# 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

## ABSTRACT

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.

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

## 1 INTRODUCTION

Although recent advances in neuroimagining techniques, particularly radio-nuclear and radiological scanning methods (Schomer and Silva, 2010), 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 (De Vos and Debener, 2014).

A key contribution to this expansion has been the field of Brain Computer Interfaces (BCI) (Wolpaw and E., 2012) which is the pursuit of the development of a new channel of communication particularly aimed to persons affected by neurodegenerative diseases.

One noteworthy aspect of this novel communication channel is the ability to transmit information from the Central Nervous System (CNS) to a computer device and from there use that information to control a wheelchair (Carlson and del R. Millan, 2013), as input to a speller application (Guger et al., 2009), in a Virtual Reality environment (Lotte et al., 2013) or as aiding tool in a rehabilitation procedure (Jure et al., 2016). The holly grail of BCI is to implement a new complete and alternative pathway to restore lost locomotion (Wolpaw and E., 2012).

EEG signals are remarkably complex and have been characterized as a multichannel non-stationary stochastic process. Additionally, they have high variability between different subjects and even between different moments for the same subject, requiring adaptive and co-adaptive calibration and learning procedures (Clerc et al., 2016). Hence, this imposes an outstanding challenge that is necessary to overcome in order to extract information from raw EEG signals.

BCI has gained mainstream public awareness with worldwide challenge competitions like Cybathlon (Riener and Seward, 2014; Novak et al., 2018) and even been broadcasted during the inauguration ceremony of the 2014 Soccer World Cup. New developments have overcome the out-of-the-lab high-bar and they are starting to be used in real world environments (Guger et al., 2017; Huggins et al., 2016). However, they still lack the necessary robustness, and its performance is well behind any other method of human computer interaction, including any kind of detection of residual muscular movement (Clerc et al., 2016).

A few works have explored the idea of exploiting the signal waveform to analyze the EEG signal. In (Alvarado-González et al., 2016) an approach based on Slope Horizontal Chain Code is presented, whereas in (Yamaguchi et al., 2009) a similar procedure was implemented based on Mathematical Morphological Analysis. The seminal work of Bandt-Pompe Permutation Entropy (Berger et al., 2017) also explores succinctly this idea as a basis to establish the time series ordinal patterns. In the article (Ramele et al., 2016), the authors introduce a method for classification of rhythmic EEG events like Visual Occipital Alpha Waves and Motor Imagery Rolandic Central $\mu$ Rhythms using the histogram of gradient orientations of signal plots. Inspired in that work, we propose a novel application of the developed method to classify and describe transient events, particularly the P300 Event Related Potential. The proposed approach is based on the waveform analysis of the shape of the EEG signal, but using histogram of gradient orientations. The method is built by mimicking what regularly electroencephalographers have been performing for almost a century as it is described in (Hartman, 2005): visually inspecting raw signal plots.

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

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

## 2 MATERIALS AND METHODS

The P300 (Farwell and Donchin, 1988; Knuth et al., 2006) is a positive deflection of the EEG signal which occurs around 300 ms after the onset of a rare and deviant stimulus that the subject is expected to attend. It is produced under the oddball paradigm (Wolpaw and E., 2012) 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 (Hu et al., 2010). This signal can be used to implement a speller application by means of a Speller Matrix (Farwell and Donchin, 1988). This matrix is composed of 6 rows and 6 columns of numbers and letters. The subject can focus on one character of the matrix. Fig. 1 shows an example of the Speller Matrix used in the OpenVibe open source software (Renard et al., 2010), 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. Fig. 3 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.

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

- **Signal Enhancement**: The preprocessing 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(Krusienski et al., 2006).

- **Artifact Removal**: For every complete sequence of 12 intensifications of 6 rows and 6 columns, a basic artifact elimination procedure is implemented by removing the entire sequence when any signal deviates above/bellow $\pm 70 \mu V$.

- **Segmentation**: For each of the 12 intensifications of one intensification sequence, a segment $S_i^l$ of a window of $t_{max}$ seconds 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 randomly permuted, 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 EEG so the 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 (Liang and Bougrain, 2008).

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

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

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

117    2. Repeat step 1 $k_a$ times, obtaining the $1 \leq l \leq 12$ segments $S_1^l(n,c), \ldots, S_{k_a}^l(n,c)$, of the EEG
118       signal where the variables $1 \leq n \leq n_{max}$ and $1 \leq n \leq C$ correspond to sample points and channel,
119       respectively. The parameter $C$ is the number of available EEG channels whereas $n_{max} = F_s.t_{max}$
120       is the segment length and $F_s$ is the sampling frequency. The parameter $k_a$ is the number of
121       repetitions of intensifications and it is an input parameter of the algorithm.

122    3. Compute the Ensemble Average by

$$x^l(n,c) = \frac{1}{k_a} \sum_{i=1}^{k_a} S_i^l(n,c) \tag{1}$$

123       for $1 \leq n \leq n_{max}$ and for the channels $1 \leq n \leq C$. This provide an averaged signal $x^l(n,c)$ for
124       the twelve locations.

## 2.1.2   Signal Plotting

126    Averaged signal segments are standardized and scaled by

$$\tilde{x}^l(n,c) = \left\lfloor \gamma \cdot \frac{(x^l(n,c) - \bar{x}^l(c))}{\hat{\sigma}^l(c)} \right\rfloor, \; 1 \leq n \leq n_{max}, 1 \leq n \leq C \tag{2}$$

where $\gamma > 0$ is an input parameter of the algorithm and it is related to the image scale. In addition, $x^l(n,c)$
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

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

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

128    Consequently, for a pixel $(z_1, z_2)$, the image $I^l$ is constructed according to

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

129    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
130    on the peak-to-peak amplitude) of the newly created image, $1 \leq n \leq n_{max}$ and $1 \leq c \leq C$. The values

131   $z^l(c)$, are the location on the image where the signal's zero value has to be located in order to fit the entire
132   signal within the image for each c:

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

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

135       In order to complete the plot $I^{(l,c)}$ from the pixels, the Bresenham (Bresenham, 1965; Ramele et al.,
136   2016) algorithm is used to interpolate straight lines between each pair of consecutive pixels.

