Single-Trial P300 Detection in Healthy and ALS Subjects by Means of a Genetic Algorithm

B. Dal Seno[1], M. Matteucci[1], L. Mainardi[2], F. Piccione[3], S. Silvoni[3]

[1]Department of Electronics and Information, IIT-Unit, Politecnico di Milano, Italy
[2]Department of Bioengineering, IIT-Unit, Politecnico di Milano, Italy
[3]IRCSS San Camillo, Venezia, Italy

bernardo.dalseno@polimi.it

Abstract

P300 is a potential widely used in brain-computer interfaces (*BCI*), as P300 is an innate response that does not require training on the part of the user. In the literature several classification algorithms have been used (e.g., Linear Discriminant Analysis, Stepwise Discriminant Analysis, Support Vector Machines), and, typically, first the P300 relevant features are extracted from the EEG signal, then they are fed into the classifier. From this, it becomes clear that feature extraction is the key point, and doing it by hand can be at the same time cumbersome and suboptimal. In this paper, we face the issue of automatic feature extraction by using a genetic algorithm (*GA*) able to retrieve the relevant aspects of the signal to be classified in an automatic fashion. We do not use GA for feature selection or classifier optimization; instead, we learn directly from the signal which are the features we should use in our classifier. The approach has been used for single-sweep classification with a logistic classifier on a group of 10 subjects affected by ALS (amyotrophic lateral sclerosis), hospitalized in the S. Camillo structure, and a group of 4 healthy subjects, voluntarily participating to the study. Results are promising, reaching up to 95% accuracy for some subjects; moreover, the features extracted by the GA turn out to be related to the P300 activity and can provide insights about the most interesting regions and time to classify P300s.

1 Introduction

A brain-computer interface (BCI) [1] is an interface that does not entail muscle movements, but it bypasses any muscle or nerve mediation and connects a computer directly with the brain by picking up signals generated by the brain activity.

In this study, we focus on the *P300* [2], an event-related potential (ERP) that can be recorded through an electroencephalogram (EEG). This potential is a late positive wave that occurs between 250 and 800 ms after the onset of a meaningful stimulus; the wave elicitation occurs in response to task-relevant events, and its latency depends on the stimulation paradigm.

The P300 has been widely used for BCIs, with many variations, but in all cases the paradigm is the same: the BCI system presents the user with some choices, one at a time; when it detects a P300 potential, the associated choice is selected. The user is normally asked to count the number of times the choice of interest is presented, so as to remain concentrated on the task. As the P300 is an innate response, it does not require training on part of the user.

In [3], Donchin and colleagues presented the first P300-based BCI, called also P300 speller, which permits to spell words. A grid of letters and symbols is presented to the user, and entire columns or rows are flashed one after the other in random order. Classification is made through stepwise discriminant analysis (SWDA) applied to averages of samples from epochs relative to the same stimulation (same row or same column).

In [4], a virtual-reality system is presented where users operates objects selected through the P300. Classification is made by comparing the correlation of single responses with the averages of all target and nontarget responses.

In [5], tests have been made both with healthy and impaired subjects. The subjects control a cursor by choosing among four commands (up, down, left, right) via the P300. Single-sweep detection is performed; independent component analysis (ICA) is used to decompose the EEG signal, a fuzzy classifier identifies a candidate P300 component among the ones extracted by ICA, and a neural network classifies it as target or non-target. The system is more effective with healthy subjects, though no exact reason could be pinpointed.

In [6], an initial attempt at using a BCI in a home environment is reported: a person with ALS uses a P300 speller on a daily basis. The system is very similar to the original Donchin's speller, with a few differences in the detection algorithm.

Many techniques for detecting the P300 extract relevant features from the EEG signal and feed those features into a classifier. In these approaches, feature extraction becomes the key point, and doing it by hand can be at the same time cumbersome and suboptimal. In this paper we face the issue of feature extraction by using a genetic algorithm (GA) able to retrieve the relevant aspects of the signal to be classified in an automatic fashion.

