Single-trial P300 detection with Kalman ltering and SVMs

Lucie Daubigney, Olivier Pietquin

ESANN 2011 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2011, i6doc.com publ., ISBN 978-2-87419-044-5. Available from http://www.i6doc.com/en/livre/?GCOI=28001100817300.

Single-trial P300 detection with Kalman filtering and SVMs

Lucie Daubigney and Olivier Pietquin

SUPELEC / UMI 2958 (GeorgiaTech - CNRS)
2 rue Edouard Belin, 57070 Metz - France
firstname.lastname@supelec.fr

Abstract. Brain Computer Interfaces (BCI) are systems enabling humans to communicate with machines through signals generated by the brain. Several kinds of signals can be envisioned as well as means to measure them. In this paper we are particularly interested in even-related brain potentials (ERP) and especially visually-evoked potential signals (P300) measured with surface electroencephalograms (EEG). When the human is stimulated with visual inputs, the P300 signals arise about 300 ms after the visual stimulus has been received. Yet, the EEG signal is often very noisy which makes the P300 detection hard. It is customary to use an average of several trials to enhance the P300 signal and reduce the random noise but this results in a lower bit rate of the interface. In this contribution, we propose a novel approach to P300 detection using Kalman filtering and SVMs. Experiments show that this method is a promising step toward single-trial detection of P300.

1 Introduction

Event-related brain potentials (ERP) are modifications of the electrical activity of the brain happening after some stimulation (*e.g.* visual or audio stimuli) [1]. They can serve in the framework of Brain-Computer Interfaces (BCI) [2], or in the analysis of brain diseases [3]. There are several reasons for which the detection of $P300$ is difficult among which, the Signal to Noise Ratio (SNR) is one of the most important. For this reason, it is very frequent that a synchronous averaging of several signals is used to enhance the SNR [4], considering that the time evolution of EEG is similar for all ERPs. This assumption is actually false and ERPs can suffer from distorsions which are not only due to additive noise. For this reason, among other methods, Kalman filtering [5] has been studied to model in an online fashion the ERPs [6, 7] and has proven to be an efficient method. The Kalman filter can be combined with a classifier for the $P300$ to be automatically detected [8]. We propose here an original way to detect the $P300$ by using the coefficients found by the Kalman filter to represent the $P300$ as features for a Support Vector Machine (SVM).

2 Signal modeling

The EEG signal arises from the sum of several electrical activities produced simultaneously by the brain. An exemple is given Figure 1. The activity produced

ESANN 2011 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2011, i6doc.com publ., ISBN 978-2-87419-044-5. Available from http://www.i6doc.com/en/livre/?GCOI=28001100817300.

by a resting brain will be considered as noise (white, centered, gaussian) and the *P300* wave as the signal of interest. Notice that on the Figure 1, the maximum amplitude is about $25\mu V$. The maximum amplitude for a *P300* wave is about $10\mu V$. The signal to noise ratio is thus very small.

A parametric model of this signal is proposed under some assumptions and the parameters are adapted online according to observations of the brain activity. The signal is supposed to be C_1, bounded, defined on a segment of $\mathbb{R}$. Under these assumptions, the signal, $y(t)$, can be approximated by a sum of Gaussians : $y(t) = \alpha^T \phi_t$ with $\alpha^T = [\alpha_1, \ldots, \alpha_p]$ and $\phi_t^T = [exp(\frac{(t-t_1)^2}{2\sigma_1^2}), \cdots, exp(\frac{(t-t_p)^2}{\sigma_p^2})]$. On Figure 2 one can see the expected shape of a *P300* wave, resulting of a sum of three Gaussians.

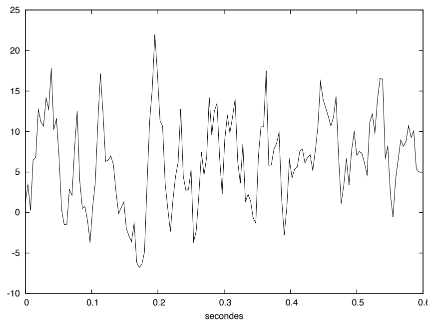


Fig. 1: Raw example of the electrical activity of the brain captured on a surface EEG (abscissa in s, ordinate in μV).