### 2.1.3   Feature Extraction: Histogram of Gradient Orientations

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

143       A local representation of the signal shape within the patch can be described by obtaining the gradient
144   orientations on each of the $16$ blocks $B_{i,j}$ and creating a histogram of gradients. This technique is based
145   on Lowe's SIFT (Lowe, 2004) method, and it is biomimetically inspired in how the visual cortex detects
146   shapes by analyzing orientations (Edelman et al., 1997). In order to calculate the histogram, the interval
147   $[0, 360]$ of possible angles is divided in $8$ bins, each one of $45$ degrees.

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

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

150   where $\mathbf{p}$ is a pixel from the image $I$, $\theta$ is the angle bin with $\theta \in \{0, 45, 90, 135, 180, 225, 270, 315\}$, $|J(\mathbf{p})|$
151   is the norm of the gradient vector in the pixel $\mathbf{p}$ and it is computed using finite differences and $\angle J(\mathbf{p})$ is the
152   angle of the gradient vector. The scalar $w_{\mathrm{ang}}(\cdot)$ and vector $w_{ij}(\cdot)$ functions are linear interpolations used
153   by Lowe (2004) and Vedaldi and Fulkerson (2010) to provide a weighting contribution to eight adjacent
154   bins. They are calculated as

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

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

$$w_{\mathrm{ang}}(\alpha) = \sum_{r \in [-1,1]} w\left(\frac{8\alpha}{2\pi} + 8r\right) \tag{7}$$

156   where $x_i$ and $y_i$ are the spatial bin centers located in $x_i, y_i \in \{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\}$, $\mathbf{v} = (v_x, v_y)$ is a dummy
157   vector variable and $\alpha$ a dummy scalar variable. On the other hand, $r$ is an integer that can vary freely

158 between $[-1, 1]$ which allows the argument $\alpha$ to be unconstrained in terms of its values in radians. The
159 interpolating function $w(\cdot)$ is defined as $w(z) = \max(0, |z| - 1)$.

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

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

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

173 ### 2.1.4 Speller Matrix letter Identification

174 #### *2.1.4.1 P300 ERP Extraction*

175 Segments corresponding to row flickering are labeled 1-6, whereas those corresponding to column
176 flickering are labeled 7-12. The extraction process has the following steps:

177 - **Step A:** First highlight randomly the rows and columns from the matrix and obtain the Ensemble
178 Average as detailed in steps 1, 2 and 3 in Section 2.1.1.
179 - **Step B:** Plot the signals $\tilde{x}^l(n, c)$, $1 \le n \le n_{max}$, $1 \le c \le C$, according Section 2.1.2 in order to
180 generate the images $I^{(l,c)}$ for rows and columns.
181 - **Step C:** Obtain the descriptors $d^{(l,c)}$ for rows and columns from $I^{(l,c)}$ in accordance to the method
182 described in Section 2.1.3.

183 #### *2.1.4.2 Calibration*

184 A trial, as defined by the BCI2000 platform (Schalk et al., 2004), is every attempt to select just one letter
185 from the speller. A set of trials is used for calibration and once the calibration is complete it can be used to
186 identify new letters from new trials.

187 During the calibration phase, two descriptors $d^{(l,c)}$ are extracted for each available channel for the
188 calibration trials, corresponding a previously known chosen letter. These descriptors are the P300 templates,
189 grouped together in a template set called $T$. The set is constructed using the steps described in Section
190 2.1.1 and the steps A,B and C of the P300 ERP extraction process.

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

193 #### *2.1.4.3 Letter identification*

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

196 - **Step C:** Match to the calibration template $T$ by computing

$$r\hat{o}w = \arg\min_{l \in \{1,...,6\}} \sum_{q \in N_T(d^{(l,bpc)})} \left\| q - d^{(l,bpc)} \right\|^2 \tag{8}$$

and

$$\hat{col} = \arg\min_{l \in \{7,...,12\}} \sum_{q \in N_T(d^{(l,bpc)})} \left\| q - d^{(l,bpc)} \right\|^2 \tag{9}$$

where $N_T(d^l)$ is defined as $N_T(d_u^l) = \{d \in T /$ is the k-nearest neighbor of $d^l\}$ for the best performing channel. This set is obtained by sorting all the elements in $T$ based on distances between them and $d^l$, choosing the $k$ with smaller values, with $k$ a parameter of the algorithm. This procedure is based on the k-NBNN algorithm (Boiman et al., 2008).

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

## 2.2 Experimental Protocol

To verify the validity of the proposed framework and method, the public dataset 008-2014 (Riccio et al., 2013) published on the BNCI-Horizon website (Brunner et al., 2014) by IRCCS Fondazione Santa Lucia, is used. Additionally, an own dataset with the same experimental conditions is generated. Both of them are utilized to perform an offline BCI Simulation to decode the spelled words from the provided signals.

The algorithm is implemented using VLFeat (Vedaldi and Fulkerson, 2010) Computer Vision libraries on MATLAB V2014a (Mathworks Inc., Natick, MA, USA).

