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

ABSTRACT

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 the images of the plots of the signals. 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 for healthy subjects.

Keywords: EEG, BCI, SIFT, P300, ALS, NBNN, HOG

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 impeled for a revamping of this hundred years old technology. Their versatility, ease of use, temporal resolution, ease of development and fabrication, 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.

Moreover, EEG markers (Clerc et al., 2016) that can be used to transmit volitional information are limited, and each one of them has a particular combination of appropriate methods to decode them. Inevitably, it is necessary to implement distinct and specialized algorithmic methods, to filter the signal, enhance its Signal to Noise Ratio (SNR), and try to determine some meaning out of it.

BCI has gained mainstream public awareness with worldwide challenge competitions like Cybathlon (Riener and Seward, 2014) 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 μ 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 traditionally 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) classify P300 signals based on the identification of their structure in the shape domain using histograms of gradient orientations extracted from the image of signal plots, and (2) describe the way in which this classification 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 2.1.3 describes the image generation of the signal plot, Section 2.1.4 presents the feature extraction procedure while Section 2.1.5 introduces the whole Speller Matrix Letter Identification procedure including the classification algorithm based on Naive Bayes Nearest Neighbor (NBNN) (Boiman et al., 2008). In Section 2.2, the experimental protocol is expounded. Section 3 shows the results of applying the proposal technique and a discussion is presented as well. In the final Section **??** 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 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). 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 based on Histogram of Gradient Orientations of the Signal Plot

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.

2.1.1 Preprocessing Pipeline

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

- **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**: The EEG signal matrix is processed on a channel by channel basis. For every complete sequence of 12 intensification 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, a window of $t_{max} = 1$ second of the multichannel signal is extracted, starting from the stimulus onset, corresponding to each row/column intensification. 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 traditional 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).

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

2.1.2 Ensemble Average

The procedure to obtain the averaged signal goes as follows:

1. Highlight randomly the rows and columns from the matrix. There is one row and one column that should match the letter selected by the subject.
2. Repeat step **??** k_a times, obtaining the single trial segments $S_1(n,c),\ldots,S_{k_a}(n,c)$, of the EEG signal where the variables $n \in \{1,\ldots,n_{max}\}$ and $c \in \{1,2,\ldots,Ch\}$ correspond to sample points and channel, respectively. The parameter k_a is the number of repetitions of intensifications and it is an input parameter of the algorithm.
3. Compute the Ensemble Average by

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

for each row and column.

2.1.3 Signal Plotting

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 EEG matrix for the sample point n and for channel c, $n_{max} = F_s.t_{max}$ and F_s is the sampling frequency. The parameter Ch is the number of available EEG channels. Lastly,

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

and

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

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

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

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

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

$$z(c) = \left\lfloor \frac{\max_{1\le n\le n_{max}} \tilde{x}(n,c) - \min_{1\le n\le n_{max}} \tilde{x}(n,c)}{2} \right\rfloor - \left\lfloor \frac{\max_{1\le n\le n_{max}} \tilde{x}(n,c) + \min_{1\le n\le n_{max}} \tilde{x}(n,c)}{2} \right\rfloor \tag{4}$$

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

2.1.4 Feature Extraction: Histogram of Gradient Orientations

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 around the keypoint is considered. A local image patch of size $S_p \times S_p$ pixels is constructed by dividing the window in 16 blocks of size $3s$ each one, where s is the scale of the local patch and it is an input parameter of the algorithm. It is arranged in a 4×4 grid and the pixel $\mathbf{kp}$ is the patch center, thus $S_p = 4.3.s$ pixels.

