1

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

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 neuroimagining techniques, particularly radio-nuclear and radiological
20 scanning methods (Schomer and Silva, 2010), have diminished the prospects of the traditional
21 Electroencephalography (EEG), the advent and development of digitized devices has impelled for a
22 revamping of this hundred years old technology. Their versatility, ease of use, temporal resolution, ease of
23 development and production, and its proliferation as consumer devices, are pushing EEG to become the
24 de-facto non invasive portable or ambulatory method to access and harness brain information (De Vos and
25 Debener, 2014).

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

29   One noteworthy aspect of this novel communication channel is the ability to transmit information from
30   the Central Nervous System (CNS) to a computer device and from there use that information to control a
31   wheelchair (Carlson and del R. Millan, 2013), as input to a speller application (Guger et al., 2009), in a
32   Virtual Reality environment (Lotte et al., 2013) or as aiding tool in a rehabilitation procedure (Jure et al.,
33   2016). The holly grail of BCI is to implement a new complete and alternative pathway to restore lost
34   locomotion (Wolpaw and E., 2012).

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

40   BCI has gained mainstream public awareness with worldwide challenge competitions like
41   Cybathlon (Riener and Seward, 2014; Novak et al., 2018) and even been broadcasted during the inauguration
42   ceremony of the 2014 Soccer World Cup. New developments have overcome the out-of-the-lab high-bar
43   and they are starting to be used in real world environments (Guger et al., 2017; Huggins et al., 2016).
44   However, they still lack the necessary robustness, and its performance is well behind any other method of
45   human computer interaction, including any kind of detection of residual muscular movement (Clerc et al.,
46   2016).

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

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

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

2 MATERIALS AND METHODS

70 The P300 (Farwell and Donchin, 1988; Knuth et al., 2006) is a positive deflection of the EEG signal which
71 occurs around 300 ms after the onset of a rare and deviant stimulus that the subject is expected to attend. It
72 is produced under the oddball paradigm (Wolpaw and E., 2012) and it is consistent across different subjects.
73 It has a lower amplitude ($\pm 5\mu V$) compared to basal EEG activity, reaching a Signal to Noise Ratio (SNR)
74 of around -15 db estimated based on the amplitude of the P300 response signal divided by the standard
75 deviation of the background EEG activity (Hu et al., 2010). This signal can be used to implement a speller
76 application by means of a Speller Matrix (Farwell and Donchin, 1988). Fig. 1 shows an example of the
77 Speller Matrix used in the OpenVibe open source software (Renard et al., 2010), where the flashes of rows
78 and columns provide the deviant stimulus required to elicit this physiological response. Each time a row or
79 a column that contains the desired letter flashes, the corresponding synchronized EEG signal should also
80 contain the P300 signature and by detecting it, the selected letter can be identified.

2.1 Feature Extraction from Signal Plots

81

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

2.1.1 Preprocessing Pipeline

84

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

87 • **Signal Enhancement**: The preprocessing stage consists of the enhancement of the SNR of the P300
88 pattern above the level of basal EEG. The pipeline starts by applying a notch filter to the raw digital
89 signal, a 4th degree 10 Hz lowpass Butterworth filter and finally a decimation with a Finite Impulse
90 Response (FIR) filter of order 30 from the original sampling frequency down to 16 Hz(Krusienski
91 et al., 2006).

92 • **Artifact Removal**: The multichannel EEG signal is processed on a channel by channel basis. For
93 every complete sequence of 12 intensification of 6 rows and 6 columns, a basic artifact elimination
94 procedure is implemented by removing the entire sequence when any signal deviates above/bellow
95 $\pm 70\mu V$.

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

100 • **Signal Averaging**: The P300 ERP is deeply buried under background EEG so the traditional approach
101 to identify it is by point-to-point averaging the time-locked stacked signal segments. Hence the values
102 which are not related to, and not time-locked to the onset of the stimulus are canceled out (Liang and
103 Bougrain, 2008).