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

### 2.2.1 P300 ALS Public Dataset

The experimental protocol used to generate this dataset is explained in (Riccio et al., 2013) but can be summarized as follows: 8 subjects with confirmed diagnoses but on different stages of ALS disease, were recruited and accepted to perform the experiments. The Visual P300 detection task designed for this experiment consisted of spelling 7 words of 5 letters each, using the traditional P300 Speller Matrix (Farwell and Donchin, 1988). The flashing of rows and columns provide the deviant stimulus required to elicit this physiological response. The first 3 words are used for calibration and the remaining 4 words, for testing with visual feedback. A trial, as defined by the BCI2000 platform (Schalk et al., 2004), is every attempt to select a letter from the speller. It is composed of signal segments corresponding to $k_a = 10$ repetitions of flashes of 6 rows and $k_a = 10$ repetitions of flashes of 6 columns of the matrix, yielding 120 repetitions. Flashing of a row or a column is performed for 0.125 s, following by a resting period (i.e. inter-stimulus interval) of the same length. After 120 repetitions an inter-trial pause is included before resuming with the following letter.

The recorded dataset was sampled at 256 Hz and it consisted of a scalp multichannel EEG signal for electrode channels Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8, identified according to the 10-20 International System, for each one of the 8 subjects. The recording device was a research-oriented digital EEG device (g.Mobilab, g.Tec, Austria) and the data acquisition and stimuli delivery were handled by the BCI2000 open source software (Schalk et al., 2004).

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

### 2.2.2  P300 for healthy subjects

236    We replicate the same experiment on healthy subjects (Ramele et al., 2017) using a wireless digital EEG
237    device (g.Nautilus, g.Tec, Austria). The experimental conditions are the same as those used for the previous
238    dataset, as detailed in section 2.2.1.

239    Participants are recruited voluntarily and the experiment is conducted anonymously in accordance with
240    the Declaration of Helsinki published by the World Health Organization. No monetary compensation
241    is handed out and all participants agree and sign a written informed consent. This study is approved
242    by the *Departamento de Investigación y Doctorado, Instituto Tecnológico de Buenos Aires (ITBA)*. All
243    healthy subjects have normal or corrected-to-normal vision and no history of neurological disorders. The
244    experiment is performed with 8 subjects, 6 males, 2 females, 6 right-handed, 2 left-handed, average age
245    29.00 years, standard deviation 11.56 years, range 20-56 years.

246    EEG data is collected in a single recording session. Participants are seated in a comfortable chair, with
247    their vision aligned to a computer screen located one meter in front of them. The handling and processing
248    of the data and stimuli is conducted by the OpenVibe platform (Renard et al., 2010).

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

### 2.2.3  Parameters

254    The patch size is $X_P = 12s \times 12s$ pixels, where $s$ is the scale of the local patch and it is an input parameter
255    of the algorithm. The P300 event can have a span of $400$ ms and its amplitude can reach $10\mu V$ (Rao, 2013).
256    Hence it is necessary to utilize a signal segment of size $t_max = 1$ second and a size patch $X_P$ that could
257    capture an entire transient event. With this purpose in consideration, the $s$ value election is essential.

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

$$s_x = \frac{\gamma \cdot \lambda \cdot F_s}{12} \tag{10}$$

$$s_y = \frac{\gamma \cdot \Delta \mu V}{12} \tag{11}$$

260    where $\lambda$ is the length in seconds covered by the patch, $F_s$ is the sampling frequency of the EEG signal
261    (downsampled to 16 Hz) and $\Delta\mu V$ corresponds to the amplitude in microvolts that can be covered by the
262    height of the patch. The geometric structure of the patch forces a squared configuration, then we discerned
263    that by using $s = s_x = s_y = 3$ and $\gamma = 4$, the local patch and the descriptor can identify events of 9 $\mu V$
264    of amplitude, with a span of $\lambda = 0.56$ seconds. This also determines that 1 pixel represents $\frac{1}{\gamma} = \frac{1}{4}\mu V$
265    on the vertical direction and $\frac{1}{F_s \cdot \gamma} = \frac{1}{64}$ seconds on the horizontal direction. Descriptors $\mathbf{p_k}$ are located
266    at $(x_{p_k}, y_{p_k}) = (0.55 F_s.\gamma, z^l(c)) = (35, z^l(c))$ for the corresponding channel $c$ and location $l$ (see Eq. 4).

267   In this way the whole transient event is captured. Figure 4 shows a patch of a signal plot covering the
268   complete amplitude (vertical direction) and the complete span of the signal event (horizontal direction).

269   Lastly, the number of channels $C$ is equal to $8$ for both datasets, and the number of intensification
270   sequences $k_a$ is statically assigned to $10$. The parameter $k$ used to construct the set $N_T(d^l)$ is assigned to
271   $k = 7$, which was found empirically to achieve better results. In addition, the norm used on Equations 8
272   and 9 is the cosine norm, and descriptors are normalized to $[-1, 1]$.

## 3   RESULTS

273   Table 2 shows the results of applying the algorithm to the subjects of the public dataset of ALS patients.
274   The percentage of correctly spelled letters is calculated while performing an offline BCI Simulation. From
275   the seven words for each subject, the first three are used as calibration, and the remaining four for testing.
276   The best performing channel is informed as well. The target ratio is $1 : 36$; hence chance level is $2.8\%$. It
277   can be observed that the best performance of the letter identification method is reached in various channels
278   depending on the subject been studied.

279   The Information Transfer Rate (ITR), or Bit Transfer Rate (BTR), in the case of reactive BCIs (Wolpaw
280   and E., 2012) depends on the amount of signal averaging required to transmit a valid and robust selection.
281   Fig. 5 shows the performance curves for varying intensification sequences. It can be noticed that the
282   percentage of correctly identified letters depends on the number of intensification sequences $k_a$ that are
283   used to obtain the averaged signal. Moreover, when the number of intensification sequences tend to 1, which
284   corresponds to single-intensification character recognition, the performance is reduced. As mentioned
285   before, the SNR of the P300 obtained from only one segment is very low and the shape of its P300
286   component is not very well defined.

