expected to attend. It is produced under the oddball paradigm (**?**)
68   and it is consistent across different subjects. It has a lower amplitude ($\pm 5\mu V$) compared to basal EEG

69 activity, reaching a Signal to Noise Ratio (SNR) of around -15 db estimated based on the amplitude of the
70 P300 response signal divided by the standard deviation of the background EEG activity (**?**). This signal can
71 be used to implement a speller application by means of a Speller Matrix (**?**). Fig. 1 shows an example of
72 the Speller Matrix used in the OpenVibe open source software (**?**), where the flashes of rows and columns
73 provide the deviant stimulus required to elicit this physiological response. Each time a row or a column
74 that contains the desired letter flashes, the corresponding synchronized EEG signal should also contain the
75 P300 signature and by detecting it, the selected letter can be identified.

2.1 Feature Extraction from Signal Plots

77 In this section, the signal preprocessing, the method for generating images from signal plots, the feature
78 extraction procedure and the Speller Matrix identification are described.

2.1.1 Preprocessing Pipeline

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

82 • **Signal Enhancement**: The preprocessing stage consists of the enhancement of the SNR of the P300
83   pattern above the level of basal EEG. The pipeline starts by applying a notch filter to the raw digital
84   signal, a 4th degree 10 Hz lowpass Butterworth filter and finally a decimation with a Finite Impulse
85   Response (FIR) filter of order 30 from the original sampling frequency down to 16 Hz(**?**).

86 • **Artifact Removal**: The multichannel EEG signal is processed on a channel by channel basis. For
87   every complete sequence of 12 intensification of 6 rows and 6 columns, a basic artifact elimination
88   procedure is implemented by removing the entire sequence when any signal deviates above/bellow
89   $\pm 70 \mu V$.

90 • **Segmentation**: For each of the 12 intensifications, a window of $t_{max} = 1$ second of the multichannel
91   signal is extracted, starting from the stimulus onset, corresponding to each row/column intensification.
92   Two of these segments should contain the P300 ERP signature time-locked to the flashing stimulus,
93   one for the row, and one for the column.

94 • **Signal Averaging**: The P300 ERP is deeply buried under background EEG so the traditional approach
95   to identify it is by point-to-point averaging the time-locked stacked signal segments. Hence the values
96   which are not related to, and not time-locked to the onset of the stimulus are canceled out (**?**).

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

2.1.2 Ensemble Average

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

103 1. Highlight randomly the rows and columns from the matrix. There is one row and one column that
104   should match the letter selected by the subject.

105 2. Repeat step 1 k_a times, obtaining the single trial segments $S_1(n,c), \ldots, S_{k_a}(n,c)$, of the EEG signal
106   where the variables $n \in \{1, \ldots, n_{max}\}$ and $c \in \{1, 2, \ldots, Ch\}$ correspond to sample points and
107   channel, respectively. The parameter Ch is the number of available EEG channels whereas $n_{max} =$

108    $F_s.t_{max}$ is the segment length and F_s is the sampling frequency. The parameter k_a is the number of
109   repetitions of intensifications and it is an input parameter of the algorithm.

110   3. Compute the Ensemble Average by

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

111   for each row and column on the Speller Matrix.

2.1.3 Signal Plotting

113   Averaged signal segments are standardized and scaled by

$$\tilde{x}(n, c) = \left\lfloor \gamma \cdot \frac{(x(n, c) - \bar{x}(c))}{\hat{\sigma}(c)} \right\rfloor, \ n \in \{1, \ldots, n_{max}\}, \ c \in \{1, 2, \ldots, Ch\} \tag{2}$$

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

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

and

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

114   are the mean and estimated standard deviation of $x(n, c), n \in \{1, \ldots, n_{max}\}$, for each channel c.

115   Consequently, the image is constructed by placing the sample points according to

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

116   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
117   on the peak-to-peak amplitude) of the newly created image, $n \in \{1, \ldots, n_{max}\}$ and $c \in \{1, 2, \ldots, Ch\}$.
118   The values $z(c), c \in \{1, 2, \ldots, Ch\}$ are the location on the image where the signal's zero value has to be
119   located in order to fit the entire signal within the image for each c:

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

120   where the minimization and maximization are carried out for n varying between $1 \leq n \leq n_{max}$.

