Reducing Calibration Time for the P300 Brain-Computer Interface Speller

H. Cecotti [1], G. Saint-Ellier [1], B. Rivet [1], M. Perrin [2], E. Maby [2], J. Mattout [2]
[1] GIPSA-lab CNRS UMR 5216
Grenoble Universities
Saint Martin d'Hères, 38402, France
[2] INSERM, U821, Lyon, 69500, France
Institut Fédératif des Neurosciences, Lyon, 69000, France
Université Lyon 1, Lyon, 69000, France

Abstract

With a Brain-Computer Interface (BCI), it is nowadays possible to achieve a direct pathway between the brain and computers thanks to the analysis of some particular brain activities. The detection of even-related potentials, like the P300, provides a way to create BCIs. The generation of the P300 wave is achieved with the oddball paradigm, which allows detecting targets selected by the user on a screen. The P300-Speller is based on this principle. The detection of the P300 requires efficient signal processing and machine learning techniques. Thus, a calibration step is needed for training the models. However, the duration of this calibration can be a drawback. We propose to evaluate the optimal number of characters that should be spelt in order to provide a working system with a minimum calibration duration. The evaluation has been tested on data recorded on 20 healthy subjects. It is possible to spell only seven symbols during the calibration to reach an initialized system with an average accuracy of at least 80%.

Keywords

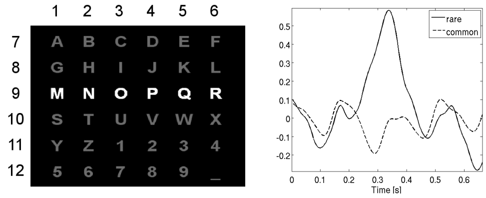
Brain-Computer Interface, P300 Speller, Spatial filtering, Signal detection, Calibration duration.

1. Introduction

The goal of a Brain-Computer Interface (BCI) is to provide a direct communication pathway between the brain and computers. The performances of a BCI are both related to the performance of the signal processing techniques that are used for assigning some EEG signal to a command and to the possibility of the user to adapt him/herself to the system over time. Indeed, advanced machine learning and signal processing and classifier techniques have been widely used for improving BCIs [1, 12, 18]. In spite of these recent improvements in the BCI community, several obstacles remain to fully transfer laboratory demonstrator BCIs to real commercial/clinical applications. Whereas tuning the different parameters of a BCI in relation to a specific individual can improve the performance of the system, this procedure is time consuming. For training a classifier, a large database containing labeled EEG signal is often needed. To obtain these data, a training session is required where the subject has to follow a specific protocol.

We can distinguish several types of potential BCI users. First, BCIs are usually dedicated to persons with severe motor disabilities who are unable to communicate through any other means. For these persons, a BCI is the only way to communicate and the main challenge is to have a functional BCI. While reducing the time of the training session can be a certain advantage, it is not the main objective. For people who suffer from severe disabilities, like the locked-in syndrome, having a working BCI can still be challenging. Second, BCIs can be used by persons with disabilities or elderly persons who need a device to facilitate their daily life. In this situation, current challenges in BCI are to get a BCI to work outside of the laboratory, in real condition. For this group of potential users, the time dedicated to the training session and the expected performance shall be well balanced. Third, BCIs can be used by healthy people as an alternative device for controlling video games or other applications. In this case, a BCI should ideally be flawless: the calibration step shall be as short as possible. Indeed, once the excitement of controlling something with the mind has passed, the usability stays an important satisfaction criterion. Some potential BCI users are indeed highly demanding in term of performance and usability. The past decade has shown that BCIs can effectively work, they should now ideally become plug'n'play [4].