287   In Table 3 results obtained for 8 healthy subjects are shown. The obtained performance were slightly
288   inferior than those obtained for ALS patients but well above chance level.

289   For comparison, in both tables results for Rate Character Recognition with the SWLDA algorithm are
290   added. The feature was obtained by concatenating all the channels (Krusienski et al., 2006) and the SWLDA
291   algorithm is used in accordance to the publishers of the ALS dataset (Riccio et al., 2013). The rate obtained
292   for both datasets is slightly improved in relation to SWLDA. Additionally Tables  **??** and  **??** show the
293   obtained performance for the algorithms Permutation Entropy PE and SVM support vector machine. These
294   algorithms also works on a channel by channel basis and particularly PE extracts information from the
295   waveform.

296   The P300 ERP consists of two overlapping components: the P3a and P3b, the former with frontocentral
297   distribution while the later stronger on centroparietal region (Polich, 2007). Hence, the standard practice is
298   to find the stronger response on the central channel Cz (Riccio et al., 2013). However, Krusienski et al.
299   (2006) show that the response may also arise in occipital regions. We found that by analyzing only the
300   waveforms, occipital channels PO8 and PO7 show higher performances for some subjects.

301   As subjects have varying *latencies* and *amplitudes* of their P300 components, they also have a varying
302   stability of the *shape* of the generated ERP (Nam et al., 2010). Figure 6 shows the P300 templates patches
303   for patients 8 and 3 from the dataset of ALS patients. It can be discerned that in coincidence with the
304   performance results, the P300 signature is more clear and consistent for subject 8 (A) while for subject 3
305   (B) the characteristic pattern is more difficult to perceive.

306      Additionally, the stability of the P300 component waveform has been extensively studied in patients
307   with ALS (Sellers et al., 2006; Madarame et al., 2008; Nijboer and Broermann, 2009; Mak et al., 2012;
308   McCane et al., 2015) where it was found that these patients have a stable P300 component, which were
309   also sustained across different sessions. In line with these results we do not find evidence of a difference in
310   terms of the performance obtained for the group of patients with ALS and the healthy group of volunteers.
311   Particularly, the best performance is obtained for a subject from the ALS dataset for which, based on visual
312   observation, the shape of they P300 component is consistently identified.

313      It is important to remark that when applied to binary images obtained from signal plots, the feature
314   extraction method described in Section 2.1.3 generates sparse descriptors. Under this subspace we found
315   that using the cosine metric yielded a significant performance improvement. On the other hand, the unary
316   classification scheme based on the NBNN algorithm proved very beneficial for the P300 Speller Matrix.
317   This is due to the fact that this approach solves the unbalance dataset problem which is inherent to the
318   oddball paradigm (Tibon and Levy, 2015).

## 4   DISCUSSION

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

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

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

330      We believe that the use of descriptors based on histogram of gradient orientation, presented in this work,
331   can also be utilized for deriving a shape metric in the space of the P300 signals which can complement
332   other metrics based on time-domain as those defined by Mak et al. (2012). It is important to notice
333   that the analysis of waveform shapes is usually performed in a qualitative approach based on visual
334   inspection (Sellers et al., 2006).

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

338      First, the stability of the P300 in terms of its shape is crucial: the averaging procedure, montages, the
339   signal to noise ratio and spatial filters all of them are non-physiological factors that affect the stability
340   of the shape of the P300 ERP. We tested a preliminary approach to assess if the morphological shape of
341   the P300 can be stabilized by applying different latency shifts to segments and we verified that there is a
342   better performance when a correct segment alignment is applied. We also applied Dynamic Time Warping
343   (DTW) (Casarotto et al., 2005) but we were unable to find a substantial improvement. Further work to
344   study the stability of the P300 signature component needs to be addressed.

345    The second problem is the amplitude variation of the P300. We propose a solution by standardizing the
346  signal, shown in Eq. 2. It has the effect of normalizing the peak-to-peak amplitude, moderating its variation.
347  It has also the advantage of reducing noise that was not reduced by the averaging procedure. It is important
348  to remark that the averaged signal variance depends on the number of segments used to compute it (Van
349  Drongelen, 2006). The standardizing process converts the signal to unit signal variance which makes it
350  independent of the number $k_a$ of signals averaged. Although this is initially an advantageous approach, the
351  standardizing process reduces the amplitude of any significant P300 complex diminishing its automatic
352  interpretation capability.

353    In our opinion, the best benefit of the presented method is that a closer collaboration with physicians can
354  be fostered, since this procedure intent to imitate human visual observation. Automatic classification of
355  patterns in EEG that are specifically identified by their shapes like K-Complex, Vertex Waves, Positive
356  Occipital Sharp Transient (Hartman, 2005) are a prospect future work to be considered. We are currently
357  working in unpublished material analyzing K-Complex components that could eventually provide assistance
358  to physicians to locate these EEG patterns, specially in long recording periods, frequent in sleep research.
359  Additionally, it can be used for artifact removal which is performed on many occasions by visually
360  inspecting signals. This is due to the fact that the descriptors are a direct representation of the shape
361  of signal waveforms. In line with these applications, it can be used to build a database (Chavarriaga
362  et al., 2017) of quantitative representations of waveforms and improve atlases (Hartman, 2005), which are
363  currently based on qualitative descriptions of signal shapes.