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

108 ## 2.1.2 Ensemble Average

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

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

112   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
113   where the variables $n \in \{1, \ldots, n_{max}\}$ and $c \in \{1, 2, \ldots, Ch\}$ correspond to sample points and
114   channel, respectively. The parameter Ch is the number of available EEG channels whereas $n_{max} =$
115      $F_s.t_{max}$ is the segment length and F_s is the sampling frequency. The parameter k_a is the number of
116   repetitions of intensifications and it is an input parameter of the algorithm.

117   3. Compute the Ensemble Average by

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

118   for each row and column on the Speller Matrix.

119 ## 2.1.3 Signal Plotting

120   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

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

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

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

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

123   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
124   on the peak-to-peak amplitude) of the newly created image, $n \in \{1, \ldots, n_{max}\}$ and $c \in \{1, 2, \ldots, Ch\}$.
125   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
126   located in order to fit the entire signal within the image for each c:

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

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

128   In order to complete the plot from the pixels, the Bresenham (Bresenham, 1965; Ramele et al., 2016)
129   algorithm is used to interpolate straight lines between each pair of consecutive pixels.

2.1.4   Feature Extraction: Histogram of Gradient Orientations

131   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
132   around the keypoint is considered. A local image patch of size $S_p \times S_p$ pixels is constructed by dividing the
133   window in 16 blocks of size $3s$ each one, where s is the scale of the local patch and it is an input parameter
134   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.

135   A local representation of the signal shape within the patch can be described by obtaining the gradient
136   orientations on each of the 16 blocks and creating a histogram of gradients. This technique is based on
137   Lowe's SIFT (Lowe, 2004) method, and it is biomimetically inspired in how the visual cortex detects
138   shapes by analyzing orientations (Edelman et al., 1997). In order to calculate the histogram, the interval
139   $[0 - 360]$ of possible angles is divided in 8 bins, each one at 45 degrees.

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

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

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

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

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

147   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
148   vector variable and α a dummy scalar variable. On the other hand, r is an integer that can vary freely which
149   allows the argument α to be unconstrained in terms of its values in radians. The interpolating function $w(\cdot)$
150   is defined as:

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

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

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

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

164 ### 2.1.5 Speller Matrix letter Identification

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

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

173 - **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
174   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.
175 - **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
176   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.
177 - **Step C:** Match to the Template T by computing

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

178   and

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

179   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
180   descriptor that belongs to it. This set is obtained by sorting all the elements in T based on the distances
181   between them and d_u^l, choosing the k smaller elements. This procedure is a modification of the
182   k-NBNN algorithm (Boiman et al., 2008).

183 By computing the aforementioned equations, the letter of the matrix can be determined from the intersection
184 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 training 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).

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

2.2.2 P300 for healthy subjects

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

Participants are recruited voluntarily and the experiment is conducted anonymously in accordance with the declaration of Helsinki published by the World Health Organization. No monetary compensation is handed out and all participants agree and sign a written informed consent. All healthy subjects have normal or corrected-to-normal vision and no history of neurological disorders. The experiment is performed with 8 subjects, 6 males, 2 females, 6 right-handed, 2 left-handed, average age 29.00 years, standard deviation 11.56 years, range 20-56 years.

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

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

2.2.3   Parameters

234   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
235 of the algorithm. The P300 event can have a span of 400 ms and its amplitude can reach $10\mu V$ (Rao,
236 2013). Hence it is necessary to utilize a size patch S_P that could capture an entire transient event. With this
237 purpose in consideration, the s value election is essential.