121   In order to complete the plot from the pixels, the Bresenham (**??**) algorithm is used to interpolate straight
122   lines between each pair of consecutive pixels.

123 2.1.4 Feature Extraction: Histogram of Gradient Orientations

124   On the generated image I, a keypoint $\mathbf{kp}$ is placed on a pixel (x_{kp}, y_{kp}) over the image plot and a window
125 around the keypoint is considered. A local image patch of size $S_p \times S_p$ pixels is constructed by dividing the
126 window in 16 blocks of size $3s$ each one, where s is the scale of the local patch and it is an input parameter
127 of the algorithm. It is arranged in a 4×4 grid and the pixel $\mathbf{kp}$ is the patch center, thus $S_p = 12s$ pixels.

128   A local representation of the signal shape within the patch can be described by obtaining the gradient
129 orientations on each of the 16 blocks and creating a histogram of gradients. This technique is based on
130 Lowe's SIFT (**?**) method, and it is biomimetically inspired in how the visual cortex detects shapes by
131 analyzing orientations (**?**). In order to calculate the histogram, the interval $[0 - 360]$ of possible angles is
132 divided in 8 bins, each one at 45 degrees.

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

$$h(\theta, i, j) = 3s \sum_{\mathbf{p}} w_{\text{ang}}(\angle J(\mathbf{p}) - \theta)\, w_{ij}\left(\frac{\mathbf{p} - \mathbf{kp}}{3s}\right) |J(\mathbf{p})| \tag{5}$$

135 where $\mathbf{p}$ is a pixel from within the patch, θ is the angle bin with $\theta \in \{0, 45, 90, 135, 180, 225, 270, 315\}$,
136 $|J(\mathbf{p})|$ is the norm of the gradient vector in the pixel $\mathbf{p}$ and it is computed using finite differences and $\angle J(\mathbf{p})$
137 is the angle of the gradient vector. The scalar $w_{\text{ang}}(\cdot)$ and vector $w_{ij}(\cdot)$ functions are linear interpolations
138 used by **?** and **?** to provide a weighting contribution to eight adjacent bins. They are calculated as

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

$$w_{\text{ang}}(\alpha) = \sum_{k} w\left(\frac{8\alpha}{2\pi} + 8r\right) \tag{7}$$

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

$$w(z) = \max(0, |z| - 1) \tag{8}$$

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

146   Lastly, the fixed value of 3 is a magnification factor which corresponds to the number of pixels per each
147 block when $s = 1$. As the patch has 16 blocks and 8 bin angles are considered, a feature called *descriptor*
148 of 128 dimension is obtained.

149   Fig. 2 shows an example of a patch and a scheme of the histogram computation. In (A) a plot of the
150 signal and the patch centered around the keypoint is shown. In (B) the possible orientations on each patch

151 are illustrated. Only the upper-left four blocks are visible. The first eight orientations of the first block, are
152 labeled from 1 to 8 clockwise. The orientations of the second block $B_{1,2}$ are labeled from 9 to 16. This
153 labeling continues left-to-right, up-down until the eight orientations for all the sixteen blocks are assigned.
154 They form the corresponding **kp**-descriptor of 128 coordinates. Finally, in (C) an enlarged image plot is
155 shown where the oriented gradient vector for each pixel can be seen.

2.1.5 Speller Matrix letter Identification

157 The aim is to identify the selected letter from the matrix. Previously, during the training phase, two
158 descriptors are extracted from averaged signal segments which correspond to the letter where the user was
159 supposed to be focusing onto. These descriptors are the P300 templates which are grouped in a template
160 set called T. This set is constructed using the steps described in Section 2.1.2 and the steps A and B of
161 the following algorithm. Segments corresponding to rows are labeled 1-6, whereas those corresponding to
162 columns are labeled 7-12. The whole process has the following steps:

163 First highlight randomly the rows and columns from the matrix and obtain the Ensemble Average as
164 detailed in steps 1, 2 and 3 in Section 2.1.2.