## CONFLICT OF INTEREST STATEMENT

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

## AUTHOR CONTRIBUTIONS

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

## FUNDING

## REFERENCES

369  Alvarado-González, M., Garduño, E., Bribiesca, E., Yáñez-Suárez, O., and Medina-Bañuelos, V. (2016).
370    P300 Detection Based on EEG Shape Features. *Computational and Mathematical Methods in Medicine* ,
371    1–14doi:10.1155/2016/2029791
372  Berger, S., Schneider, G., Kochs, E., and Jordan, D. (2017). Permutation Entropy: Too Complex a Measure
373    for EEG Time Series? *Entropy 2017, Vol. 19, Page 692* 19, 692. doi:10.3390/E19120692
374  Boiman, O., Shechtman, E., and Irani, M. (2008). In defense of nearest-neighbor based image classification.
375    *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR* doi:10.1109/CVPR.2008.
376    4587598
377  Bresenham, J. E. (1965). Algorithm for computer control of a digital plotter. *IBM Systems Journal* 4,
378    25–30

379 Brunner, C., Blankertz, B., Cincotti, F., Kübler, A., Mattia, D., Miralles, F., et al. (2014). BNCI Horizon
380     2020 – Towards a Roadmap for Brain / Neural Computer Interaction. *Lecture Notes in Computer Science*
381     8513, 475–486

382 Carlson, T. and del R. Millan, J. (2013). Brain-controlled wheelchairs: A robotic architecture. *IEEE*
383     *Robotics & Automation Magazine* 20, 65–73. doi:10.1109/MRA.2012.2229936

384 Casarotto, S., Bianchi, A., Cerutti, S., and Chiarenza, G. (2005). Dynamic time warping in the analysis of
385     event-related potentials. *IEEE Engineering in Medicine and Biology Magazine* 24, 68–77. doi:10.1109/
386     MEMB.2005.1384103

387 Chavarriaga, R., Fried-Oken, M., Kleih, S., Lotte, F., and Scherer, R. (2017). Heading for new shores!
388     Overcoming pitfalls in BCI design. *Brain-Computer Interfaces* 4, 60–73. doi:10.1080/2326263X.2016.
389     1263916

390 Clerc, M., Bougrain, L., and Lotte, F. (2016). *Brain-computer interfaces, Technology and applications*
391     *2(Cognitive Science)* (ISTE Ltd. and Wiley)

392 De Vos, M. and Debener, S. (2014). Mobile EEG: Towards brain activity monitoring during natural action
393     and cognition. *International Journal of Psychophysiology* 91, 1–2. doi:10.1016/j.ijpsycho.2013.10.008

394 Edelman, S., Intrator, N., and Poggio, T. (1997). Complex cells and object recognition

395 Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing
396     event-related brain potentials. *Electroencephalography and clinical neurophysiology* 70, 510–23

397 Guger, C., Allison, B. Z., and Lebedev, M. A. (2017). Introduction. In *Brain Computer Interface Research:*
398     *A State of the Art Summary 6* (Springer, Cham). 1–8. doi:10.1007/978-3-319-64373-1_1

399 Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., et al. (2009). How many people
400     are able to control a P300-based brain-computer interface (BCI)? *Neuroscience Letters* 462, 94–98.
401     doi:10.1016/j.neulet.2009.06.045

402 Hartman, a. L. (2005). *Atlas of EEG Patterns*, vol. 65 (Lippincott Williams & Wilkins). doi:10.1212/01.
403     wnl.0000174180.41994.39

404 Hu, L., Mouraux, A., Hu, Y., and Iannetti, G. D. (2010). A novel approach for enhancing the signal-to-noise
405     ratio and detecting automatically event-related potentials (ERPs) in single trials. *NeuroImage* 50, 99–111.
406     doi:10.1016/j.neuroimage.2009.12.010

407 Huggins, J. E., Alcaide-Aguirre, R. E., and Hill, K. (2016). Effects of text generation on P300 brain-
408     computer interface performance. *Brain-Computer Interfaces* 3, 112–120. doi:10.1080/2326263X.2016.
409     1203629

410 Jure, F., Carrere, L., Gentiletti, G., and Tabernig, C. (2016). BCI-FES system for neuro-rehabilitation of
411     stroke patients. *Journal of Physics: Conference Series* 705, 1–8. doi:10.1088/1742-6596/705/1/012058

412 Knuth, K. H., Shah, A. S., Truccolo, W. A., Ding, M., Bressler, S. L., and Schroeder, C. E. (2006).
413     Differentially variable component analysis: Identifying multiple evoked components using trial-to-trial
414     variability. *Journal of Neurophysiology* 95, 3257–3276. doi:10.1152/jn.00663.2005

415 Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., et al.
416     (2006). A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering*
417     3, 299–305. doi:10.1088/1741-2560/3/4/007

418 Liang, N. and Bougrain, L. (2008). Averaging techniques for single-trial analysis of oddball event-related
419     potentials. *4th International Brain-Computer* , 1–6

420 Lotte, F., Faller, J., Guger, C., Renard, Y., Pfurtscheller, G., Lécuyer, A., et al. (2013). *Combining BCI*
421     *with Virtual Reality: Towards New Applications and Improved BCI* (Berlin, Heidelberg: Springer Berlin
422     Heidelberg). 197–220. doi:10.1007/978-3-642-29746-5_10

423 Lowe, G. (2004). SIFT - The Scale Invariant Feature Transform. *International Journal* 2, 91–110

424  Madarame, T., Tanaka, H., Inoue, T., Kamata, M., and Shino, M. (2008). The development of a brain
425      computer interface device for amyotrophic lateral sclerosis patients. In *Conference Proceedings - IEEE*
426      *International Conference on Systems, Man and Cybernetics* (IEEE), 2401–2406. doi:10.1109/ICSMC.
427      2008.4811654