GAs have been used already in the BCI field, although differently from the present work: in [7], the best combination between different features and different classifiers is sought for a motor-imagery task, while in [8], a classifier operating on P300 features is selected by a GA.

In the following section, we present the paradigm used to collect the EEG data for the present study, while Section 3 gives a brief overview of the GA. Section 4 presents the performance achieved by the GA and a graphical interpretation of evolved classifiers.

2 Experimental Setup

2.1 Subjects

A group of 10 subjects affected by ALS, hospitalized in the S. Camillo structure, and a group of 4 healthy subjects voluntarily participated to the study (ALS group: 3 females and 7 males, mean age of 55 years, range 31–73 years; control group: 2 females and 2 males, mean age of 36 years, range 27–41 years). The research was approved by the ethical committee of the S. Camillo Hospital; informed consent was obtained according to the Declaration of Helsinki. All participants underwent neuropsychological evaluation and auditory odd-ball P300 testing, in order to exclude cognitive deficits. We assessed that all participants had preserved auditory, visual, and cognitive functions, including adequate language comprehension.

2.2 BCI Paradigm

An experiment was carried on to test the ability of the subjects to use a BCI based on P300 elicitation with an on-line single-sweep classifier. The paradigm consisted of a presentation of finite sequences of visual stimuli on a computer screen to the subjects. They were asked to control the movement of a virtual object (a blue ball) from the center of the monitor to one out of four peripheral target images representing generic needs. The initial distance between the virtual object and the target image encompassed four discrete steps. Upward, rightward, downward and leftward arrows in peripheral positions of the monitor (see Figure 1 a) were flashed in random order. Each arrow indicated one out of four possible directions for the movement of the ball. Participants had to pay attention to the arrow indicating the target image direction (target arrow; probability of occurrence: 25%), but to ignore the arrows indicating wrong directions (distracting arrows; probability of occurrence: 75%). The subjects had to move the blue ball along only one direction, according to the target image specified by the examiner.

Each trial comprised the flashing of one arrow for 150 ms (see Figure 1 b), followed by the data processing necessary for P300 recognition, and finally the generation of the feedback consisting in

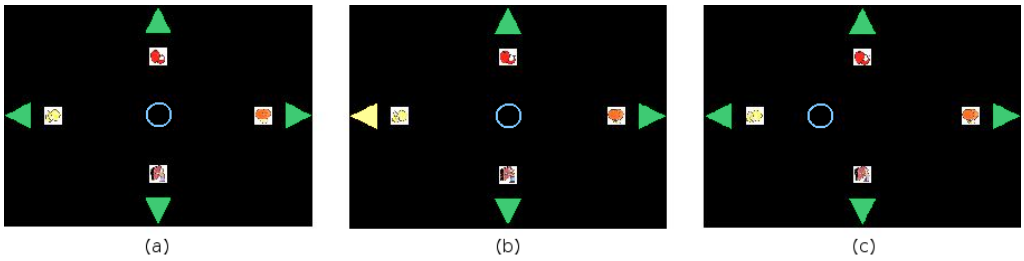

Figure 1: Representation of a trial. (a) The blue ball, the target images and the four directions arrows; (b) the flashed arrow; (c) the movement of the ball after a P300 recognition.

the movement of the ball (see Figure 1 c). The interval between the presentation of two arrows (inter-trial interval — ITI) was 2.5 s, in order to achieve optimal on-line data processing. A session was defined as the complete sequence of trials sufficient to reach the target image (range: 13-92 trials). We hypothesized that every target arrow should elicit a P300 wave. Every time a P300 was detected during the trial, the ball moved on the graphical interface in the direction of the flashed arrow. Each participant performed eight learning sessions (LS) in the first day, and sixteen testing sessions (TS) spread over the following 11 days (more precisely, first day: 8 LS → second day: 4 TS → two days interval → fifth day: 4 TS → two days interval → eighth day: 4 TS → two days interval → eleventh day 4 TS). Learning sessions were characterized by an ideal feedback (after each target stimulus the ball moved), while all testing sessions were characterized by a real feedback (the movement of the ball depended on the classification algorithm).

