

# Impact of the time segment analysis for P300 detection with spatial filtering

Hubert Cecotti, Ronald Phlypo, Bertrand Rivet, and Marco Congedo

GIPSA-Lab CNRS UMR 5216

961 rue de la Houille Blanche, Domaine universitaire - BP 46

38402 Saint Martin d'Hères, France

{hubert.cecotti,ronald.phlypo,bertrand.rivet}@gipsa-lab.inpg.fr

marco.congedo@gmail.com

<http://www.gipsa-lab.inpg.fr/>

**Abstract.** A Brain-Computer Interface (BCI) allows the direct communication between humans and computers by analyzing brain activity. The oddball paradigm allows detecting event-related potentials (ERPs), like the P300 wave, on targets selected by the user. While this paradigm provides the location of the P300 wave in the signal, its exact location remains a hypothesis and depends on the subject. This paper deals with the choice of the time segment for the signal analysis and its impact on the classification. A method for selecting the relevant part of the signal that contains the P300 wave is proposed. First, spatial filters are estimated for enhancing the signal. Second, a part of the enhanced P300 wave is selected based on its magnitude. This selection aims at providing an optimal start for the time window representing the P300 wave. Three window lengths are compared. We show that a window length of 500ms provides on average the best results, but the optimal window length should be set individually. The proposed technique has been validated on data recorded on 20 healthy subjects.

**Key words:** Brain-Computer Interface, P300 speller, Signal detection, Spatial filtering, Latency estimation

## 1 Introduction

A Brain-Computer Interface (BCI) allows the direct communication between human and computers by analyzing brain activity. This is thanks to particular brain responses to which a specific command can be assigned when they are detected. The detection of event related potentials (ERP) is one way for creating a BCI. A typical ERP based BCI is the P300 speller, which allows people to spell characters. Oddball paradigms are used in BCI to generate event-related potentials (ERPs), like the P300 wave, on targets selected by the user. These paradigms provide random visual stimuli that give a surprise effect to the subject. In this paper, we consider the classical P300 speller in which a  $6 \times 6$  matrix containing all the available characters is presented to the user on a computer screen [1, 2].

During the experiments, the user has to focus on the character s/he wants to spell. When the user focuses on a cell of the matrix, it is possible to detect a P300 (a positive deflection in voltage at a latency of about 300 ms relative to the stimuli onset in the EEG) time-locked to the onset of the cell intensification. To generate ERPs, the rows and columns are intensified randomly. Row/column intensifications are block randomized in 12 events (6 rows and 6 columns). The sets of 12 intensifications is repeated  $N_{epoch}$  times for each character. Therefore,  $2N_{epoch}$  possible P300 responses should be detected for the recognition of one character ( $N_{epoch}$  times for each row and column).

The first step in the P300 speller represents the detection of P300 waves in the electroencephalogram (EEG) signal. The second one combines several P300 responses for determining the right character to spell. The order of the intensifications in the paradigm during the experiment allows estimating when a P300 response is expected. In the character recognition step, the outputs of the P300 classification are combined to classify the main classes of the application (characters). In the oddball paradigm, a character is defined by a couple (row,column). The character is supposed to correspond to the intersection (row/column) of the accumulation of several P300 waves ( $2N_{epoch}$ ). The best accumulation of P300 waves for the horizontal (resp. vertical) flashing lights determines the row (resp. the column) of the desired character. For the P300 detection, we consider a signal  $X \in \mathbb{R}^{N_t \times N_s}$  where  $N_t$  is the number of sampling points in the time domain, and  $N_s$  is the number of electrodes that are used for the signal acquisition.  $N_t$  corresponds to a duration  $T$  of the recorded signal and a sampling frequency  $F_s$ . We denote by  $t_0$  the moment when a visual stimulus is presented to the zone on the screen observed by the subject. We define by  $\Delta t$  the duration between  $t_0$  and the moment where an extract of the signal  $X$  of duration  $N_1/F_s$  seconds is recorded. A challenge in P300 based BCI is to find that  $\Delta t$  and  $N_1$ , *i.e.* the latency and the duration of the P300 wave, which result in the best P300 detection. While it is possible to detect the P300 wave with a large  $N_1$ , finding the location of the P300 wave in relation to a subject can reduce the signal length to analyze and thus increase the signal-to-noise ratio by focusing only on the relevant part of the signal.