165 • **Step A:** Plot the signals $x(n, c)$, $n \in \{1, \ldots, n_{max}\}$, $c \in \{1, \ldots, Ch\}$, according Section 2.1.3 in
166   order to generate the images $I_1^{row}, \ldots, I_6^{row}$ and $I_7^{col}, \ldots, I_{12}^{col}$ for rows and columns, respectively.
167 • **Step B:** Obtain the descriptors $d_1^{row}, \ldots, d_6^{row}$ and $d_7^{col}, \ldots, d_{12}^{col}$ for rows and columns, respectively
168   from $I_1^{row}, \ldots, I_6^{row}$ and $I_7^{col}, \ldots, I_{12}^{col}$ in accordance to the method described in Section 2.1.4.
169 • **Step C:** Match to the Template T by computing

$$\hat{row} = \arg \min_{u \in \{1,\ldots,6\}} \sum_{q \in NN_T(d_u^{row})} \|q - d_u^{row}\|^2 \tag{9}$$

170 and

$$\hat{col} = \arg \min_{u \in \{7,\ldots,12\}} \sum_{q \in NN_T(d_u^{col})} \left\|q - d_u^{col}\right\|^2 \tag{10}$$

171 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
172 descriptor that belongs to it. This set is obtained by sorting all the elements in T based on the distances
173 between them and d_u^l, choosing the k smaller elements. This procedure is a modification of the
174 k-NBNN algorithm (**?**).

175 By computing the aforementioned equations, the letter of the matrix can be determined from the intersection
176 of the row $\hat{row}$ and column $\hat{col}$. Figure 3 shows a scheme of this process.

2.2 Experimental Protocol

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

182 The algorithm is implemented using VLFeat (**?**) Computer Vision libraries on MATLAB V2014a
183 (Mathworks Inc., Natick, MA, USA).

184 In the following sections the characteristics of the datasets and parameters of the identification algorithm
185 are described.

2.2.1 P300 ALS Public Dataset

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

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

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

2.2.2 P300 for healthy subjects

208 We replicate the same experiment on healthy subjects (**?**) using a wireless digital EEG device (g.Nautilus,
209 g.Tec, Austria). The experimental conditions are the same as those used for the previous dataset, as detailed
210 in section 2.2.1.

211 Participants are recruited voluntarily and the experiment is conducted anonymously in accordance with
212 the Declaration of Helsinki published by the World Health Organization. No monetary compensation
213 is handed out and all participants agree and sign a written informed consent. This study is approved
214 by the *Departamento de Investigación y Doctorado, Instituto Tecnológico de Buenos Aires (ITBA)*. All
215 healthy subjects have normal or corrected-to-normal vision and no history of neurological disorders. The
216 experiment is performed with 8 subjects, 6 males, 2 females, 6 right-handed, 2 left-handed, average age
217 29.00 years, standard deviation 11.56 years, range 20-56 years.

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

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

225 ### 2.2.3   Parameters

226 The patch size is $S_P = 12s \times 12s$ pixels, where s is the scale of the local patch and it is an input parameter
227 of the algorithm. The P300 event can have a span of 400 ms and its amplitude can reach $10\mu V$ (**?**). Hence
228 it is necessary to utilize a size patch S_P that could capture an entire transient event. With this purpose in
229 consideration, the s value election is essential.