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

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

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

240 where λ is the length in seconds covered by the patch, Fs is the sampling frequency of the EEG signal
241 (downsampled to 16 Hz) and $\Delta \mu V$ corresponds to the amplitude in microvolts that can be covered by the
242 height of the patch. The geometric structure of the patch forces a squared configuration, then we discerned
243 that by using $s = s_x = s_y = 3$ and $\gamma = 4$, the local patch and the descriptor can identify events of 9 μV
244 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
245 the vertical direction and $\frac{1}{Fs.\gamma} = \frac{1}{64}$ seconds on the horizontal direction. Descriptors **kp** are located at
246 $(x_{kp}, y_{kp}) = (0.55Fs.\gamma, z(c)) = (35, z(c))$ for the corresponding channel c (see Eq. 4). In this way the
247 whole transient event is captured. Figure 4 shows a patch of a signal plot covering the complete amplitude
248 (vertical direction) and the complete span of the signal event (horizontal direction).

249   Lastly, the number of channels Ch is equal to 8 for both datasets, and the number of intensification
250 sequences k_a is statically assigned to 10. The parameter k used to construct the set $NN_T(d_u^l)$, $l \in$
251 $\{row, col\}$ is assigned to $k = 7$, which was found empirically to achieve better results. In addition, the
252 norm used on Equations 9 and 10 is the cosine norm, and descriptors are normalized to $[-1, 1]$.

3   RESULTS

253 Table 1 shows the results of applying the algorithm to the subjects of the public dataset of ALS patients.
254 The percentage of correctly spelled letters is calculated while performing an offline BCI Simulation. From
255 the seven words for each subject, the first three are used as training, and the remaining four for testing.
256 The best performing channel is informed as well. The chance level is 2%. It can be observed that the best
257 performance of the letter identification method is reached in various channels depending on the subject
258 been studied.

259   The Information Transfer Rate (ITR), or Bit Transfer Rate (BTR), in the case of reactive BCIs (Wolpaw
260 and E., 2012) depends on the amount of signal averaging required to transmit a valid and robust selection.
261 Fig. 5 shows the performance curves for varying intensification sequences. It can be noticed that the
262 percentage of correctly identified letters depends on the number of intensification sequences k_a that are
263 used to obtain the averaged signal. Moreover, when the number of intensification sequences tend to 1,
264 which corresponds to single-trial letter identification, the performance is reduced. As mentioned before, the
265 SNR of the single-trial P300 is very low and the shape of its P300 component is not very well defined.

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

268   The P300 ERP consists of two overlapping components: the P3a and P3b, the former with frontocentral
269 distribution while the later stronger on centroparietal region (Polich, 2007). Hence, the standard practice is
270 to find the stronger response on the central channel Cz (Riccio et al., 2013). However, Krusienski et al.
271 (2006) show that the response may also arise in occipital regions. We found that by analyzing only the
272 waveforms, occipital channels PO8 and PO7 show higher performances for some subjects.

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

278   Additionally, the stability of the P300 component waveform has been extensively studied in patients
279 with ALS (Eric W. Sellers and Emanuel Donchin, 2006; Madarame et al., 2008; Nijboer and Broermann,
280 2009; Mak et al., 2012; McCane et al., 2015) where it was found that these patients have a stable P300
281 component, which were also sustained across different sessions. In line with these results we do not find
282 evidence of a difference in terms of the performance obtained for the group of patients with ALS and the
283 healthy group of volunteers. Particularly, the best performance is obtained for a subject from the ALS
284 dataset for which, based on visual observation, the shape of they P300 component is consistently identified.

285   It is important to remark that when applied to binary images obtained from signal plots, the feature
286 extraction method described in Section 2.1.4 generates sparse descriptors. Under this subspace we found
287 that using the cosine metric yielded a significant performance improvement. On the other hand, the unary
288 classification scheme based on the NBNN algorithm proved very beneficial for the P300 Speller Matrix.
289 This is due to the fact that this approach solves the unbalance dataset problem which is inherent to the
290 oddball paradigm (Tibon and Levy, 2015).