2.3 Data acquisition

EEG electrodes were placed according to the international 10-20 system at Fz, Cz, Pz and Oz; the EOG was placed at SO2; all electrodes were referenced to the left earlobe. The five channels were amplified, band-pass filtered between 0.15 Hz and 30 Hz, sampled at 200 Hz, and digitized (with a 16-bit resolution). Every ERP epoch, synchronized with the stimulus, began 500 ms before the stimulus onset, and ended 1000 ms after the stimulus onset (1500 ms total). Thus, after each stimulus (trial) the system recorded 300 samples per each of the 5 channels, available for on-line and off-line processing.

3 The Genetic Algorithm

We applied the genetic algorithm described in [9] to the data described in the previous section in an offline fashion. In this section, only a very brief description of the algorithm is given; details are given for the fitness function, as it differs from the one used in the cited work.

Genetic algorithms are a class of optimization algorithms that mimic the way natural evolution works. In a genetic algorithm, a set of possible solutions to an optimization problem are coded in strings called *chromosomes*; solutions are evaluated, and the best ones (those with highest *fitness*) are selected and combined together to form new possible solutions, in a process that mimics evolution among living beings. After some repetitions of the procedure, good solutions emerge.

In the genetic algorithm used in this work, each individual (chromosome) represents a set of possible features for discriminating the presence of a P300 in EEG recordings. Each gene encodes a feature and an EEG channel from which to extract it; a feature is obtained by multiplying the EEG channel by a weight function, whose exact shape is encoded by parameters in genes (see Figure 2 for examples of weight functions). Genetic operators are a variant of 1-point crossover and uniform mutation, and tournament selection with elitism is used.

The fitness of a chromosome is computed by evaluating the performance of a logistic classifier on the features it encodes. To have a fair estimate of the performance, a 4-fold cross-validation

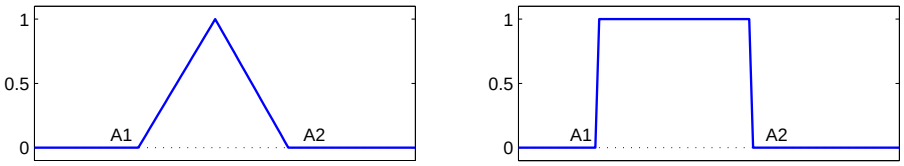

Figure 2: Weight functions used for feature extraction

scheme on the training set is used, and the mean performance on the 4 folds is used as the fitness. The "performance" f of a classifier is obtained by combining precision p_{T} and recall r_{T} for targets according to this formula:

$$f = \frac{2}{3}\, p_{\mathrm{T}} + \frac{1}{3}\, r_{\mathrm{T}} \tag{1}$$

The definitions of precision and recall in terms of *true positives* (TP), *false positives* (FP), and *false negatives* (FN) are:

$$p_{\mathrm{T}} = \frac{TP}{TP + FP} \qquad\qquad r_{\mathrm{T}} = \frac{TP}{TP + FN} \tag{2}$$

An analysis of the combination of the features extracted by the genetic algorithm and the classifier trained on the training set allows to compute weights assigned to individual EEG samples.

4 Results and Conclusions

The GA was applied in an offline fashion to the data recorded as described in Section 2. EEG data were decimated to half of the original frequency; epochs were trimmed to the interval from $-.2\,\mathrm{s}$ to $+.8\,\mathrm{s}$ (i.e., half a second was thrown away), and the linear trend was removed. No normalization of data was performed; we tried to normalize EEG data before running the GA, but it did not change the test performance significantly, so we chose the way that required less computation.

The GA was trained on the first three quarters of the available data for each subject, and the features encoded by the best chromosome and the corresponding classifier trained on the training set were applied to the remaining quarter of the data. This procedure was repeated with 14 independent runs of the GA on each subjects.