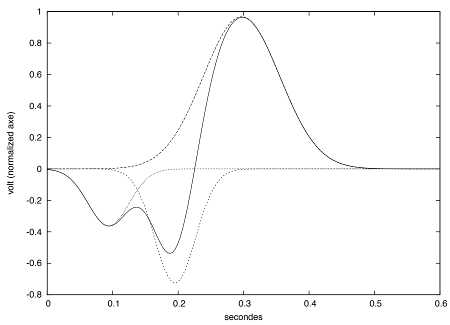


Fig. 2: Modeling of a *P300* as a sum of three gaussians (abscissa in s, ordinate in normalized V). The plain curve is the expected shape of the *P300*.

The α parameters have to be estimated by using the observations of the brain activity. In the following a method based on the Kalman filtering paradigm is proposed, after the works presented in [7, 6]. After the α parameters are found they are used by a SVM to detect the $P300$ in the signal.

3 Parameters estimation

The *P300* wave is thus described by a generic model and the model parameters have to be adapted to detect *P300* in the signal. As said before, the information used to learn the parameters come from noisy observations. This estimation problem is cast into the Kalman filtering paradigm. In this paradigm, the hidden state of a dynamic system has to be estimated online from noisy observations. This is done given that the evolution of the state is ruled by a known equation (*evolution equation*, Equation 1) and that the state is linked to the observations by another known equation (*observation equation*, Equation 2). These two equations form the so called *state-space* representation of the problem. In the general case, Equation 1 x represents the state, F the process matrix and v is the

ESANN 2011 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2011, i6doc.com publ., ISBN 978-2-87419-044-5. Available from http://www.i6doc.com/en/livre/?GCOI=28001100817300.

process noise. In the Equation 2 which describes the observations, y represents the observations, H the matrix of observations and w observation noise.

$$x_{t+1} = F_t x_t + v_t \tag{1}$$

$$y_{t+1} = H_t x_t + w_t \tag{2}$$

The Kalman filter computes the estimation of the state at time t by using the previous estimation, at time $t-1$ and the current observation. The estimation of the state is computed according to Equation 3.

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(\hat{y}_t - y_t) \tag{3}$$

$$E\left[\left\|\hat{x}_{t|t} - x_t\right\|^2 |y_1, \ldots, y_t\right] \tag{4}$$

In Equation 3, $\hat{x}_{t|t-1} = E[x_t|Equation\ 1] = F_t\hat{x}_{t-1|t-1}$ represents the prediction obtained according to Equation 1 and $\hat{x}_{t-1|t-1}$ the estimation at the previous time and given that the process noise is considered as white and centered ($E[v] = 0$). $\hat{y}_t = H_t\hat{x}_{t|t-1}$ is the prediction of the observation ($E[y_t|Equation\ 2]$), computed according to Equation 2 and y_t is the current observation. K_t is the so-called *Kalman gain*. It is computed such as that the estimates minimise the expression 4 which is explained later in this section.

First, the parameter estimation problem is cast into a state-space formulation. According to section 2 the signal can be approximated by a sum of Gaussians. The observation equation is thus Equation 5 with Φ a matrix of known Gaussians and w the observation noise (which reflects the fact that the real signal is not exactly a sum of Gaussians). An equation describing the evolution of the system is also needed. Here, the hidden state we are trying to estimate is the vector containing the parameters which we will note as α_t. As the evolution of the parameters α over time is not well known (it depends on the $P300$ evolution), a random walk is used (see Equation 6). The sample at the time t is the same as the one at the time $t-1$ plus a small random variation. The aim of the filter is to find the hidden parameters α_t^i (i from 0 to p, the number of Gaussians chosen to estimate the signal) from observations. The bigger the number of observations, the better the estimation of the parameters. For an infinite number of observations, the estimated parameters tends towards the actual ones.

$$y_t = \alpha_t^T\Phi + w_t \tag{5}$$

$$\alpha_t = \alpha_{t-1} + v_t \tag{6}$$

Three steps are needed to compute the gain. The *a priori* estimation of the covariance of the error between the hidden variable α_t and its estimation $\hat{\alpha}_t$ is computed according to the Equation 7, Q being the matrix of covariance of the process noise whose expression is $E\left[vv^T\right]$. This Q matrix takes into accounts the small variations between the *P300* waves. Q is usually chosen equal to $\eta P_{t-1|t-1}$, η depending of the differences between the *P300*. Then the gain[1] is computed using the Equation 8, with $\sigma_{w_i}^2$, variance of the observation

[1]The initial value, for a normalized signal are : $P_0 = \lambda Id$ with $\lambda = 1$, $\sigma_{w_0} = 10^{-3}$, $\alpha_0 = 0$ or α_0 fixed with the values of a previous estimation.