The calibration step can be a drawback for some potential BCI users. With a lot of available data, it is possible to train classifiers and improve the reliability of the detection procedure. However, one may wonder when the signal processing steps show their limit and it becomes useless to pursue the training procedure. Therefore, a challenge is to determine the ideal training session duration for the personalization of a BCI. On one hand, a too short training session would not



(a) P300 Speller - Graphical User Interface (b) P300 evoked potentials

Figure 1. P300-BCI Speller. Fig. 1: screen display, Fig. 1(b): average P300 response on one sensor located on Cz

be enough and would involve the creation of a poorly trained classifier. On the other hand a too long training session would frustrate the user, decrease her/his motivation. The calibration step shall not be considered as the final state of the system adaptation but only one first step for providing an efficient system. For instance, classifiers can be trained in an unsupervised way over time in such fashion that the system adaptability stays invisible to the user. In addition, the user can improve his performance by finding appropriate ways to adapt his/her behavior to the system or to some feedbacks [15, 16, 17].

In this paper, we address the problem of the required time that is needed to train a classifier and provide efficient results for the P300-Speller. In the P300-Speller, the duration of the training session can be evaluated as a number of characters to spell. The duration for spelling a character is indeed determined by initial setting of the speller like the inter-stimuli interval. This research is part of an ongoing effort in the French BCI community to reduce the time dedicated to the calibration step of the P300-Speller.

The paper is organized as follows: each step of the P300 detection, from the spatial filtering methods till the classification, is described in the second section. The experimental protocol is then detailed in the third section. Finally, the results of the offline classification are presented and discussed in the last section.

2. System overview

2.1 P300-Speller

The P300-Speller allows people to select symbols where each symbol is depicted in a cell of a matrix on a computer screen. The P300-Speller is one of the oldest BCI paradigm. It is based on the oddball paradigm to generate event-related potentials (ERPs), like the P300, on targets selected by the user. This paradigm provides random visual stimuli that shall give a surprise effect to the subject. The classical P300 speller is considered thereafter: a 6×6 matrix that contains all the available characters is presented to the subject on a computer screen [10, 9]. 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, *i.e.*, the character the person wants to spell, it is possible to detect a P300. It corresponds to a positive deflection in voltage at a latency of about 300 ms relative to the stimuli onset in the electroencephalogram (EEG) signal. This deflection is time-locked to the onset of the cell intensification. The rows and columns are intensified randomly to generate ERPs. Row/column intensifications are block randomized in 12 events (6 rows and 6 columns). This set of 12 intensifications is repeated N_{epoch} times for each character. Hence, $2N_{epoch}$ possible P300 responses should be detected for the recognition of one character.

The P300-Speller is decomposed into several steps. The most important step in the P300 speller represents the detection of P300 waves in the EEG. The next step 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). This 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.

2.2 Spatial filtering

The EEG signal containing ERPs 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 [19]. Adaptive spatial filters obtained through ICA [8, 22] and Common Spatial Pattern (CSP) [2, 3] are also commonly used. Spatial filters can also be determined during the training of the classifier [5].

The spatial filtering method that is considered in this paper is based on the xDAWN algorithm [6, 21]. This method allows estimating a set of spatial filters that optimize the signal to signal-plus-noise (SSNR) ratio by considering Rayleigh quotients. This technique is based on two 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. The optimal location of the P300 wave may be difficult to identify. The responses are not always correlated 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 EEG signals X. X is composed of three terms: the P300 responses ($D_1 A_1$), a response common to the P300 and non P300 waves ($D_2 A_2$) and the residual noise (N)

$$X \quad = \quad D_1 A_1 + D_2 A_2 + H. \qquad (1)$$

where $X \in \mathbb{R}^{N_t \times N_s}$. N_t represents the window length that contains the P300. 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, which are set 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 SSNR of the enhanced P300 responses ($D_1 A_1 U$), where N_f is the number of spatial filters

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

2.3 Classifier

The input of the classifier for the P300 detection corresponds to the four first components of the enhanced signal ($N_f = 10$). The Bayesian linear discriminant analysis (BLDA) classifier is considered for the detection of the P300 wave [11, 14]. It finds a discriminant vector w such that the following expression is minimized:

$$|w^T p - O(c)| \qquad (3)$$

where p, the feature vector representing the filtered signal, 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. $O(c) = 0$).