230 We propose the Equations 11 and 12 to compute the scale value in horizontal and vertical directions,
231 respectively.

$$s_x = \frac{\lambda \cdot Fs}{12} \cdot \gamma \tag{11}$$

$$s_y = \frac{\Delta \mu V}{12} \cdot \gamma \tag{12}$$

232 where λ is the length in seconds covered by the patch, Fs is the sampling frequency of the EEG signal
233 (downsampled to 16 Hz) and $\Delta \mu V$ corresponds to the amplitude in microvolts that can be covered by the
234 height of the patch. The geometric structure of the patch forces a squared configuration, then we discerned
235 that by using $s = s_x = s_y = 3$ and $\gamma = 4$, the local patch and the descriptor can identify events of $9\ \mu V$
236 of amplitude, with a span of $\lambda = 0.56$ seconds. This also determines that 1 pixel represents $\frac{1}{\gamma} = \frac{1}{4}\mu V$ on
237 the vertical direction and $\frac{1}{Fs.\gamma} = \frac{1}{64}$ seconds on the horizontal direction. Descriptors **kp** are located at
238 $(x_{kp}, y_{kp}) = (0.55Fs.\gamma, z(c)) = (35, z(c))$ for the corresponding channel c (see Eq. 4). In this way the
239 whole transient event is captured. Figure 4 shows a patch of a signal plot covering the complete amplitude
240 (vertical direction) and the complete span of the signal event (horizontal direction).

241 Lastly, the number of channels Ch is equal to 8 for both datasets, and the number of intensification
242 sequences k_a is statically assigned to 10. The parameter k used to construct the set $NN_T(d_u^l),\ l \in$
243 $\{row, col\}$ is assigned to $k = 7$, which was found empirically to achieve better results. In addition, the
244 norm used on Equations 9 and 10 is the cosine norm, and descriptors are normalized to $[-1, 1]$.

3   RESULTS

245 Table 1 shows the results of applying the algorithm to the subjects of the public dataset of ALS patients.
246 The percentage of correctly spelled letters is calculated while performing an offline BCI Simulation. From
247 the seven words for each subject, the first three are used as training, and the remaining four for testing. The
248 best performing channel is informed as well. The target ratio is $1 : 36$; hence chance level is 2.8%. It can
249 be observed that the best performance of the letter identification method is reached in various channels
250 depending on the subject been studied.

251 The Information Transfer Rate (ITR), or Bit Transfer Rate (BTR), in the case of reactive BCIs (**?**)
252 depends on the amount of signal averaging required to transmit a valid and robust selection. Fig. 5 shows
253 the performance curves for varying intensification sequences. It can be noticed that the percentage of
254 correctly identified letters depends on the number of intensification sequences k_a that are used to obtain the
255 averaged signal. Moreover, when the number of intensification sequences tend to 1, which corresponds
256 to single-trial letter identification, the performance is reduced. As mentioned before, the SNR of the
257 single-trial P300 is very low and the shape of its P300 component is not very well defined.

258 In Table 2 results obtained for 8 healthy subjects are shown. The obtained performance were slightly
259 inferior than those obtained for ALS patients but well above chance level.

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

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

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

276   It is important to remark that when applied to binary images obtained from signal plots, the feature
277 extraction method described in Section 2.1.4 generates sparse descriptors. Under this subspace we found
278 that using the cosine metric yielded a significant performance improvement. On the other hand, the unary
279 classification scheme based on the NBNN algorithm proved very beneficial for the P300 Speller Matrix.
280 This is due to the fact that this approach solves the unbalance dataset problem which is inherent to the
281 oddball paradigm (**?**).

4 DISCUSSION

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

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

288   This method has the advantage that shapes of waveforms can be analyzed in an objective way. We
289 observed that the shape of the P300 component is more stable in occipital channels, where the performance
290 for identifying letters is higher. We additionally verified that ALS P300 signatures are stable in comparison
291 to those of healthy subjects. Further work should be conducted over larger samples to cross-check the
292 validity of these results.

293   We believe that the use of descriptors based on histogram of gradient orientation, presented in this work,
294 can also be utilized for deriving a shape metric in the space of the P300 signals which can complement
295 other metrics based on time-domain as those defined by **?**. It is important to notice that the analysis of
296 waveform shapes is usually performed in a qualitative approach based on visual inspection (**?**).