4 DISCUSSION

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

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

297   This method has the advantage that shapes of waveforms can be analyzed in an objective way. We
298 observed that the shape of the P300 component is more stable in occipital channels, where the performance

299 for identifying letters is higher. We additionally verified that ALS P300 signatures are stable in comparison
300 to those of healthy subjects. Further work should be conducted over larger samples to cross-check the
301 validity of these results.

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

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

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

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

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

CONFLICT OF INTEREST STATEMENT

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

AUTHOR CONTRIBUTIONS

FUNDING

REFERENCES

341   Alvarado-González, M., Garduño, E., Bribiesca, E., Yáñez-Suárez, O., and Medina-Bañuelos, V. (2016).
342   P300 Detection Based on EEG Shape Features. *Computational and Mathematical Methods in Medicine* ,
343   1–14doi:10.1155/2016/2029791
344   Berger, S., Schneider, G., Kochs, E., and Jordan, D. (2017). Permutation Entropy: Too Complex a Measure
345   for EEG Time Series? *Entropy 2017, Vol. 19, Page 692* 19, 692. doi:10.3390/E19120692
346   Boiman, O., Shechtman, E., and Irani, M. (2008). In defense of nearest-neighbor based image classification.
347   *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR* doi:10.1109/CVPR.2008.
348   4587598
349   Bresenham, J. E. (1965). Algorithm for computer control of a digital plotter. *IBM Systems Journal* 4,
350   25–30
351   Brunner, C., Blankertz, B., Cincotti, F., Kübler, A., Mattia, D., Miralles, F., et al. (2014). BNCI Horizon
352   2020 – Towards a Roadmap for Brain / Neural Computer Interaction. *Lecture Notes in Computer Science*
353   8513, 475–486
354   Carlson, T. and del R. Millan, J. (2013). Brain-controlled wheelchairs: A robotic architecture. *IEEE*
355   *Robotics & Automation Magazine* 20, 65–73. doi:10.1109/MRA.2012.2229936
356   Casarotto, S., Bianchi, A., Cerutti, S., and Chiarenza, G. (2005). Dynamic time warping in the analysis of
357   event-related potentials. *IEEE Engineering in Medicine and Biology Magazine* 24, 68–77. doi:10.1109/
358   MEMB.2005.1384103
359   Chavarriaga, R., Fried-Oken, M., Kleih, S., Lotte, F., and Scherer, R. (2017). Heading for new shores!
360   Overcoming pitfalls in BCI design. *Brain-Computer Interfaces* 4, 60–73. doi:10.1080/2326263X.2016.
361   1263916
362   Clerc, M., Bougrain, L., and Lotte, F. (2016). *Brain-computer interfaces, Technology and applications*
363   *2(Cognitive Science)* (ISTE Ltd. and Wiley)
364   De Vos, M. and Debener, S. (2014). Mobile EEG: Towards brain activity monitoring during natural action
365   and cognition. *International Journal of Psychophysiology* 91, 1–2. doi:10.1016/j.ijpsycho.2013.10.008
366   Edelman, S., Intrator, N., and Poggio, T. (1997). Complex cells and object recognition
367   Eric W. Sellers and Emanuel Donchin, A. K. (2006). Brain?Computer Interface Research at the University
368   of South Florida Cognitive Psychophysiology Laboratory:The P300 Speller , 221–224doi:10.1109/
369   TNSRE.2006.875580
370   Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing
371   event-related brain potentials. *Electroencephalography and clinical neurophysiology* 70, 510–23
372   Guger, C., Allison, B. Z., and Lebedev, M. A. (2017). Introduction. In *Brain Computer Interface Research:*
373   *A State of the Art Summary 6* (Springer, Cham). 1–8. doi:10.1007/978-3-319-64373-1_1

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

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

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

382 Huggins, J. E., Alcaide-Aguirre, R. E., and Hill, K. (2016). Effects of text generation on P300 brain-
383   computer interface performance. *Brain-Computer Interfaces* 3, 112–120. doi:10.1080/2326263X.2016.
384   1203629