2.4 Evaluation of the training session

We propose to evaluate the optimal number of characters that is needed during the training session to provide some desired performance. We consider a training session of N_{symb}^{max} characters. We define N_{symb}^{i} as the i^{th} character to spell, $1 \leq i \leq N_{symb}^{max}$. The optimal number of characters, N_{symb}^{j} shall be the smallest j respecting one of these following rules:

$$|Acc(N_{symb}^{j}) - Acc(N_{symb}^{j+1})| \quad \leq \quad \alpha \qquad (4)$$

$$Acc(N_{symb}^{j}) \quad \geq \quad \tau \qquad (5)$$

where $1 \leq j \leq N_{symb}^{max}$. $Acc(N_{symb}^{j})$ represents the accuracy (in %) of the speller when the training session is limited to the i first characters. τ represents the desired accuracy to reach. The advantage of considering one more character for training is determined by α. During the further signal analysis, we consider $\alpha = 1\%$.

3. Experiments

The EEG signal was recorded on 20 healthy subjects (average age= 26 years, standard deviation= 5.7 years, 13 males, 7 females). Subjects were wearing an EEG cap with 32 electrodes [7]. The OpenViBE framework was used to record EEG and perform the different experiments [13, 20]. For each subject, we consider the recorded signals of three sessions. The first one is dedicated to the training part of the classifier, with an inter-stimuli interval of 170ms and 10 repetitions. 50 characters were used for each subject. In the second and third session, 60 characters were spelled by each subject. The second session is used for estimating the optimal number of characters that should be spelt in the first session. The third session is for evaluating the relevance of the approach. 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 10Hz. 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.

4. Results

For the evaluation of the accuracy in relation to the number of characters, the characters were taken in the chronological order during the training, *i.e.*, when i characters are selected, they correspond to the first i characters in the training session. Some evaluations were first performed by comparing the order of the selected characters, *i.e.*, if we obtain the same average accuracy by selecting the first five or the last five characters. An ANOVA test proved that there exists no statistical difference between the order in which the characters are selected.

Two criteria were tested for evaluating the optimal number of characters in the training session. They correspond to two desired accuracies of the P300-Speller: $\tau = 80\%$ and $\tau = 90\%$. In Figure 2, the average recognition rate of the P300 speller across the 20 subjects is presented in relation to the number of characters used in the training session. According to the desired average accuracy, it is possible to consider only 7 and 15 characters to obtain in average an accuracy of 80% and 90%.

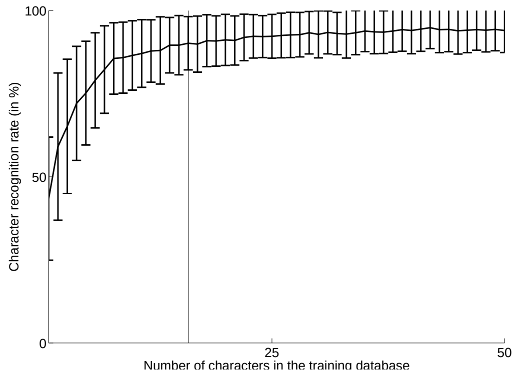


Figure 2. Average speller accuracy in relation to the number of characters used in the training session.

The evolution of the speller accuracy over time is detailed for each subject in Figure 3, from left to right, then top to bottom. The accuracy over the number of characters is low-pass filtered. With only five characters in the training session, Subject 13 is able to achieve good results. However, some subjects could neither reach the desired performance nor stabilize their performance. For these subjects, the more characters there are, the best the accuracy becomes: the speller accuracy never reaches the desired performance but the speller accuracy still improves over time. It is the case of subjects 8, 10, 12, and 18.

Figure 4 presents the recognition rate of the P300-Speller for the two chosen criteria, compared to the case where the whole training database (50 characters) would be used. The mean recognition rate across the 20 subjects is 79.58%, 85.00% and 88.92% when the training session is limited to 7, 15 and 50 characters respectively. While the accuracy naturally decreases with a limited number of characters for training, the performance stays decent.

