A ROBUST CONSTRAINED METHOD FOR THE EXTRACTION OF P300 SUBCOMPONENTS

Loukianos Spyrou and Saeid Sanei

Centre of Digital Signal Processing, School of Engineering, Cardiff University, Cardiff, UK, CF24 3AA, emails: spyroul, saneis@cf.ac.uk

ABSTRACT

Separation of event related potentials (ERPs) is being investigated in this paper by means of a new constrained blind source extraction (BSE) technique. Specifically, the P3a and P3b subcomponents are extracted from an ERP incorporating some prior knowledge of the desired signals. The method reliably extracts the desired P300 subcomponents with great accuracy. One advantage of BSE algorithms over Blind Signal Separation (BSS), which have been used in EEG research, is that the algorithm focuses on one source instead of trying to extract all of the sources simultaneously which involves indeterminacy in the number of sources. To testify the performance of the algorithm, some experiments are shown which are performed on real EEG data.

1. INTRODUCTION

Event-related potentials (ERPs) are electroencephalograms (EEGs) which directly measure the electrical response of the cortex to sensory, affective and/or cognitive events. The good temporal resolution offered by ERPs allows accurate study of the time course of information processing unavailable to other neuroimaging techniques. On the other hand, spatial resolution in that approach is limited. In addition, overlapping components of the ERP which represent specific stages of information processing are difficult to distinguish [1] [2]. An example is the P300 wave, a positive ERP component which occurs with a latency of about 300ms after novel stimuli, or task relevant stimuli, requiring an effortful response on the part of the individual under test [3] [4] [1] [5] [2]. The P300 wave represents cognitive functions involved in orientation of attention, contextual updating, response modulation, and response resolution [3] [1], and consists of multiple overlapping subcomponents, two of which are identified as P3a and P3b [5] [2]. P3a reflects an automatic orientation of attention to novel or salient stimuli independent of task relevance [5] [6]. Prefrontal, frontal and anterior temporal brain regions play a major role in generating P3a giving it a frontocentral distribution [1] [5]. In contrast, P3b has a greater centro-parietal distribution due to its reliance on posterior temporal, parietal and posterior cingulate mechanisms [1] [2]. P3a is also characterised by a shorter latency and more rapid habituation than P3b [5] [2]. Figure 1 illustrates some typical P3a and P3b waveforms from temporal-basal and temporo-superior dipoles [7].

P300 has significant diagnostic and prognostic potential especially when combined with other clinical examination [4] [2]. However, in order for this to be fully realised, efficient and reliable methods for separating P300 sources and its subcomponents must be established [4].

The most common method of averaging the EEG over a number of trials[1] cannot reliably distinguish the small differences between the various subcomponents since they temporally overlap. Some common signal processing methods used to overcome this limitation are based on Blind Signal Separation (BSS) [8] [9] [10] and Least squares approaches such as [11] [12]. In this paper we use a BSE method which extracts only one signal at a time and can be modified to extract only the sources of interest.

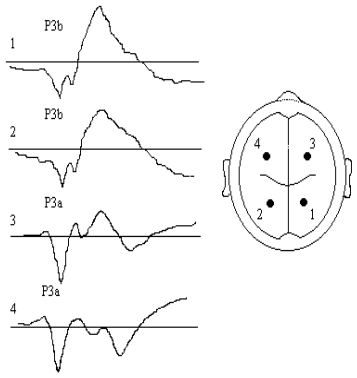

Fig. 1. Some examples for P3b (1 and 2) and P3a (3 and 4) signals and their corresponding typical locations.

2. CONSTRAINED BLIND SIGNAL EXTRACTION

The main advantages of BSE over BSS can be the following: (a) The algorithm concentrates only on the signals of interest

[1] in this context trial means the recorded response of the brain after a stimuli has been applied

based on their properties (b) in general BSE algorithms are much simpler and (c) they can be easily modified for a number of situations. In this paper we use prior knowledge of the shape and latency of P300 signals to obtain the signals. This is done by using a constraint function, which is imposed on the original BSE cost function.

We start with the normalised kurtosis cost function, which estimates the deviation of a random variable from Gaussianity. For kurtosis equal to zero the signal is Gaussian, for positive values it is super-Gaussian and for negative values sub-Gaussian. From the Central Limit Theorem it is known that a signal consisting of a mixture of different signals with different probability density functions tends to have a Gaussian distribution. So, if we try to maximise the absolute value of the kurtosis of the signal it will be separated from the mixture.