Several individual characteristics significantly influence the amplitude and the latency of the P300 wave. Indeed, the age, gender, intelligence and the personality are reported parameters that have an impact on the P300 wave [3]. Elderly have a higher latency and it increases linearly over the years. It is indeed reported that the latency increases of about 17ms every decade. In addition, the latency of the P300 wave is usually more important in the morning than in the evening [4]. A depression can also involve modifications in the amplitude and the latency of the P300 wave. Such psychological disorder is often present for disabled people with locked-in syndrome. In this paper, we address the problem of the window length choice for detecting the P300 wave. The starting point and the length of this window are usually set empirically and centered around the expected P300 wave occurrence at 300ms after the visual stimulus. The purpose of this paper is to identify the impact of  $\Delta t$  and  $N_1$  on the classification, *i.e.*, on

the P300 speller accuracy. This research is part of an ongoing effort in the BCI community to determine what parameter of a P300 speller should be adapted to each individual.

## 2 System overview

### 2.1 Spatial filtering

The EEG signal containing ERP is very noisy. One usual step for enhancing a particular brain response is to use spatial filters. Several methods for spatial filtering are described in the literature. The bipolar and Laplacian operators are usually used on sets of the electrodes for canceling the common nuisance signals [5]. Adaptive spatial filters obtained through Independent Component Analysis [6, 7] and Common Spatial Pattern [8, 9] are also commonly used. Spatial filters can also be embedded in the classification procedure as described in [10].

The spatial filtering method that is considered in this paper is based on the xDAWN algorithm [11]. This method assumes two main hypotheses:

- There exists a typical response synchronized with the target stimuli superimposed on an evoked response to all the stimuli (target and non-target). This hypothesis assumes the presence of a P300 wave only after the flashing light corresponding to the target on the screen. This hypothesis is common to P300 classifiers. Nevertheless, we can point out the relative confidence of the ground truth for training the classifier. In addition, the optimal location of the P300 wave may be difficult to identify. The responses are not always correlated in time to certain stimuli.
- The evoked responses to target stimuli could be enhanced by spatial filtering. This hypothesis is validated by several previous works that proved the interest of enhancing the input signal. The P300 is a spatially stationary waveform that has origins different from the background, *i.e.*, the current ongoing EEG.

We consider an algebraic model of the recorded signals  $X$  that is composed of three terms: the P300 responses ( $D_1A_1$ ), a response common to all stimuli, targets and non-targets confound ( $D_2A_2$ ) and the residual noise ( $H$ )

$$X = D_1A_1 + D_2A_2 + H. \quad (1)$$

where  $X \in \mathbb{R}^{N_t \times N_s}$ . The choice of  $N_t$  is one of the problems addressed in this work.  $N_t$  should be large enough to contain the P300 wave but ideally, it should not contain the noise parts of the signal.  $D_1$  and  $D_2$  are two real Toeplitz matrices of size  $N_t \times N_1$  and  $N_t \times N_2$  respectively.  $D_1$  has its first column elements set to zero except for those that correspond to a target onset shifted by  $\Delta t$ , which are represented with a value equal to one. For  $D_2$ , its first column elements are set to zero except for those that correspond to stimuli onset.  $N_1$  and  $N_2$  are the number of sampling points representing the target (the P300 response) and superimposed evoked potentials, respectively.  $H$  is a real matrix of size  $N_t \times N_s$ .

The purpose of applying spatial filters  $U \in \mathbb{R}^{N_s \times N_f}$  is to enhance the signal to signal-plus-noise ratio (SSNR) of the enhanced P300 responses ( $D_1 A_1 U$ ), where  $N_f$  is the number of spatial filters

$$XU = D_1 A_1 U + D_2 A_2 U + HU. \quad (2)$$

We define the SSNR in relation to the spatial filters by:

$$\text{SSNR}(U) = \frac{\text{Tr}(U^T \hat{A}_1^T D_1^T D_1 \hat{A}_1 U)}{\text{Tr}(U^T X^T X U)} \quad (3)$$

where  $\hat{A}_1$  corresponds to the least mean square estimation of  $A_1$  :

$$\hat{A} = \begin{bmatrix} \hat{A}_1 \\ \hat{A}_2 \end{bmatrix} = ([D_1; D_2]^T [D_1; D_2])^{-1} [D_1; D_2]^T X \quad (4)$$

where  $[D_1; D_2]$  is a matrix of size  $N_t \times (N_1 + N_2)$  obtained by concatenation of  $D_1$  and  $D_2$ .