ESANN 2011 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2011, i6doc.com publ., ISBN 978-2-87419-044-5. Available from http://www.i6doc.com/en/livre/?GCOI=28001100817300.

noise, which represents the accepted gap between the estimated parameters and the real parameters. This expression is determined to minimize the expression given in 4. Finally the *a posteriori* estimation of the covariance of the error is computed using Equation 9.

$$P_{t|t-1} = P_{t-1|t-1} + Q \tag{7}$$

$$K_t = \frac{P_{t|t-1}\phi_t^T}{\phi_t \cdot P_{t|t-1} \cdot \phi_t^T + \sigma_{w_i}^2} \tag{8}$$

$$P_{t|t} = (Id - K_t\phi_t) \cdot P_{t|t-1} \tag{9}$$

The Kalman filter is convenient to use in that case because the computation of the gain does not require a lot of resources. The filter can be used "online" to estimate the signal from the observations. Once a model of a signal without noise is obtained (like in [6, 7]), it can be used to detect if the signal contains or not a *P300* wave.

4 Detection

The *P300* wave is created by the brain after it has been submitted at a random time to an expected visual stimulus among other stimuli. The target stimuli is chosen before the beginning of the experiment. The aim is thus to determine in the EEG signal if the target stimulus provoked a *P300*. The times at which the stimuli had been displayed are known but the moment at which the *P300* may appear is not necessarily known precisely. This moment may vary by several milli secondes depending on the subject (gender, age, tiredness, etc). The *P300* wave is supposed to last around $400ms$.

An automatic detection of *P300* waves is made using the result of the Kalman filtering by a SVM. Here we want to classify the data in two categories : $P300$ or noise. The SVM is first trained with two sets of data in a supervised manner, then new data are submitted to the classifier to be classified. The originality of this work is to provide the estimated parameters α to the SVM classifier.

5 Experiment

5.1 Data

The processing of the data, the estimation of the signal and the classification, have been tested with two different datasets like in [6] : a set of artificial data and a set of real EEG signals. The artificial data is a constructed *P300* wave added to real background noise got from a human brain. This solution has fist been tested to be sure that the set of data contains *P300* and avoid problem of generating *P300*.

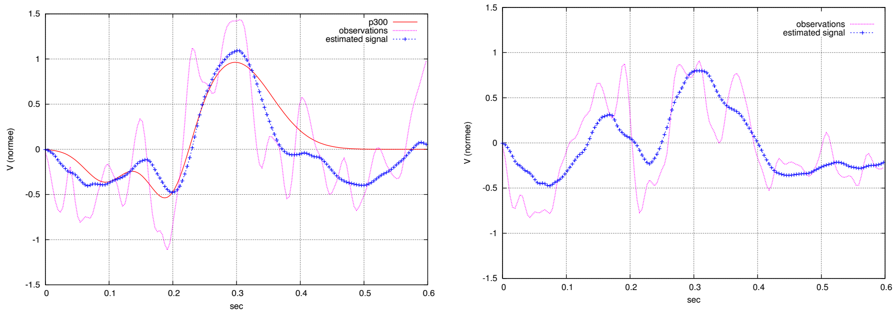
The real data comes from an experiment led with a human watching at several pictures displayed one after another. Among the displayed pictures, the subject has to find the target picture (s)he was asked to recognize. This experiment

ESANN 2011 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2011, i6doc.com publ., ISBN 978-2-87419-044-5.
Available from http://www.i6doc.com/en/livre/?GCOI=28001100817300.

is supposed to make the brain generate *P300* waves since the subject does not know at which moment the picture will be displayed.