5. Conclusion

Helping people with disabilities should and will stay a major focus in the BCI community. Current research directions are nevertheless moving beyond the restricted user group of disabled people for a broader audience. Such research directions imply a better focus on interfaces and improvement on the usability in non-invasive BCIs. In this study, we have estimated the number of characters that should be spelt in order to provide a good initialization of a P300-Speller. Such initialization can allow a subject to start using a P300-Speller without frustration. It provides a good base for online adaptations and improvements. Further works will deal with more evaluations with different type of sessions spread over time to qualify the keypoints when the system should be adapted online.

Acknowledgments.

This work has been supported by French National Research Agency (ANR) through TecSan program (project RoBIK ANR-09-TECS-013) and DEFIS program (project Co-Adapt nřANR-09-EMER-002).

References

[1] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch. A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *Journal of Neural Engineering*, 4:32–57, 2007.

[2] B. Blankertz, M. Kawanabe, R. Tomioka, F. Hohlefeld, V. Nikulin, and K.-R. Müller. Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing. *Advances in Neural Information Processing Systems*, 20, 2008.

[3] C. Brunner, M. Naeem, R. Leeb, B. Graimann, and G. Pfurtscheller. Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis. *Pattern Recognition Letters*, 28(8):957–964, 2007.

[4] H. Cecotti. A self-paced and calibration-less SSVEP based brain-computer interface speller. *IEEE Trans. on Neural Systems and Rehab. Eng.*, 18:127–133, 2010.

[5] H. Cecotti and A. Gräser. Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2010.

[6] H. Cecotti, B. Rivet, M. Congedo, C. Jutten, O. Bertrand, E. Maby, and J. Mattout. Suboptimal sensor subset evaluation in a p300 brain-computer interface. *Proc. of the 18th European Signal Processing Conference*, 2010.

[7] G. E. Chatrian, E. Lettich, and P. L. Nelson. Ten percent electrode system for topographic studies of spontaneous and evoked eeg activity. *Am J EEG Technol*, 25:83–92, 1985.

[8] P. Comon. Independent component analysis, a new concept ? *Signal Processing*, 36(3):287–314, 1994.

[9] E. Donchin, K. M. Spencer, and R. Wijesinghe. Assessing the speed of a P300-based brain-computer interface. *IEEE Trans. Neural Sys. Rehab. Eng.*, 8(2):174–179, 2000.

[10] L. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr. Clin. Neurophysiol.*, 70:510–523, 1988.

[11] U. Hoffmann, J. Vesin, K. Diserens, and T. Ebrahimi. An efficient P300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods*, 167(1):115–125, 2008.

[12] F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 4:R1–R13, 2007.

[13] E. Maby, G. Gibert, P.-E. Aguera, M. Perrin, O. Bertrand, and J. Mattout. The OpenViBE P300-Speller scenario: a thorough online evaluation. *In Human Brain Mapping Conference*, 2010.

[14] D. J. C. MacKay. Bayesian interpolation. *Neural Comput.*, 4(3):415–447, 1992.