The SSNR is maximized by:

$$\hat{U} = \underset{U}{\text{argmax}} \text{SSNR}(U). \quad (5)$$

In the definition of the SSNR, we replace  $\hat{A}_1$  by  $B_1^T X$  where  $B_1^T$  is a part of the least mean squares estimation Eq. (4). Then, we apply two QR decompositions yielding  $D_1 = Q_1 R_1$  and  $X = Q_x R_x$ . Finally, one can express Eq. (3) as:

$$\text{SSNR}(V) = \frac{\text{Tr}(V^T (Q_x^T B_1 R_1^T R_1 B_1^T Q_x) V)}{\text{Tr}(V^T V)}, \quad (6)$$

where  $V = R_x U$ .  $V$  is therefore obtained from the Rayleigh quotient, whose solution is the concatenation of  $N_f$  eigenvectors associated with the  $N_f$  largest eigenvalues of  $Q_x^T B_1 R_1^T R_1 B_1^T Q_x$  [12]. These vectors are estimated thanks to a singular value decomposition (SVD) of  $R_1 B_1^T Q_x = \Phi \Lambda \Psi^T$ ,  $\Phi$  and  $\Psi$  being two orthogonal matrices and  $\Lambda$  being a diagonal matrix with nonnegative diagonal elements in decreasing order. The solution of Eq. (5) provides the spatial filters. They are ordered in decreasing order by relevance impact. Finally, the enhanced signal  $\hat{X}$  can be estimated:

$$\hat{U} = R_x^{-1} \Psi \quad \hat{X} = X \hat{U}. \quad (7)$$

## 2.2 Classifier

The input of the classifier for the P300 detection corresponds to the four first components of the enhanced signal ( $N_f = 4$ ). The Bayesian linear discriminant analysis (BLDA) classifier is considered for the detection of the P300 wave [13]. This classifier is fast to train and does not require hyperparameters to adjust [14]. It finds a discriminant vector  $w$  such that the following expression is minimized:

$$|w^T p - O(c)| \quad (8)$$

where  $p$  belongs to the class  $c$  and  $O(c)$  represents the associated scalar of a class  $c$ . For the class representing the P300 (resp. non P300),  $O(c) = 1$  (resp.  $O(c) = 0$ ). The vector  $p$  is obtained by the concatenation of the different time-course signals across the different spatial filters.

### 3 Latency estimation

The latency is estimated thanks to the enhanced P300 response defined by  $\widehat{A_1U} = \widehat{A_1\check{U}}$ . The first component of  $\widehat{A_1U}$  represents by interpretation of Eq. (3) the application of the filter that maximizes the SSNR on the wave form of the P300. Therefore,  $\widehat{A_1U}$  represents an enhanced wave form of the P300 wave. Thus, only the first component of  $\widehat{A_1U}$  is taken into account for estimating the P300 latency. We consider the vector  $(\zeta_1, \dots, \zeta_{N_1})$  that contains the magnitude of the P300 wave enhanced by the first spatial filter, where  $\zeta = |\widehat{A_1U}(:, 1)|$ .

The  $p$  highest values of  $\zeta$  are chosen as a subset of the best observations that describe the P300 wave. In the experiments  $p$  is chosen as  $p = N_1/2$  or  $p = N_1/4$ , which represents the best 500ms or 250ms part of the signal to consider. The selected signal parts are denoted by  $\hat{\zeta} = (\zeta_{j_1}, \dots, \zeta_{j_p})$ .  $\zeta_{j_1}$  corresponds to the first sampling point to consider after a visual stimulus. Therefore, we have that  $\Delta t = \arg(\zeta_{j_1})/F_s$  and  $N_t = \arg(\zeta_{j_p}) - \arg(\zeta_{j_1})$ .  $\arg(\cdot)$  provides the index of the selected  $\zeta$  value. The newly selected window is then used for creating new spatial filters. The selected signal is estimated again with the procedure described before. These steps are iterated until the selected signal parts converge.

### 4 Experiments

The EEG signal was recorded on 20 healthy subjects (average age= 26 years, standard deviation = 5.7 years, 17 males, 7 females). Subjects were wearing an EEG cap with 32 electrodes [15]. For the further analysis, we consider also a subset of eight electrodes:  $F_Z, C_Z, P_7, P_3, P_Z, P_4, P_8, O_Z$ . Such electrode subset represents an acceptable comparison for what could be a commercial BCI. For each subject, two sessions are recorded. The first session is dedicated to the training part of the classifier. 50 characters were used for each subject. In the second session, the test session, 60 characters were spelled by each subject. Before the classification steps, the signal initially acquired with a sampling rate of 100Hz was first filtered by a bandpass filter with cut-off frequencies at 1Hz and 20Hz. Finally, the signals were normalized independently for each sensor and for each character as to have a zero mean and standard deviation equal to one.