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

300   First, the stability of the P300 in terms of its shape is crucial: the averaging procedure, montages, the
301   signal to noise ratio and spatial filters all of them are non-physiological factors that affect the stability of
302   the shape of the P300 ERP. We tested a preliminary approach to assess if the morphological shape of the
303   P300 can be stabilized by applying different latency shifts to segments and we verified that there is a better
304   performance when a correct single-trial alignment is applied. We also applied Dynamic Time Warping
305   (DTW) (**?**) but we were unable to find a substantial improvement. Further work to study the stability of the
306   P300 signature component needs to be addressed.

307   The second problem is the amplitude variation of the P300. We propose a solution by standardizing
308   the signal, shown in Eq. 2. It has the effect of normalizing the peak-to-peak amplitude, moderating its
309   variation. It has also the advantage of reducing noise that was not reduced by the averaging procedure. It
310   is important to remark that the averaged signal variance depends on the number of single-trials segments
311   used to compute it (**?**). The standardizing process converts the signal to unit signal variance which makes it
312   independent of the number k_a of signals averaged. Although this is initially an advantageous approach, the
313   standardizing process reduces the amplitude of any significant P300 complex diminishing its automatic
314   interpretation capability.

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

CONFLICT OF INTEREST STATEMENT

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

AUTHOR CONTRIBUTIONS

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

FUNDING

Table 1. Percentage of correctly predicted letters while performing an offline BCI Simulation for the best performing channel for each subject of the public dataset of ALS patients. The spelled words are *GATTO*, *MENTE*, *VIOLA* and *REBUS*.

Participant	BPC	Performance
1	Cz	35%
2	Fz	85%
3	Cz	25%
4	PO8	55%
5	PO7	40%
6	PO7	60%
7	PO8	80%
8	PO7	95%

Table 2. Percentage of correctly predicted letters while performing an offline BCI Simulation for the best performing channel for each subject of the own dataset. The spelled words are *MANSO*,*CINCO*,*JUEGO* and *QUESO*.

Participant	BPC	Performance
1	Oz	40%
2	PO7	30%
3	P4	40%
4	P4	45%
5	P4	60%
6	Pz	50%
7	PO7	70%
8	P4	50%

Figure 1. Example of the 6×6 Speller Matrix used in the study. Rows and columns flash intermittently in random permutations.
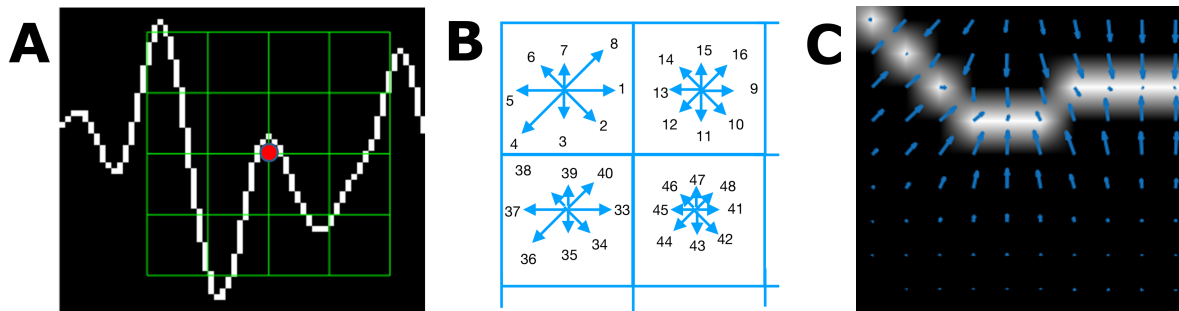