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

For each spacial bin $i, j = \{1, 2, 3, 4\}$, corresponding to the indexes of each block $B_{i,j}$, the orientations are accumulated in a 3-dimensional histogram h through the following equation:

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

where $\mathbf{p}$ is a pixel from the i, j-block $B_{i,j}$, $i, j \in \{1, 2, 3, 4\}$, θ is the angle bin with $\theta \in \{0, 45, 90, 135, 180, 225, 270, 315\}$, $|J(\mathbf{p})|$ is the norm of the gradient vector in the pixel $\mathbf{p}$ and it is computed using finite differences, $\angle J(\mathbf{p})$ is the angle of the gradient vector, and $w_{\text{ang}}(\cdot)$ and $w_{ij}(\cdot)$ are linear interpolation functions used by Lowe and Vedaldi et al. in (Lowe, 2004; Vedaldi and Fulkerson, 2010). Lastly, the fixed value of 3 is a magnification factor which corresponds to the number of pixels per each block when $s = 1$. As the patch has 16 blocks and 8 bin angles are considered, a descriptor of 128 dimension is obtained. It can be observed that the histogram is computed by multiplying by $|J(\mathbf{p})|$, so the method considers both, the magnitude and the orientation of the gradient vector.

Fig. **??** shows an example of a patch and a scheme of the histogram computation. Fig. **??** is a plot of the signal and the patch centered in the keypoint. In Fig. **??** the possible orientations on each patch are illustrated. They form the corresponding $\mathbf{kp}$-descriptor of 128 coordinates. The first two blocks are shown. Following this procedure for every assigned keypoint, we obtain N_{kp} descriptors.

2.1.5 Speller Matrix letter Identification

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

1. Highlight randomly the rows and columns from the matrix. There is one row and one column that should match the letter selected by the subject.
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 where the variables $n \in \{1, \ldots, n_{max}\}$ and $c \in \{1, 2, \ldots, Ch\}$ correspond to sample points and

channel, respectively. The parameter k_a is the number of repetitions and it is an input parameter of the algorithm.

3. Compute the Ensemble Average by

$$x(n,c) = \frac{1}{k_a}\sum_{i=1}^{k_a} S_i(n,c), n \in \{1,\dots,n_{max}\}, c \in \{1,\dots,Ch\} \tag{6}$$

for each row and column.

4. Plot the signals $x(n,c)$, $n \in \{1,\dots,n_{max}\}$, $c \in \{1,\dots,Ch\}$, according Section 2.1.3.
5. Repeat steps 2, 3 and 4 in order to generate the images $I_1^{row},\dots,I_6^{row}$ and $I_7^{col},\dots,I_{12}^{col}$ for rows and columns, respectively.
6. Obtain the descriptors $d_1^{row},\dots,d_6^{row}$ and $d_7^{col},\dots,d_{12}^{col}$ for rows and columns, respectively from $I_1^{row},\dots,I_6^{row}$ and $I_7^{col},\dots,I_{12}^{col}$ in accordance to the method described in Section 2.1.4.
7. Match to the Template T by computing

$$\hat{row} = \arg \min_{u\in\{1,\dots,6\}} \sum_{q\in NN_T(d_u^{row})} \|q - d_u^{row}\|^2 \tag{7}$$

and

$$\hat{col} = \arg \min_{u\in\{7,\dots,12\}} \sum_{q\in NN_T(d_u^{col})} \left\|q - d_u^{col}\right\|^2 \tag{8}$$

where $NN_T(d_u^l)$ is the set of the k nearest neighbors to d_u^l and q is a template descriptor that belongs to $NN_T(d_u^l),\ l \in \{row, col\}$. This set is obtained according to 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 $\hat{row}$ and column $\hat{col}$. Figure 4 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 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 scalp EEG matrix 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 asses 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 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.

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

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

2.2.3 Parameters

The patch size is $S_P = 4.3.s \times 4.3.s$ pixels, where s is the scale of the local patch and it is an input parameter of the algorithm. The P300 event can have a span of 400 ms and its amplitude can reach $10\mu V$ (Rao, 2013). Hence it is necessary to utilize a size patch S_P that could capture an entire transient event. With this purpose in consideration, the s value election is essential.