### 5 Results

Figure 1 presents the speller accuracy (in %) across the 20 subjects and for four different ways to select the window for classifying the P300 waves. The speller accuracy corresponds to the character accuracy after combination of the

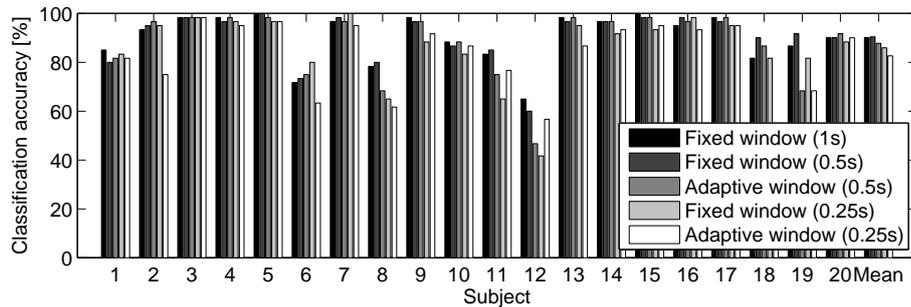
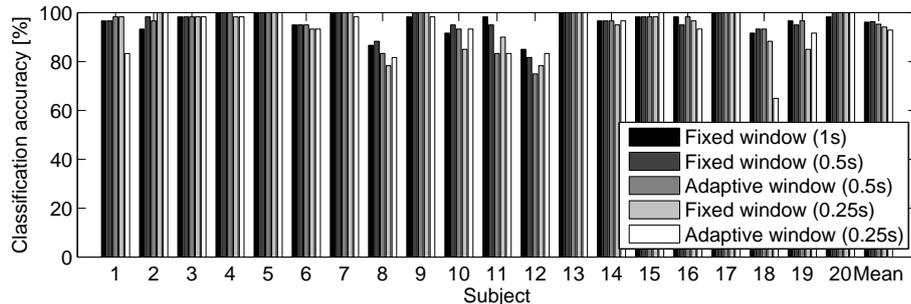
different P300 waves. Three fixed window lengths are considered: 1s, 500ms, and 250ms. With a window of 1s, the classifier input signal is taken directly after the start of a visual stimulus. The windows of size 500ms and 250ms are centered on 300ms after the start of a visual stimulus. The adaptive windows are defined in relation to the latency estimation described in section 3. For a set of eight electrodes, the average accuracy of the speller across the 20 subjects is 90.17%, 90.42% and 85.92% with a fixed window length of 1s, 500ms and 250ms, respectively. With an adaptive window, the average accuracy is 87.75% and 82.66% respectively for 500ms and 250ms. When all sensors are considered, the average accuracy is 96.17%, 96.33% and 94.17% with a fixed window length of 1s, 500ms and 250ms, respectively; with an adaptive window, we obtain 95.33% and 92.92% for 500 and 250ms. For the non-adaptive procedure, we performed 1-way repeated measure ANOVA with random factor “window length” (three levels: 1s, 500ms and 250ms), separately for 32 and 8 electrodes. We have used Geisser-Greenhouse correction for covariance circularity violation and we have performed Bonferroni-corrected all pair-wise post-hoc comparisons by paired t-tests. For the adaptive procedure we performed paired t-tests on the factor “window length” comparing the means obtained using 500ms and 250ms.

The ANOVA for the non-adaptive procedure was significant, both for the 32 electrodes configuration ( $F(2)=4.57$ ,  $p=0.045$ ) and 8 electrodes configuration ( $F(2)=6.72$ ,  $p=0.018$ ). Table 1 presents pairwise comparison between the different window lengths with the non-adaptive method. The t-test for the adaptive procedure was significant only using the 8 electrodes configuration ( $t(19)=2.50$ ,  $p=0.021$ ).

# Electrodes	1s vs 500ms		1s vs 250ms		500ms vs 250ms	
	t(19)	p	t(19)	p	t(19)	p
8	X		2.58	0.053	5.06	0.021
32	X		2.35	0.019	2.87	0.009

**Table 1.** Pairwise comparison of means. X means that the difference is not significant.

These results suggest that using the non-adaptive procedure, a time segment of 500ms is the minimum length to keep a good accuracy, by both using the 8 and 32 electrodes configuration. With the adaptive procedure, the above result is only true when using the 8 electrode configuration. On the other hand, there is no difference between 500ms and 250ms window length by considering all available electrodes. In summary, the adaptive procedure allows us to reduce the window length down to 250ms without significantly affecting the accuracy of the P300 speller, however it appears that a large number of electrodes must be used for this purpose. This proves that not only the classifier but also the normalization of the input data like the choice of the input vector size could ideally be set in relation to each subject. Alternatively, one may also chose the electrode subset adaptively. On average, the choice of a predefined window length can be a good solution.