Figure 2. (A) Example of a plot of the signal, a keypoint and the corresponding patch. (B) A scheme of the orientation's histogram computation. Only the upper-left four blocks are visible. The first eight orientations of the first block, are labeled from 1 to 8 clockwise. The orientation of the second 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 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.
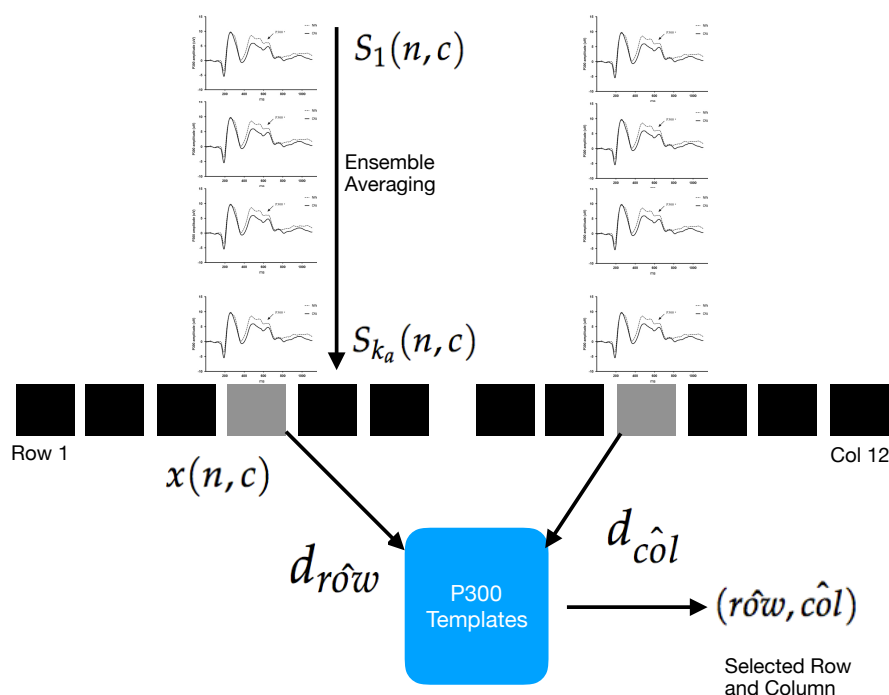
Figure 3. Single trial segments S_i are averaged for the 6 rows and 6 columns. From the averaged signal, the image of the signal plot is generated and each descriptor is computed. By comparing each descriptor against the set of templates, the P300 ERP can be detected, and finally the desired letter from the matrix can be inferred.
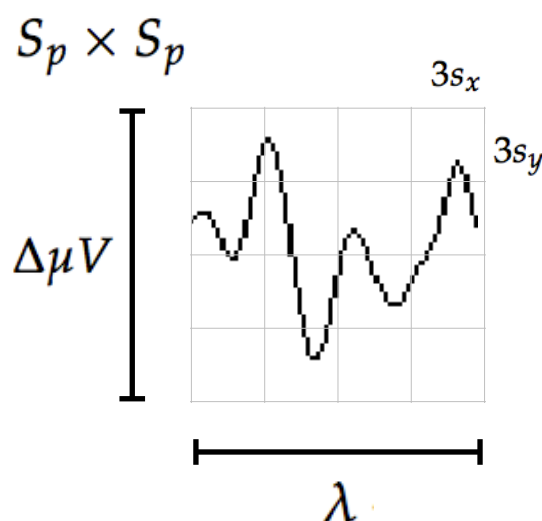