5.2 Results

Each part of the data processing chain has been tested separately in both cases *i.e.* with artificial and real data. The Kalman filter has been fed with data containing a *P300* wave. Examples of the estimated signals are provided on Figure 3. The shape of the *P300* can be recognized in the estimated signal. Then, the accuracy of the classifier is tested. It has first to be trained with



(a) Artificial *P300* (b) *P300* from a real experiment

Fig. 3: Test of the Kalman filter. The signals have been normalized to be between -1 and 1. The curves have been obtained with the following initial values : $\sigma^2_{v_0} = 1$, $\eta = 1.10^{-5}$, $\lambda = 1$ and α take the values obtained in a previous estimation. The dotted curve gives the observation of a raw EEG after a pass-band filter (0 to 20 Hz), then estimated with Kalman filter (curve with crosses). The shape of the *P300* is reminded on the left figure.

exemples from both classes which are to be recognized (*P300* vs noise). The set of examples representing the first class is composed of α parameters from sections of observations which most likely contain a *P300* wave. Each section which more likely contains a *P300* is made of samples collected from the display of the target stimulus to several hundreds of ms (here $600ms$) after the display. The set of examples representing the second class contains α parameters computed from noise. Fifty sets of α parameters selected by hands have been used in each category. The performance of the classifier alone obtained with cross-validation are the following : for the artificial *P300* the rate of 90% of success is reached and for real *P300* the rate of 87% is reached.

The accuracy has been measured on a $60s$ signal containing periodic *P300*, with a period equal to $1.8s$. Sets of α parameters were computed periodically for a $600ms$ window, with a period equal to $200ms$. We thus know here, by construction, that the *P300* appears $300ms$ after the target stimulus. The results are given Figure 4. They are promising since only the sets of α computed for a portion of signal starting $300ms$ before the *P300* were detected by the classifier as containing a *P300*. The number of *false positives* is kept quite low while the

ESANN 2011 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2011, i6doc.com publ., ISBN 978-2-87419-044-5. Available from http://www.i6doc.com/en/livre/?GCOI=28001100817300.

true positives are fairly numerous. Similar results have been obtained with the

Classifier Signal	detect a *P300*	does not detect a *P300*
contains a *P300*	88%	12%
does not contain a*P300*	1%	99%

Fig. 4: Results of the experiments obtained with artificial data

real data although it is much difficult to assess since the exact time of appearance of *P300* in the signal is unknown. Particularly, a *true negatives* rate of 98.5% could be obtained. Yet, *true positives* were difficult to assess but the method worked in more than 50% of the cases where a *P300* was expected.

6 Conclusion

This new approach for the detection of *P300* gave encouraging results since the detection is only made in a single trial. But some improvements could be brought by leading more tests to determine exactly why experiments on real data do not give the same results as those got with the artificial data. Some tests could be ran about the reliability of generating the *P300* waves. Maybe the assumption we made concerning the shape of the *P300* as a sum of gaussians is not exactly true either. The results obtained could be compared with the ones of the speller which is the reference experiment ([2]) to eliminate some causes of problem.

References

[1] M. A. B. Brazier. Evoked responses recorded from the depths of the human brain. *Annals of the New York Academy of Sciences*, 112:33–59, 1964.

[2] L.A. Farwell and E. Donchin. Talking off the top of your head : toward a mental prosthesis utilizing event-related brain potentials. In *Electroencephalography and clinical neurophysiology*, volume 70 num 6, pages 510–523. December 1988.

[3] David E. J. Linden. The P300: Where in the Brain Is It Produced and What Does It Tell Us? *The Neuroscientist*, 11(6):563–576, 2005.

[4] Leif Sörnmo and Pablo Laguna. EEG signal processing. In *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, pages 55 – 179. Academic Press, Burlington, 2005.

[5] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic engineering*, 82 (1):35–45, 1960.

[6] Esben Vedel-Larsen, Jacob Fuglo, Fouad Channir, Carsten E. Thomsen, and Helge B.D. Sorensen. A comparative study between a simplified Kalman filter and sliding window averaging for single trial dynamical estimation of event-related potentials. *Computer Methods and Programs in Biomedicine*, 99(3):252 – 260, 2010.

[7] S.D. Georgiadis, P.O. Ranta-aho, M.P. Tarvainen, and P.A. Karjalainen. Single-trial dynamical estimation of event-related potentials: a Kalman filter-based approach. *IEEE Transactions on Biomedical Engineering*, 52(8):1397–1406, 2005.

[8] H. Ishita, M. Sakai, J. Watanabe, W. Chen, and D. Wei. Development of P300 Detection Algorithm for Brain Computer Interface in Single Trial. In *7th IEEE International Conference on Computer and Information Technology, 2007. CIT 2007.*, pages 1100–1105. IEEE, 2007.