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

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

390 Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., et al.
391   (2006). A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering*
392   3, 299–305. doi:10.1088/1741-2560/3/4/007

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

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

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

399 Madarame, T., Tanaka, H., Inoue, T., Kamata, M., and Shino, M. (2008). The development of a brain
400   computer interface device for amyotrophic lateral sclerosis patients. In *Conference Proceedings - IEEE*
401   *International Conference on Systems, Man and Cybernetics* (IEEE), 2401–2406. doi:10.1109/ICSMC.
402   2008.4811654

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

406 McCane, L. M., Heckman, S. M., McFarland, D. J., Townsend, G., Mak, J. N., Sellers, E. W., et al.
407   (2015). P300-based brain-computer interface (BCI) event-related potentials (ERPs): People with
408   amyotrophic lateral sclerosis (ALS) vs. age-matched controls. *Clinical Neurophysiology* 126, 2124–2131.
409   doi:10.1016/j.clinph.2015.01.013

410 Nam, C. S., Li, Y., and Johnson, S. (2010). Evaluation of P300-based brain-computer interface in real-
411   world contexts. *International Journal of Human-Computer Interaction* 26, 621–637. doi:10.1080/
412   10447311003781326

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

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

419   Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* 118,
420   2128–2148. doi:10.1016/j.clinph.2007.04.019

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

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

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

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

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

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

436   Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., and Wolpaw, J. R. (2004). BCI2000: a
437   general-purpose brain-computer interface (BCI) system. *IEEE transactions on bio-medical engineering*
438   51, 1034–43. doi:10.1109/TBME.2004.827072

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

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

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

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

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

449   Yamaguchi, T., Fujio, M., Inoue, K., and Pfurtscheller, G. (2009). Design method of morphological
450   structural function for pattern recognition of EEG signals during motor imagery and cognition. In *Fourth
451   International Conference on Innovative Computing, Information and Control (ICICIC)*. 1558–1561.
452   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*.

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.
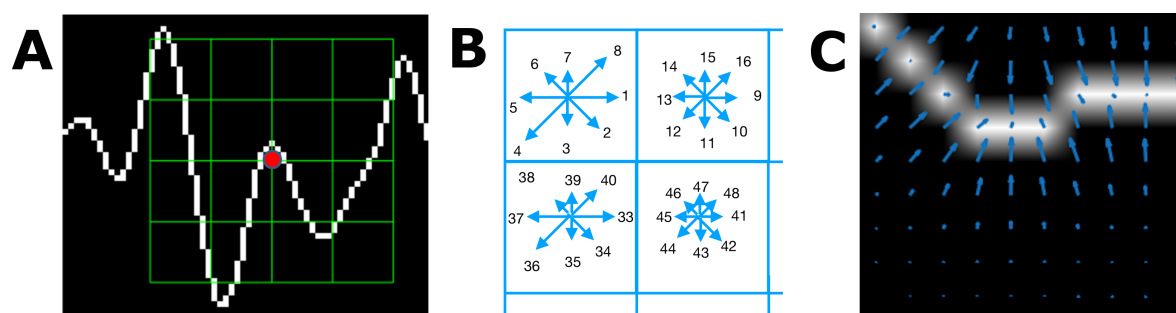