We propose the Equations 9 and 10 to compute the scale value in horizontal and vertical directions, respectively.

$$s_x = \frac{\lambda \cdot Fs}{4 \cdot 3} \cdot \gamma \tag{9}$$

$$s_y = \frac{\Delta \mu V}{4 \cdot 3} \cdot \gamma \tag{10}$$

where λ is the length in seconds covered by the patch, Fs is the sampling frequency of the EEG signal (downsampled to 16 Hz) and $\Delta\mu V$ corresponds to the amplitude in microvolts that can be covered by the height of the patch. The geometric structure of the patch forces a squared configuration, then we discerned that by using $s = s_x = s_y = 3$ and $\gamma = 4$, the local patch and the descriptor can identify events of 9 μV of amplitude, with a span of $\lambda = 0.56$ seconds. This also provides that 1 pixel represents $\frac{1}{\gamma} = \frac{1}{4}\mu V$ on the vertical direction and $\frac{1}{Fs.\gamma} = \frac{1}{64}$ seconds on the horizontal direction. Descriptors $\mathbf{kp}$ are located at $(x_{kp}, y_{kp}) = (0.55Fs.\gamma, z(c)) = (35, z(c))$ for the corresponding channel c (see Eq. 4). In this way the whole transient event is captured. Figure 5 shows a patch of a signal plot covering the complete amplitude (vertical direction) and the complete span of the signal event (horizontal direction).

Lastly, the number of channels Ch is equal to 8 for both datasets, and the number of intensification sequences k_a is statically assigned to 10. The parameter k used in the Near Neighbor $NN_T(d_u^l),\ l \in \{row, col\}$ is set to $k = 7$, following the suggestion of the article (Boiman et al., 2008). In addition, the norm used on Equations 7 and 8 is the cosine norm, and descriptors are normalized to $[-1, 1]$.

3 RESULTS

The P300 ERP consists of two overlapping components: the P3a and P3b, the former with frontocentral distribution while the later stronger on centroparietal region (Polich, 2007) then, the classical approach finds the stronger response on the central channel Cz (Riccio et al., 2013). However, Wolpaw et al (Krusienski et al., 2006) show that the response may also arise in occipital channels. In our approach, occipital channels PO8 and PO7 show higher performances for some subjects.

Table 1 shows the results of applying the algorithm to the subjects of the public dataset of ALS patients (Riccio et al., 2013). The percentage of correctly spelled letters is calculated while performing an offline BCI Simulation. From the seven words for each subject, the first three are used as training, and the remaining four for testing. The best performing channel is informed as well. The chance level is 2%. It can be observed that the best performance of the letter identification method is reached in various channels depending on the subject on study.

In Table 2 results obtained for 8 healthy subjects are shown.

The Information Transfer Rate (ITR), or Bit Transfer Rate (BTR), in the case of reactive BCIs (Wolpaw and E., 2012) depends on the amount of signal averaging required to transmit a valid and robust selection. Fig. **??** shows the performance curves for varying intensification sequences. It can be observed that the percentage of correctly identified letters depends on the number of intensification sequences k_a that are used to obtain the averaged signal.

As can be seen in the figure, as the number of intensification sequences tend to 1, which corresponds to single trial letter identification, the performance is reduced. As mentioned before, the SNR of the single trial P300 is very low. The ensemble average is carried out in order to improve the SNR but other problems arise, for example inter-trial variability due to latencies shifts that may produce an unstable P300 signature. To solve this problem, we tested an approach to assess if the morphological shape of the P300 can be stabilized by applying different shifts and we verified that there is a better performance when a correct single-trial alignment is applied, but results are still in process. We also applied Dynamic Time Warping (DTW) (Casarotto et al., 2005) but we were unable to find a substantial improvement. On the other hand amplitude, width and latency are generally affected by habituation, fatigue or level of attention and may lead to null signals(Ouyang et al., 2017). This is another source of instability of the P300 signature component that may need to be addressed.