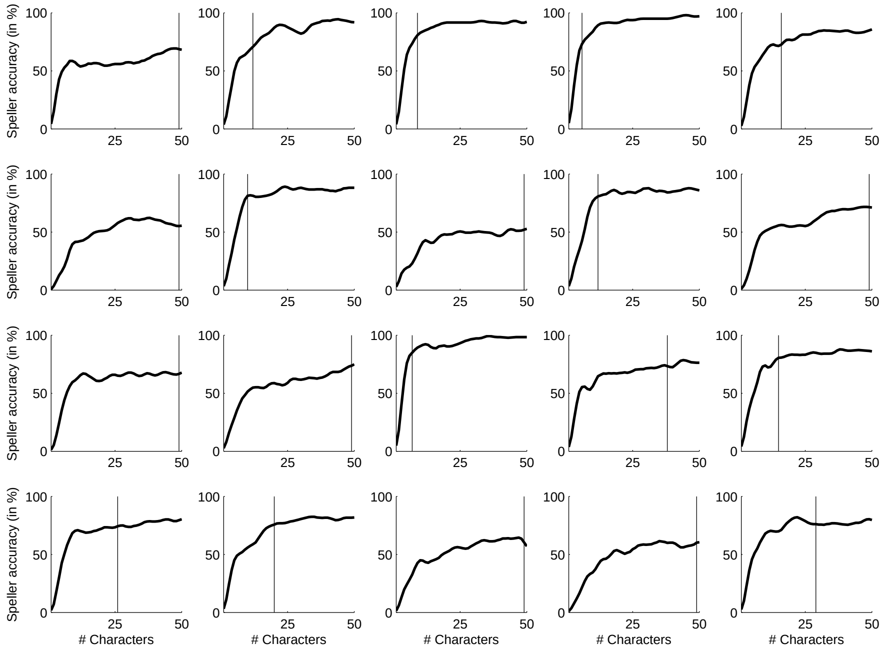


Figure 3. Speller accuracy in relation to the number of characters used in the training session, for each subject.

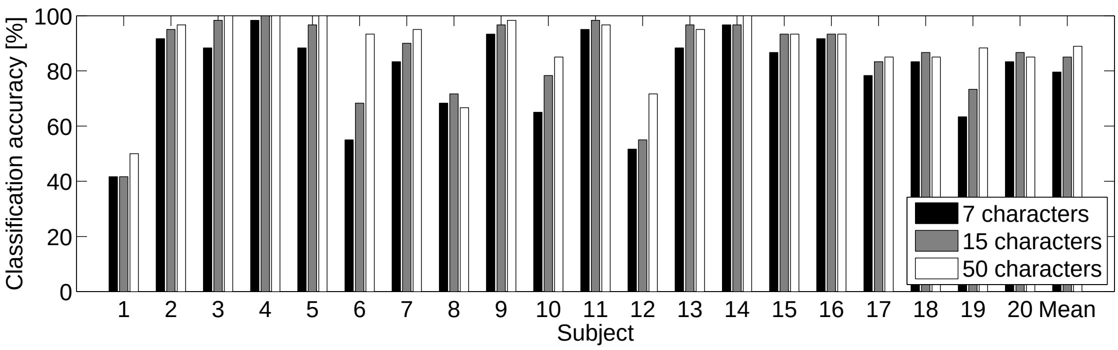


Figure 4. Speller accuracy on the test database for 20 subjects.

[15] P. Martinez, H. Bakardjian, and A. Cichocki. Fully online multicommand brain-computer interface with visual neurofeedback using SSVEP paradigm. *Computational Intelligence and Neuroscience*, 2007.

[16] D. J. McFarland, D. J. Krusienski, and J. R. Wolpaw. Brain-computer interface signal processing at the wadsworth center: mu and sensorimotor beta rhythms. *Progress in Brain Research, chapt. 26*, 159:411–419, 2006.

[17] M. M. Moore. Real-world applications for brain-computer interface technology. *IEEE Trans Neural Systems and Rehabilitation Engineering*, 11(2):162–165, 2003.

[18] K.-R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and B. Blankertz. Machine learning for real-time single-trial EEG-analysis: From brain-computer interfacing to mental state monitoring. *J Neurosci Methods*, 167(1):82–90, 2008.

[19] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller. Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. *Journal of Neural Engineering*, 2(1):123–130, 2005.

[20] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lecuyer. Openvibe: An open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence Teleoperators & Virtual Environments*, 19(1):35–53, 2010.

[21] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert. xDAWN algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Trans Biomed Eng.*, 56(8), 2009.

[22] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang. BCI competition 2003–data set iib: enhancing P300 wave detection using ICA-based subspace projections for BCI applications. *IEEE Trans Biomed Eng*, 51(6):1067–1072, 2004.