428  Mak, J. N., McFarland, D. J., Vaughan, T. M., McCane, L. M., Tsui, P. Z., Zeitlin, D. J., et al. (2012). EEG
429      correlates of P300-based brain-computer interface (BCI) performance in people with amyotrophic lateral
430      sclerosis. *Journal of Neural Engineering* 9. doi:10.1088/1741-2560/9/2/026014

431  McCane, L. M., Heckman, S. M., McFarland, D. J., Townsend, G., Mak, J. N., Sellers, E. W., et al.
432      (2015). P300-based brain-computer interface (BCI) event-related potentials (ERPs): People with
433      amyotrophic lateral sclerosis (ALS) vs. age-matched controls. *Clinical Neurophysiology* 126, 2124–2131.
434      doi:10.1016/j.clinph.2015.01.013

435  Nam, C. S., Li, Y., and Johnson, S. (2010). Evaluation of P300-based brain-computer interface in real-
436      world contexts. *International Journal of Human-Computer Interaction* 26, 621–637. doi:10.1080/
437      10447311003781326

438  Nijboer, F. and Broermann, U. (2009). Brain Computer Interfaces for Communication and Control in
439      Locked-in Patients. In *Graimann B., Pfurtscheller G., Allison B. (eds) Brain-Computer Interfaces. The*
440      *Frontiers Collection.* (Springer Berlin Heidelberg). 185–201. doi:10.1007/978-3-642-02091-9_11

441  Novak, D., Sigrist, R., Gerig, N. J., Wyss, D., Bauer, R., Gotz, U., et al. (2018). Benchmarking brain-
442      computer interfaces outside the laboratory: The cybathlon 2016. *Frontiers in Neuroscience* 11, 756.
443      doi:10.3389/fnins.2017.00756

444  Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* 118,
445      2128–2148. doi:10.1016/j.clinph.2007.04.019

446  Ramele, R., Villar, A. J., and Santos, J. M. (2016). BCI classification based on signal plots and SIFT
447      descriptors. In *4th International Winter Conference on Brain-Computer Interface, BCI 2016* (Yongpyong:
448      IEEE), 1–4. doi:10.1109/IWW-BCI.2016.7457454

449  [Dataset] Ramele, R., Villar, A. J., and Santos, J. M. (2017). P300-dataset rrid scr_015977. `https:`
450      `//www.kaggle.com/rramele/p300samplingdataset`

451  Rao, R. P. N. (2013). *Brain-Computer Interfacing: An Introduction* (New York, NY, USA: Cambridge
452      University Press)

453  Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., et al. (2010). OpenViBE: An
454      Open-Source Software Platform to Design, Test, and Use Brain–Computer Interfaces in Real and Virtual
455      Environments. *Presence: Teleoperators and Virtual Environments* 19, 35–53. doi:10.1162/pres.19.1.35

456  Riccio, A., Simione, L., Schettini, F., Pizzimenti, A., Inghilleri, M., Belardinelli, M. O., et al. (2013).
457      Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis. *Frontiers in*
458      *Human Neuroscience* 7, 732. doi:10.3389/fnhum.2013.00732

459  Riener, R. and Seward, L. J. (2014). Cybathlon 2016. *2014 IEEE International Conference on Systems,*
460      *Man, and Cybernetics (SMC)* , 2792–2794 doi:10.1109/SMC.2014.6974351

461  Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., and Wolpaw, J. R. (2004). BCI2000: a
462      general-purpose brain-computer interface (BCI) system. *IEEE transactions on bio-medical engineering*
463      51, 1034–43. doi:10.1109/TBME.2004.827072

464  Schomer, D. L. and Silva, F. L. D. (2010). *Niedermeyer's Electroencephalography: Basic Principles,*
465      *Clinical Applications, and Related Fields* (Walters Klutter -Lippincott Williams & Wilkins)

466  Sellers, E. W., Kübler, A., and Donchin, E. (2006). Brain-computer interface research at the University of
467      South Florida cognitive psychophysiology laboratory: The P300 speller. *IEEE Transactions on Neural*
468      *Systems and Rehabilitation Engineering* 14, 221–224. doi:10.1109/TNSRE.2006.875580

469 Tibon, R. and Levy, D. A. (2015). Striking a balance: analyzing unbalanced event-related potential data.
470     *Frontiers in psychology* 6, 555. doi:10.3389/fpsyg.2015.00555

471 Van Drongelen, W. (2006). *Signal processing for neuroscientists: an introduction to the analysis of*
472     *physiological signals* (Academic press)

473 Vedaldi, A. and Fulkerson, B. (2010). VLFeat - An open and portable library of computer vision algorithms.
474     *Design* 3, 1–4. doi:10.1145/1873951.1874249

475 Wolpaw, J. and E., W. (2012). *Brain-Computer Interfaces: Principles and Practice* (Oxford University
476     Press)

477 Yamaguchi, T., Fujio, M., Inoue, K., and Pfurtscheller, G. (2009). Design method of morphological
478     structural function for pattern recognition of EEG signals during motor imagery and cognition. In *Fourth*
479     *International Conference on Innovative Computing, Information and Control (ICICIC)*. 1558–1561.
480     doi:10.1109/ICICIC.2009.161

**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*. Results obtained with the SWLDA algorithm are also shown for comparison.

| Participant | BPC | Character Recognition Rates | |
| :---: | :---: | :---: | :---: |
| | | HIST | SWLDA |
| 1 | Cz | 35% | 45% |
| 2 | Fz | 85% | 30% |
| 3 | Cz | 25% | 65% |
| 4 | PO8 | 55% | 40% |
| 5 | PO7 | 40% | 35% |
| 6 | PO7 | 60% | 35% |
| 7 | PO8 | 80% | 60% |
| 8 | PO7 | 95% | 90% |