As subjects may have different *latencies*, *amplitudes* and *width* of their P300 components, they may also have distinct *shapes* of the generated ERP. Figure 6 shows the P300 templates patches for patients 8 and 3 from the public dataset. It can be observed that in coincidence with the performance results, the P300 signature is more clear and consistent for subject 8 (Fig. **??**) while for subject 3 (Fig. **??**) the characteristic pattern is more difficult to perceive.

Another problem is the amplitude variation of the P300. We propose an approach by standardizing the signal, shown in Eq. 2. First, it has the effect of normalizing the peak-to-peak amplitude, moderating its variation. It has also the advantage of reducing noise that were not reduced by the averaging procedure. It is important to remark that the signal variance depends on the number of single-trials used to compute it (Van Drongelen, 2006). The standardizing process converts the signal to unit signal variance which makes it independent of the number k_a of signals averaged. This is another advantage of this approach. On the hand, the standardizing process reduces the amplitude of any significant P300 complex diminishing its automatic interpretation capability.

For both datasets, the experimental protocol uses a very short inter-stimulus interval which has the potential to increase the ITR but at the same time it reduces the amplitude of the P300 response, hence it may be more difficult to detect it (Rao, 2013).

4 DISCUSSION

Among other applications of BCI analysis, the goal of the entire discipline is to provide communication assistance to people affected by neuro-degenerative diseases.

In this work, a method to detect transient P300 components from EEG signals based on their waveform characterization in digital time-space, is presented. Additionally, its validity is evaluated using a public dataset of ALS patients and an own dataset of healthy subjects. The objective of this article is to explore new techniques of P300 automatic interpretation. We know that we have to keep working but the results are promising.

The method works on a channel by channel basis; in this way the best performing channel can be identified and used it to reduce the number of required EEG electrodes, leading to the development of more ergonomic capturing device.

We observed that the shape of the wave is more stable in occipital channels, where the performance for identifying letters is higher. It is important to evaluate if there is any correlation between the lower performance obtained for some subjects and the characterization of their ALS stage. Although, we did not find any evidence in terms of the digital signal shape and the best response was obtained for an ALS subject. Moreover, this method can be used as an alternate BCI predictor.

The use of descriptors based on histogram of gradient orientation, presented in this work, also can be utilized for deriving a shape metric in the space of the P300 signals.

By means of the empirical experiments, we concluded that the stability of the P300 in terms of shape is crucial: synchronization averaging, inter-stimulus interval, montages, the signal to noise ratio and spatial filters can all of them affect the stability of the shape of the P300 ERP.

In our opinion, the best application of this approach is that a closer collaboration with physicians can be fostered, because our method intent to imitate human visual observation. Automatic classification of patterns in EEG that are specifically identified by their shapes (e.g. K-Complex, Vertex Waves, Positive Occipital Sharp Transient (Hartman, 2005)) is a prospect future work to be considered. We are currently working in unpublished material analyzing KComplex that also provide assistance to physician to locate these EEG patterns, specially in long recording periods, frequent in sleep research. Additionally, it can be used for artifact removal which is performed on many occasions by visually inspecting the signal. This is due to the fact that the descriptors are directly based on the signals behavior in shape domain. In line with these applications, it can be used to build a database (Chavarriaga et al., 2017) of descriptors and improve atlases (Hartman, 2005).

We also want to solve the problem of signal stabilization by applying different shifts on the averaging procedure and another methods for obtaining a correct synchronization.

4.1 Figures

Frontiers requires figures to be submitted individually, in the same order as they are referred to in the manuscript. Figures will then be automatically embedded at the bottom of the submitted manuscript. Kindly ensure that each table and figure is mentioned in the text and in numerical order. Figures must be of sufficient resolution for publication see here for examples and minimum requirements. Figures which are not according to the guidelines will cause substantial delay during the production process. Please see here for full figure guidelines. Cite figures with subfigures as figure 8B.