The results are shown in Table 1: the mean values and the standard deviation of the recall of targets and non-targets over all the 14 runs are presented. Subjects S11, S12, S13, S14 are healthy, while all the others are affected by ALS. For most subjects, the performance of the GA is good, achieving consistently more than 70% of recall in 8 subjects (i.e., the mean is at least two standard deviations over 70%). In 6 subjects, the classifiers achieved often more than 80% of correct answers.

The last two columns of Table 1 show how effective this BCI can be for the various subjects; the first of the two is the expected fraction of times a user reached the desired target, and the second is the expected number of trials needed to reach a target. These numbers have been obtained with a Montecarlo simulation of the BCI done with the mean recall values obtained by the GA. Only 3 subjects reach an accuracy in the task performance lower than 80%, so we can consider the classifier performance satisfactory in at least 11 cases. For a comparison, a perfect classifier (with 100% recall) reaches the target always, in 12 trials on average; a random classifier (50% recall) reaches the target in 25% of the cases, and each selection takes almost 21 trials on average.

Figure 3 shows what the GA has found: the continuous green lines represent the weight assigned to individual EEG samples in the final logistic classifier, while the averages of targets and non-targets epochs are given as references. Units are µV and s; the templates have been scaled to fit in the graphs. The plots regard two classifiers that reached over 90% of correct epochs for Subject S2; although the averages are not very different, the algorithm found good classifiers concentrating more on the Fz and Cz channels. By comparing the averages of targets

Subject	Training		Test		Recall		Exp.	Perf.
	Targ.	Non-T.	Targ.	Non-T.	Targ.	Non-T.	Targ.	Len.
S1	204	703	99	284	61%±3%	66%±2%	57%	22.9
S2	121	354	58	166	96%±1%	95%±1%	100%	13.1
S3	98	277	57	206	63%±5%	77%±3%	77%	23.3
S4	175	492	61	178	86%±4%	83%±3%	93%	16.4
S5	114	325	32	94	75%±7%	78%±3%	85%	19.5
S6	124	356	52	148	87%±3%	85%±5%	95%	16.0
S7	144	433	50	138	82%±4%	76%±3%	85%	17.8
S8	112	340	39	112	82%±7%	72%±5%	80%	17.9
S9	185	529	50	154	94%±2%	86%±2%	96%	14.6
S10	219	690	47	142	97%±2%	82%±2%	94%	14.5
S11	86	228	30	84	75%±4%	81%±2%	88%	19.3
S12	116	327	34	95	85%±5%	90%±3%	98%	15.8
S13	218	617	84	240	77%±3%	77%±2%	84%	19.0
S14	165	489	61	170	55%±7%	73%±4%	63%	26.0

Table 1: Means and standard deviations of recall for targets and non-targets obtained in 14 runs of the GA

and non-targets in the plots, it is possible to see that they have important differences in the Fz and Cz channels between 400 and 800 ms after the stimulus. The P300 complex includes other subcomponents besides P3b, which has its maximal amplitude in parietal regions, as P3a and the slow wave post P3b, which have their maximal amplitude over fronto-central regions. P3a occurs when the changes in physical properties of the novel stimulus are task relevant and attention is switched to the stimulus source. The slow wave post P3b occurs in the latency range from 500 to 1400 ms, and its amplitude increases as the task becomes more demanding and difficult [10, 11]. This suggests that the characteristics of the task and the attention requirement determine strong response in the front-central region, which is exploited by the classifiers found by the GA.

We think that the results obtained offline are very promising and could be confirmed in future online tests. The GA is very suitable for an online application; for the data set used in the present work and on a low-end PC, a single run of the GA takes between 5 and 15 minutes, depending on the subject, and the classification of an epoch less than 1 ms.

Acknowledgments

This work has been partially supported by the Italian Institute of Technology (IIT), and by the grant "Brain-Computer Interfaces in Everyday Applications" from Politecnico di Milano and Regione Lombardia. We are thankful to all the participants in our study.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113:767–791, 2002.