**Table 2.** 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*. Results obtained with the SWLDA algorithm are also shown for comparison.

| Participant | BPC | HIST | BPC | PE | BPC | SVM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Cz | 35% | Cz | 45% | Cz | 90% |
| 2 | Fz | 85% | Cz | 30% | Cz | 90% |
| 3 | Cz | 25% | Cz | 65% | Cz | 90% |
| 4 | PO8 | 55% | Cz | 40% | Cz | 90% |
| 5 | PO7 | 40% | Cz | 35% | Cz | 90% |
| 6 | PO7 | 60% | Cz | 35% | Cz | 90% |
| 7 | PO8 | 80% | Cz | 60% | Cz | 90% |
| 8 | PO7 | 95% | Cz | 90% | Cz | 90% |

**Table 3.** 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*. Results obtained with the SWLDA algorithm are also shown for comparison.

| Participant | BPC | Character Recognition Rates | |
| :---: | :---: | :---: | :---: |
| | | HIST | SWLDA |
| 1 | Oz | 40% | 65% |
| 2 | PO7 | 30% | 15% |
| 3 | P4 | 40% | 50% |
| 4 | P4 | 45% | 40% |
| 5 | P4 | 60% | 30% |
| 6 | Pz | 50% | 35% |
| 7 | PO7 | 70% | 25% |
| 8 | P4 | 50% | 35% |

**Figure 1.** Example of the $6 \times 6$ Speller Matrix used in the study. Rows and columns flash in random permutations.
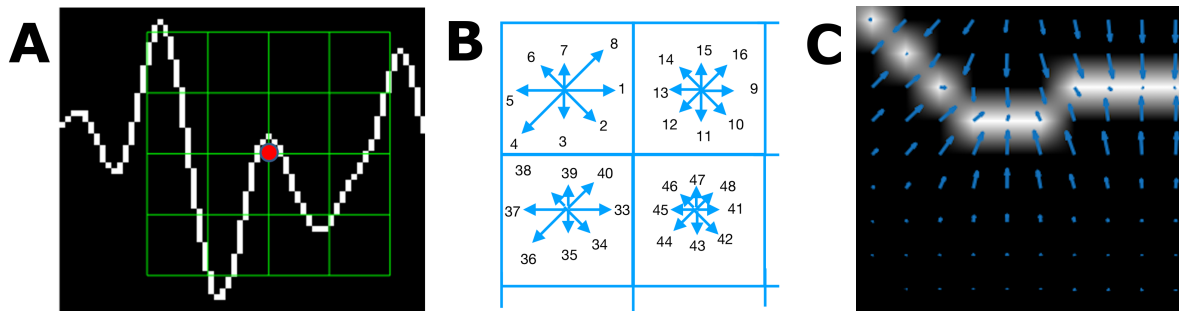


**Figure 2.** (A) Example of a plot of the signal, a keypoint and the corresponding patch. (B) A scheme of the orientation's histogram computation. Only the upper-left four blocks are visible. The first eight orientations of the first block, are labeled from 1 to 8 clockwise. The orientation of the second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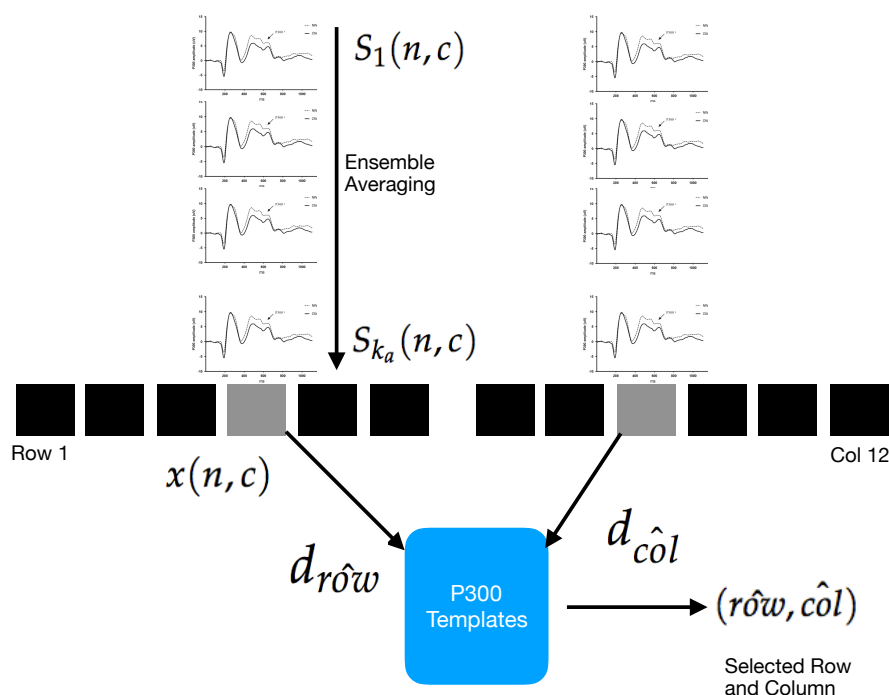
**Figure 3.** Segments $S_i$ are averaged for the 6 rows and 6 columns. From the averaged signal, the image of the signal plot is generated and each descriptor is computed. By comparing each descriptor against the set of templates, the P300 ERP can be detected, and finally the desired letter from the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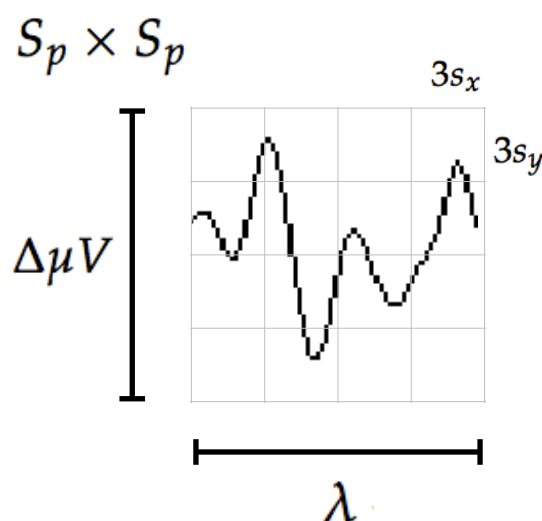