Figure 4. The scale of local patch is selected in order to capture the whole transient event. The size of the patch is $S_p \times S_p$ pixels. The vertical size consists of 4 blocks of size $3s_y$ pixels which is high enough as to contain the signal $\Delta\mu V$, the peak-to-peak amplitude of the transient event. The horizontal size includes 4 blocks of $3s_x$ and covers the entire duration in seconds of the transient signal event, λ.
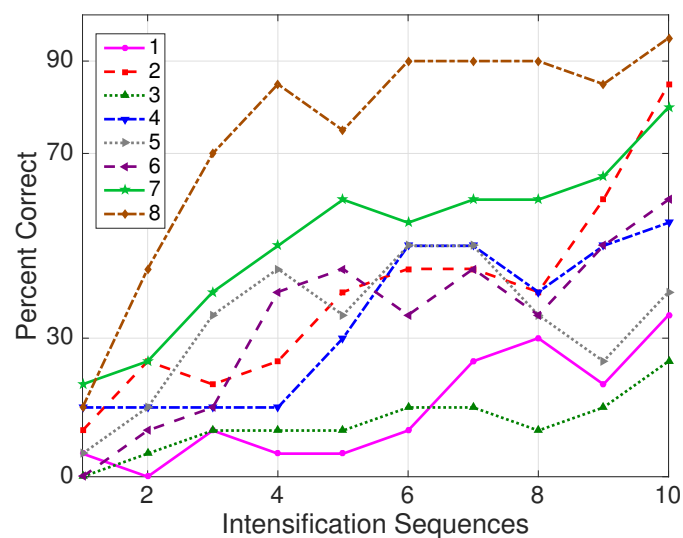
Figure 5. Performance curves for the eight subjects included in the dataset of ALS patients. Three out of eight subjects achieved the necessary performance to implement a valid P300 speller.
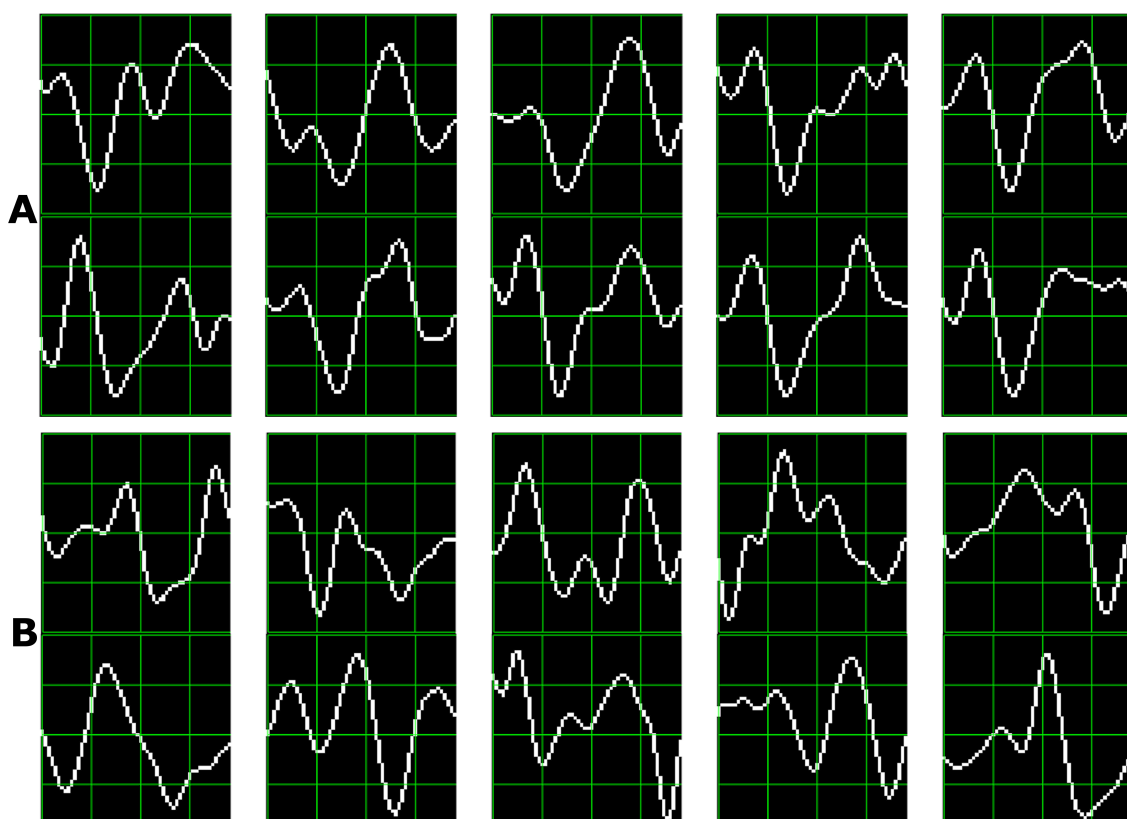

Figure 6. P300 template patches for subjects 8 (A) and 3 (B). As traditional done in neuroscience research, downward is positive polarity.