[2] S. Sutton, M. Braren, J. Zubin, and E. R. John. Evoked-potential correlates of stimulus uncertainty. *Science*, 150(3700):1187–1188, 1965.

[3] E. Donchin, K. M. Spencer, and R. Wijesinghe. The mental prosthesis: Assessing the speed of a P300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 8(2):174–179, 2000.

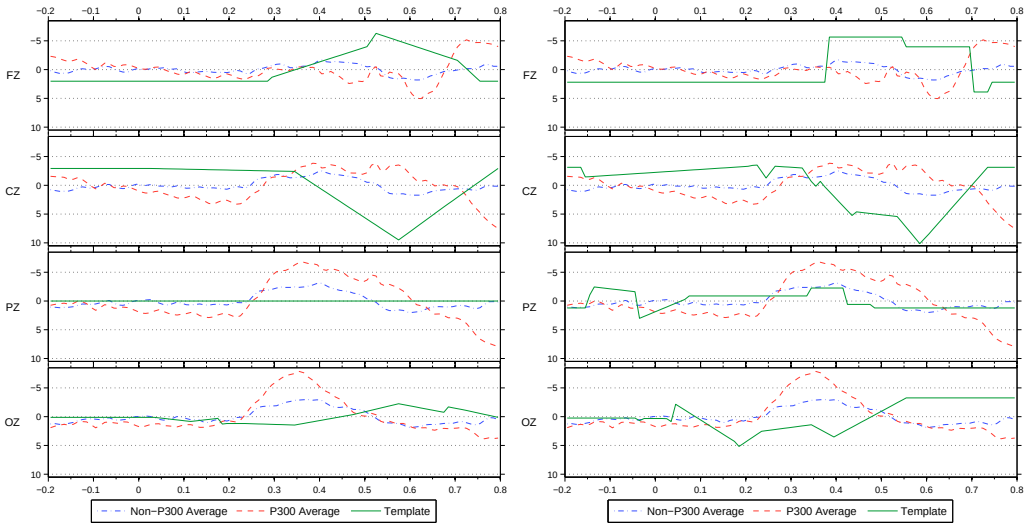

Figure 3: Examples of templates obtained in two different GA run for Subject S2

[4] J. D. Bayliss, S. A. Inverso, and A. Tentler. Changing the P300 brain computer interface. *Cyberpsychology & Behavior*, 7(6):694–704, 2004.

[5] F. Piccione, F. Giorgi, P. Tonin, K. Priftis, S. Giove, S. Silvoni, G. Palmas, and F. Beverina. P300-based brain computer interface: reliability and performance in healthy and paralysed participants. *Clinical Neurophysiology*, 117(3):531–537, 2006.

[6] T. M. Vaughan, D. J. Mcfarland, G. Schalk, W. A. Sarnacki, D. J. Krusienski, E. W. Sellers, and J. R. Wolpaw. The Wadsworth BCI research and development program: At home with BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):229–233, 2006.

[7] R. Boostani, B. Graimann, M. H. Moradi, and G. Pfurtscheller. A comparison approach toward finding the best feature and classifier in cue-based BCI. *Medical and Biological Engineering and Computing*, 45(4):403–412, 2007.

[8] L. Citi, R. Poli, C. Cinel, and F. Sepulveda. Feature selection and classification in brain computer interfaces by a genetic algorithm. In *Late-breaking papers of the Genetic and Evolutionary Computation Conference (GECCO-2004)*, volume CD-ROM, 2004.

[9] B. Dal Seno, M. Matteucci, and L. Mainardi. A genetic algorithm for automatic feature extraction in P300 detection. In *2008 International Joint Conference on Neural Networks (IJCNN)*, pages 3144–3151, 2008.

[10] J. Polich. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118:2128–2148, 2007.

[11] E. Altenmüller. Psychophysiology and EEG. In E. Niedermeyer and F. Lopes Da Silva, editors, *Electroencephalography*. Williams & Wilkins, Baltimore, USA, 3rd edition, 1993.