**Figure 4.** The scale of local patch is selected in order to capture the whole transient event. The size of the patch is $X_p \times X_p$ pixels. The vertical size consists of 4 blocks of size $3s_y$ pixels which is high enough as to contain the signal $\Delta\mu V$, the peak-to-peak amplitude of the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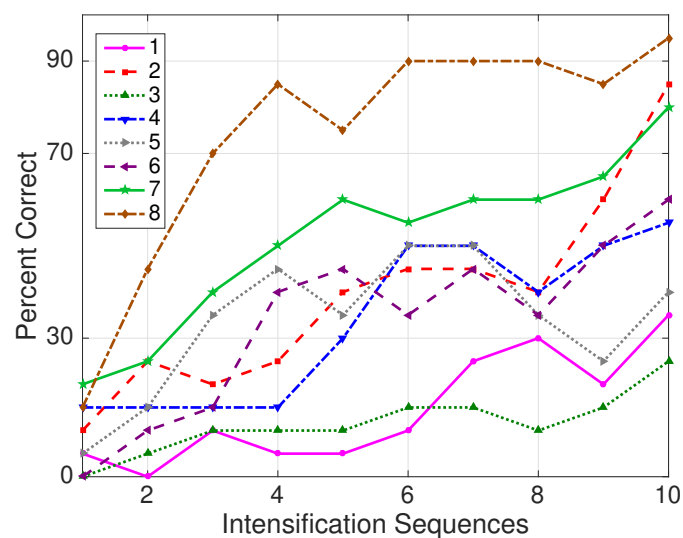
**Figure 5.** Performance curves for the eight subjects included in the dataset of ALS patients. Three out of eight subjects achieved the necessary performance to implement a valid P300 speller.
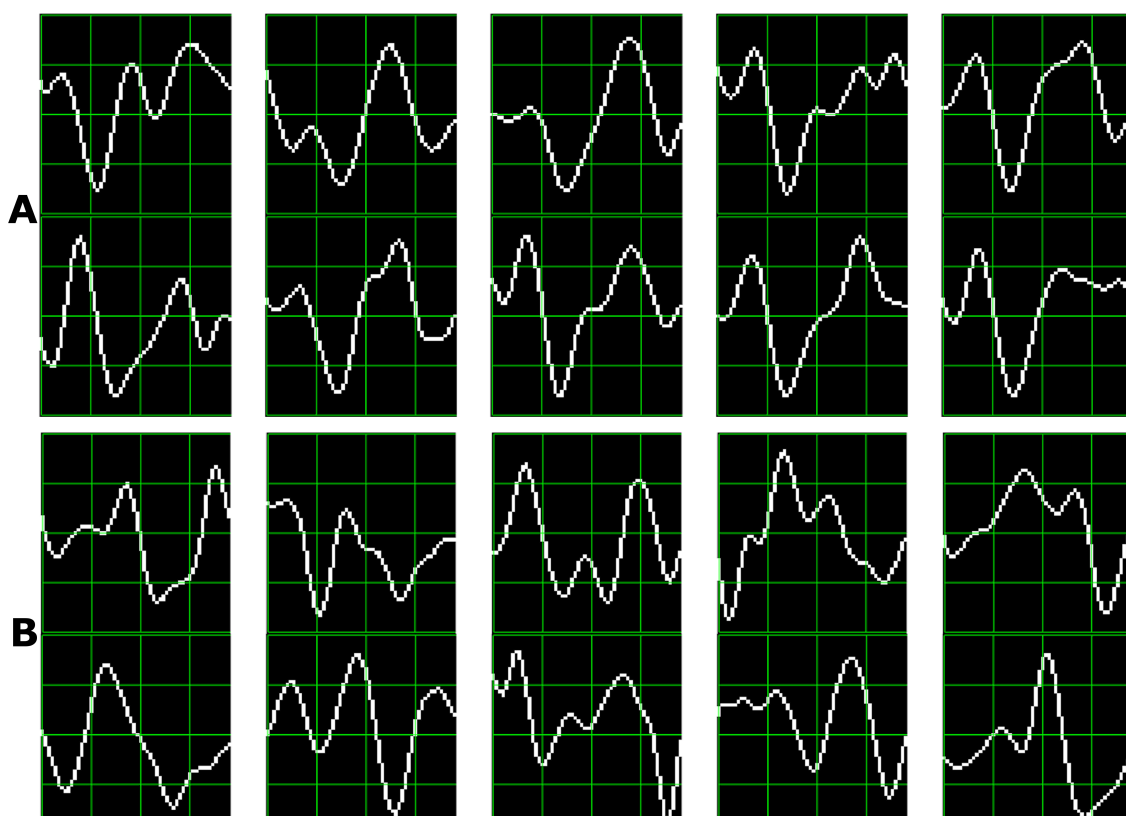


**Figure 6.** P300 template patches for subjects 8 (A) and 3 (B). Downward deflection is positive polarity.