Figure 2. (A) Example of a plot of the signal, a keypoint and the corresponding patch. (B) A scheme of the orientation's histogram computation. Only the upper-left four blocks are visible. The first eight orientations of the first block, are labeled from 1 to 8 clockwise. The orientation of the second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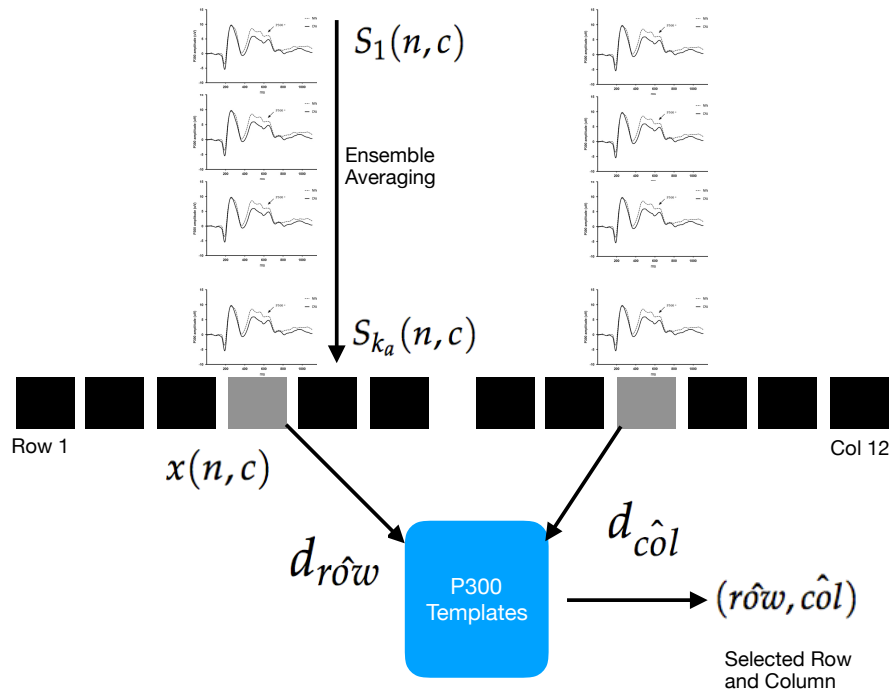
Figure 3. Single trial segments S_i are averaged for the 6 rows and 6 columns. From the averaged signal, the image of the signal plot is generated and each descriptor is computed. By comparing each descriptor against the set of templates, the P300 ERP can be detected, and finally the desired letter from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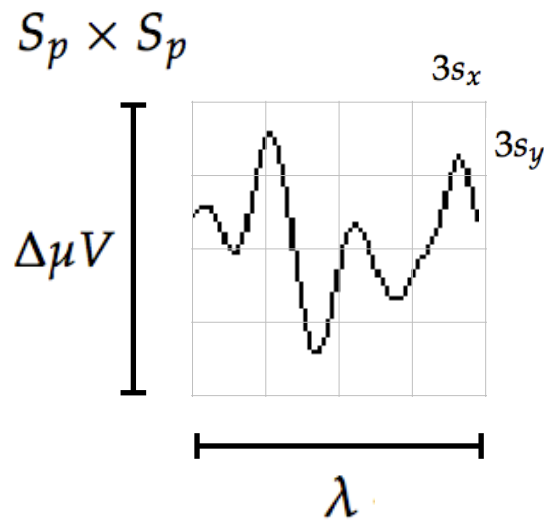