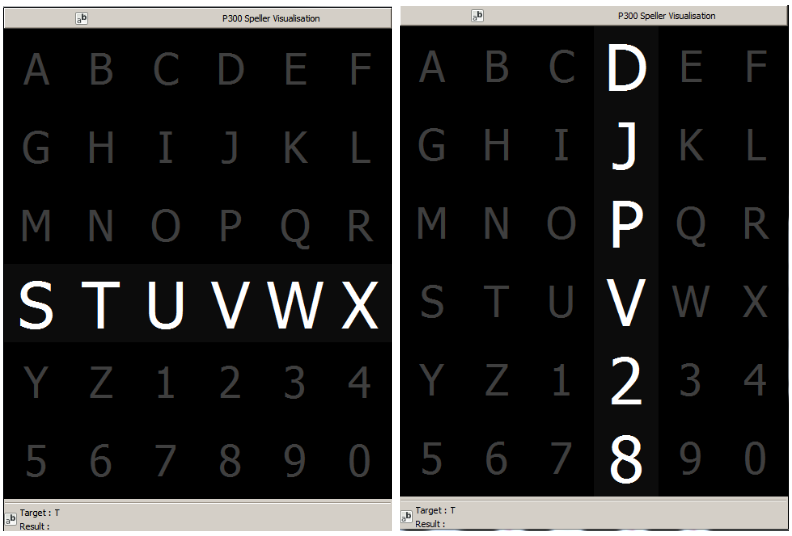

Figure 1. Example of the 6×6 Speller Matrix used in the study. Rows and columns flash intermittently in random permutations.

4.2 Tables

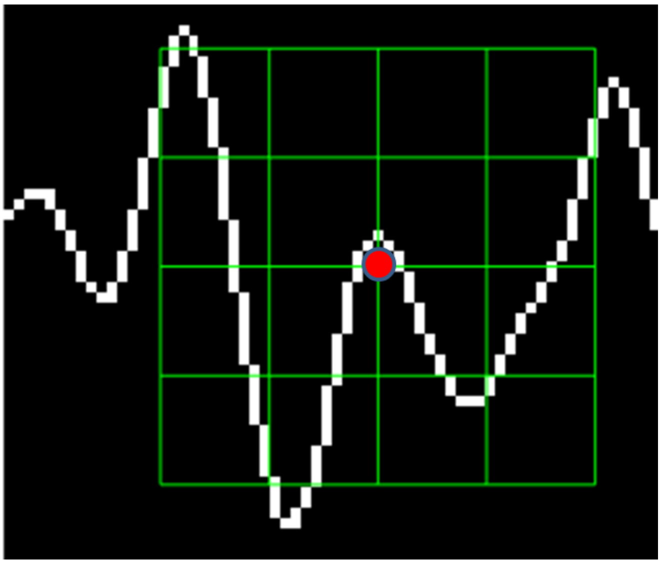

Figure 2. Example of a patch and a scheme of the orientation's histogram computation.Plot of the signal, a keypoint and the corresponding patch.

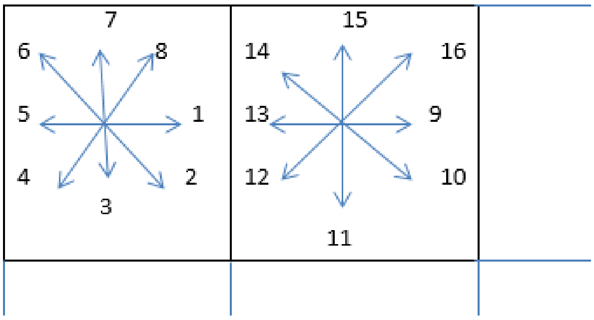

Figure 3. Example of a patch and a scheme of the orientation's histogram computation. Orientations on two blocks of the patch.

5 NOMENCLATURE

5.1 Resource Identification Initiative

To take part in the Resource Identification Initiative, please use the corresponding catalog number and RRID in your current manuscript. For more information about the project and for steps on how to search for an RRID, please click here.

6 ADDITIONAL REQUIREMENTS

For additional requirements for specific article types and further information please refer to Author Guidelines.