(a) 8 electrodes ( $F_Z, C_Z, P_7, P_3, P_Z, P_4, P_8, O_Z$ ) [15]

(b) 32 electrodes

**Fig. 1.** Accuracy of the P300 speller for the different methods. Fig. 1(a): with 8 electrodes, Fig. 1(b): with 32 electrodes

## 6 Conclusion

Current state-of-the-art methods for the classification of the P300 wave do neither exploit the latency nor the length of the P300 wave [10]. Feature selection strategies are predominantly based on an optimisation in the spatial domain, *i.e.*, electrodes selection. According to our results, the choice of the length and the starting point can have an impact on the classification rate of the speller for several individuals, particularly for persons who obtain relatively bad results with the speller. The choice of the window length and the estimated location of the P300 peak could jeopardize the BCI literacy for some subjects. A method for the estimation of the P300 latency has been proposed. It includes a step for enhancing the P300 wave and a step for selecting the best part of the signal for estimating the P300 wave. These preliminary but promising results provide hints on the need of models for solving this problem individually. Models like hidden Markov Models could help defining the temporal elasticity and elusiveness of the P300 wave over time. While the adaptive windows can provide a way for increasing the efficiency of the system, the signal variability over sessions could be an obstacle for such adaptivity. Further works could deal with the dynamics of the P300 wave over time and for each subject.

**Acknowledgments.** This work was funded by the ANR (French National Research Agency) RoBIK Project (Robust BCI Keyboard). The authors would like to thank Dr. J. Mattout, Dr. E. Maby and Prof. O. Bertrand from the INSERM and the *Institut Fédératif des Neurosciences* of Lyon, France for collecting and sharing EEG data.

## References

1. Farwell, L., Donchin, E.: Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr. Clin. Neurophysiol.* **70** (1988) 510–523
2. Donchin, E., Spencer, K.M., Wijesinghe, R.: Assessing the speed of a P300-based brain-computer interface. *IEEE Trans. Neural Sys. Rehab. Eng.* **8**(2) (2000) 174–179
3. Hansenne, M.: Le potentiel évoqué cognitif P300 (ii): variabilité interindividuelle et application clinique en psychopathologie. *Neurophysiol Clin* **30** (2000) 211–231
4. Querrioux-Coulombier, G., Gil, R.: Variations journalières des composantes des potentiels évoqués cognitifs. *Neurophysiol Clin* **21** (1991) 75–84
5. Müller-Putz, G.R., Scherer, R., Brauneis, C., Pfurtscheller, G.: Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. *Journal of Neural Engineering* **2**(1) (2005) 123–130
6. Comon, P.: Independent component analysis, a new concept ? *Signal Processing* **36**(3) (1994) 287–314
7. Xu, N., Gao, X., Hong, B., Miao, X., Gao, S., Yang, F.: BCI competition 2003–data set iib: enhancing P300 wave detection using ICA-based subspace projections for BCI applications. *IEEE Trans Biomed Eng* **51**(6) (2004) 1067–1072
8. Blankertz, B., Kawanabe, M., Tomioka, R., Hohlefeld, F., Nikulin, V., Müller, K.R.: Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing. *Advances in Neural Information Processing Systems* **20** (2008)
9. Brunner, C., Naeem, M., Leeb, R., Graimann, B., Pfurtscheller, G.: Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis. *Pattern Recognition Letters* **28**(8) (2007) 957–964
10. Cecotti, H., Gräser, A.: Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Trans. Pattern Analysis and Machine Intelligence* (2010)
11. Rivet, B., Souloumiac, A., Attina, V., Gibert, G.: xDAWN algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Trans Biomed Eng.* **56**(8)
12. Golub, G.H., Van Loan, C.F.: *Matrix computations*. 3rd ed. Johns Hopkins University Press (1996)
13. MacKay, D.J.C.: Bayesian interpolation. *Neural Comput.* **4**(3) (1992) 415–447
14. Hoffmann, U., Vesin, J., Diserens, K., Ebrahimi, T.: An efficient P300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods* **167**(1) (2008) 115–125
15. Chatrian, G.E., Lettich, E., Nelson, P.L.: Ten percent electrode system for topographic studies of spontaneous and evoked eeg activity. *Am J EEG Technol* **25** (1985) 83–92