Figure 4. The scale of local patch is selected in order to capture the whole transient event. The size of the patch is $S_p \times S_p$ pixels. The vertical size consists of 4 blocks of size $3s_y$ pixels which is high enough as to contain the signal $\Delta\mu V$, the peak-to-peak amplitude of the transient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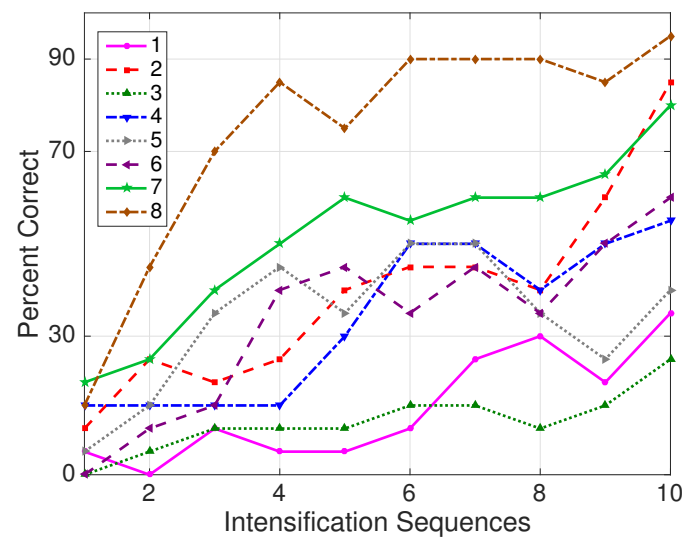
Figure 5. Performance curves for the eight subjects included in the dataset of ALS patients. Three out of eight subjects achieved the necessary performance to implement a valid P300 speller.
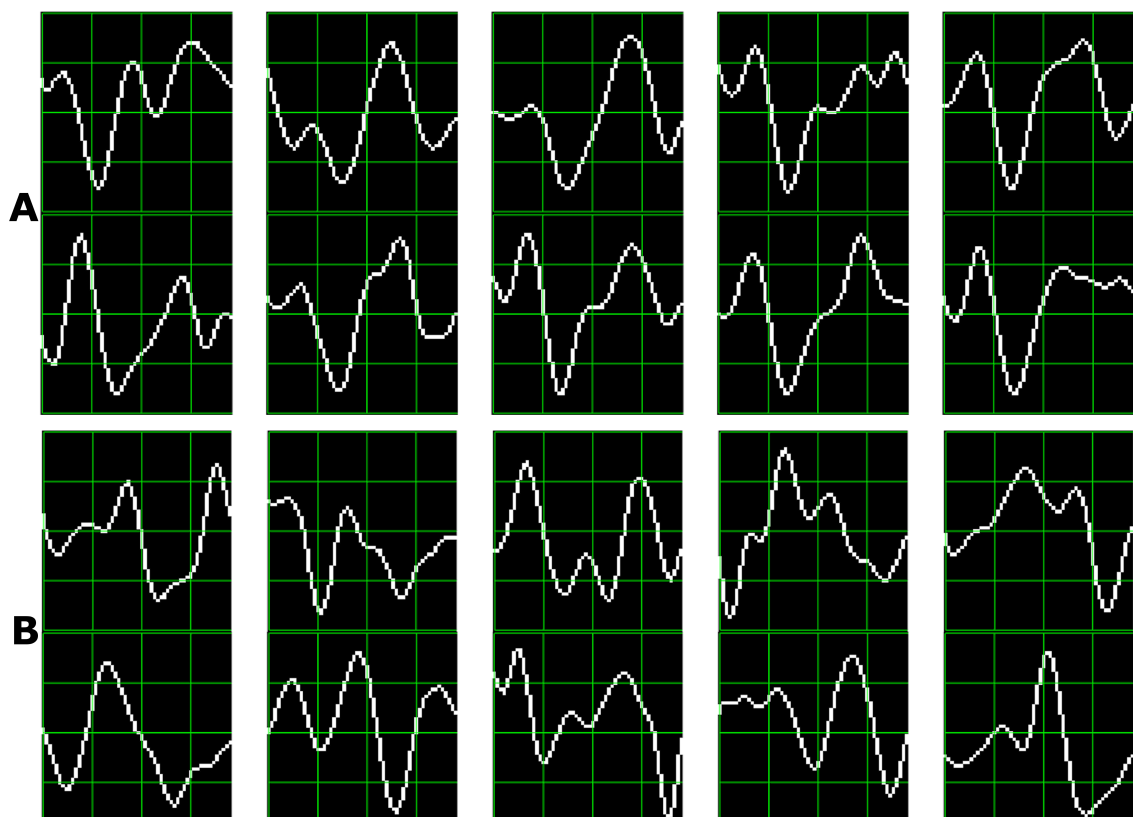

Figure 6. P300 template patches for subjects 8 (A) and 3 (B). As traditional done in neuroscience research, downward is positive polarity.