CONFLICT OF INTEREST STATEMENT

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

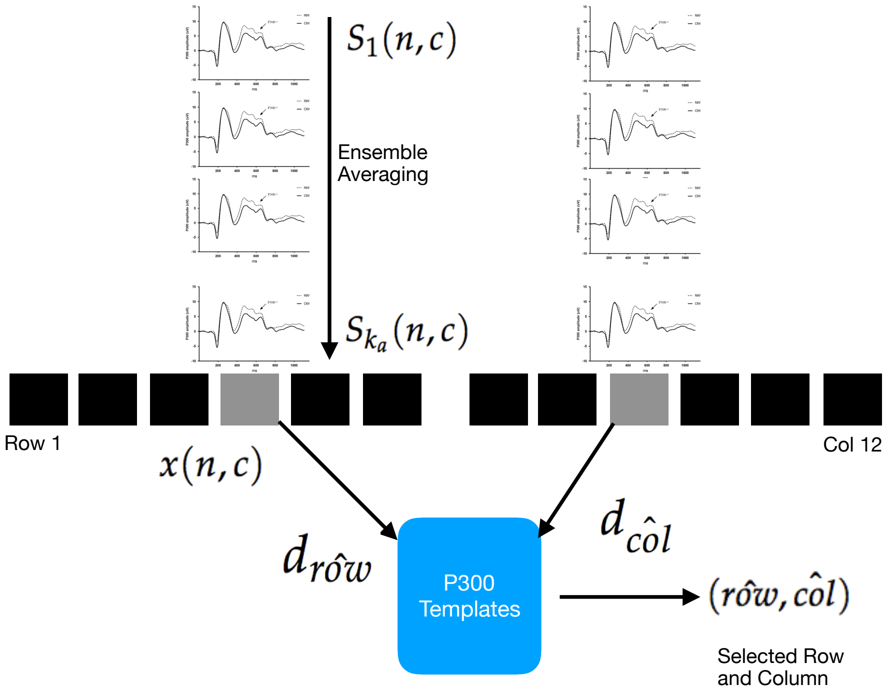

Figure 4. Single trial segments S_i are averaged for the 6 rows and 6 columns. From the averaged signal, the image of the signal plot is generated and each descriptor is computed. By comparing each descriptor against the set of templates, the P300 ERP can be detected, and finally the desired letter from the matrix can be inferred.

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 008-2014. 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%

AUTHOR CONTRIBUTIONS

This work is part of the PhD thesis of the First Author. The remaining authors contributed equally to the development of this method.

FUNDING

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

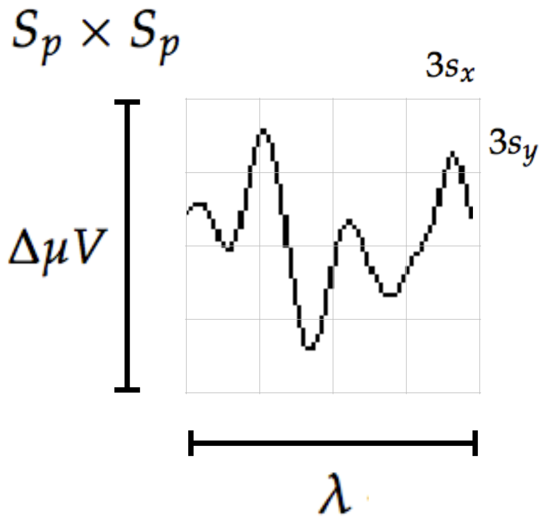

Figure 5. The scale of local patch is selected in order to capture the whole transient event. The size of the patch is $S_p \times S_p$ pixels. The vertical size consists of 4 blocks of size $3s_y$ pixels which is long enough as to contain the signal $\Delta\mu V$, the peak-to-peak amplitude of the transient event. The horizontal size includes 4 blocks of $3s_x$ and covers the entire duration in seconds of the transient signal event, λ.

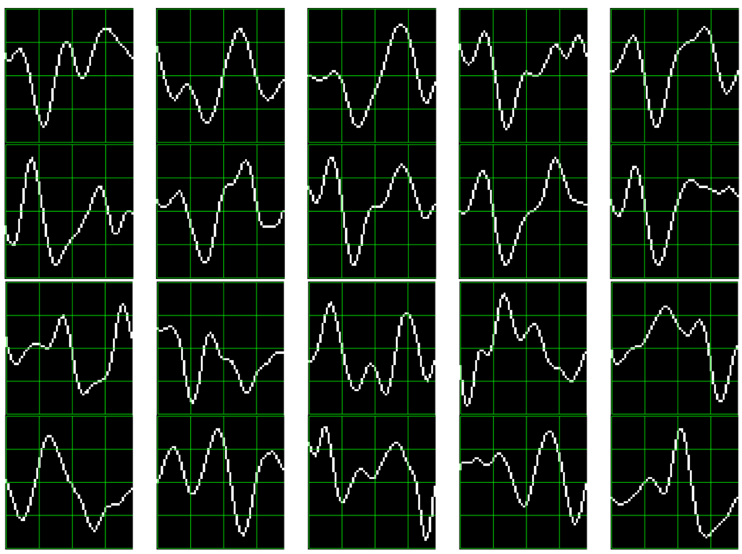

Figure 6. P300 template patches for subjects 8 and 3. As traditional done in neuroscience research, downward is positive.

SUPPLEMENTAL DATA

345 Supplementary Material should be uploaded separately on submission, if there are Supplementary Figures,
346 please include the caption in the same file as the figure. LaTeX Supplementary Material templates can be
347 found in the Frontiers LaTeX folder

Table 2. Percentage of correctly predicted letters while performing an offline BCI Simulation for the best performing channel for each healthy subject. 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%

REFERENCES

Alvarado-González, M., Garduño, E., Bribiesca, E., Yáñez-Suárez, O., and Medina-Bañuelos, V. (2016). P300 Detection Based on EEG Shape Features. *Computational and Mathematical Methods in Medicine* , 1–14doi:10.1155/2016/2029791

Berger, S., Schneider, G., Kochs, E., and Jordan, D. (2017). Permutation Entropy: Too Complex a Measure for EEG Time Series? *Entropy 2017, Vol. 19, Page 692* 19, 692. doi:10.3390/E19120692

Boiman, O., Shechtman, E., and Irani, M. (2008). In defense of nearest-neighbor based image classification. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR* doi:10.1109/CVPR.2008.4587598

Bresenham, J. E. (1965). Algorithm for computer control of a digital plotter. *IBM Systems Journal* 4, 25–30

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

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

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

Chavarriaga, R., Fried-Oken, M., Kleih, S., Lotte, F., and Scherer, R. (2017). Heading for new shores! Overcoming pitfalls in BCI design. *Brain-Computer Interfaces* 4, 60–73. doi:10.1080/2326263X.2016.1263916

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

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

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

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

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

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

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

Huggins, J. E., Alcaide-Aguirre, R. E., and Hill, K. (2016). Effects of text generation on P300 brain-computer interface performance. *Brain-Computer Interfaces* 3, 112–120. doi:10.1080/2326263X.2016.1203629

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

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

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

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

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

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

Ouyang, G., Hildebrandt, A., Sommer, W., and Zhou, C. (2017). Exploiting the intra-subject latency variability from single-trial event-related potentials in the P3 time range: A review and comparative evaluation of methods. *Neuroscience and Biobehavioral Reviews* 75, 1–21. doi:10.1016/j.neubiorev.2017.01.023

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

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

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

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

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

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

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

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

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

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

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

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

FIGURE CAPTIONS

Figure 7. Enter the caption for your figure here. Repeat as necessary for each of your figures

Figure 8. This is a figure with sub figures, (A) is one logo, (B) is a different logo.