In the BSE model the output of the algorithm is described as [13]:

$$\mathbf{y} = \mathbf{w}^T\mathbf{X} \tag{1}$$

where **y** is the output vector of size $1 \times T$ (T is the number of samples), **w** is the unmixing vector ($N \times 1$, where N is the number of electrode signals) obtained by the algorithm and **X** is the data matrix ($N \times T$) consisting of the electrode signals.

The cost function is:

$$J_m(\mathbf{w}) = -\frac{1}{4}|kurt(\mathbf{y})| \tag{2}$$

where $kurt(\mathbf{y})$ is the normalised kurtosis and is given by:

$$kurt(\mathbf{y}) = \frac{E(|\mathbf{y}|^4)}{E(|\mathbf{y}|^2)^2} - 3 \tag{3}$$

where $E()$ denotes the expectation. This leads to the following online adaptation rule [13]:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + n(k)\varphi(y(k))\mathbf{x}(k) \tag{4}$$

where

$$\varphi(y(k)) = b(\frac{m_2(y(k))}{m_4(y)(k)}y(k)^3 - y(k))\frac{m_4(y(k))}{m_2^3(y)(k)} \tag{5}$$

and

$$m_q(k) = (1-n_0)m_q(k-1) + n_0|y(k)|^q \tag{6}$$

where k is the iteration number (and sample number), m_q is the q_{th} moment and $n_0 \epsilon (0,1]$ adjusting the influence of the previous estimate of the moment and the current estimate.

The constraint is imposed upon the normalised kurtosis cost function and utilises prior knowledge about the P300 shape and latency. The procedure is to obtain a reference P300 signal and in turn a vector $\mathbf{w}_{opt}$ which minimises the Euclidean distance between that reference and the data in the following way:

$$J_c(\mathbf{w}_{opt}) = ||\mathbf{y}_{ref} - \mathbf{w}_{opt}\mathbf{X}||^2 \tag{7}$$

where $||.||^2$ represents the Euclidean distance. The solution to this is the common minimum-norm solution:

$$\mathbf{w}_{opt} = (\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}\mathbf{y}_{ref}^T \tag{8}$$

Then, we need to minimise the distance of the obtained **w** from (4) and $\mathbf{w}_{opt}$ from (8). So, we want to minimise:

$$d(\mathbf{w}) = ||\mathbf{w}_{opt} - \mathbf{w}||^2 \tag{9}$$

This is the constraint cost function placed in the original cost function (2) according to the theory of penalty parameters. Hence, the adaptation rule becomes[2]

$$\mathbf{w}(k+1) = \mathbf{w}(k) + n(k)\varphi(y(k))\mathbf{x}(k) + K(\mathbf{w}(k) - \mathbf{w}_{opt}) \tag{10}$$

where K is the penalty parameter. It should not be too high in order not to overcome the effect of the main cost function or too low so its effect is too small. The selection of the appropriate value for K will be discussed in section 3. The product $\mathbf{y}_{LS} = \mathbf{w}_{opt}\mathbf{X}$ can be considered as the closest representation of $\mathbf{y}_{ref}$ that can be obtained by unmixing the data such as described by the model in (1).

2.1. Reference signal

The reference signal $\mathbf{y}_{ref}$ used to calculate $\mathbf{w}_{opt}$ is obtained via the following method. By prior knowledge about the shape and latency of a usual P300 component and considering a number of trials of the same experiment we can provide a reference signal. This can be achieved by averaging of all the relevant trials[3] of an EEG to obtain a temporally averaged ERP of dimensions $N \times T$, where N is the number of electrodes and T the number of samples. Then we perform a spatial averaging of all the electrode signals and select the appropriate time period corresponding to the appropriate P300 subcomponent(for example 250-300ms for P3a), zero the rest and we get a signal of dimensions $1 \times T$. This is the $\mathbf{y}_{ref}$ signal which is used in (8) to obtain $\mathbf{w}_{opt}$. Two $\mathbf{y}_{LS}$ examples can be seem in Figures 3 and 5.

3. EXPERIMENTAL RESULTS

The EEG data were recorded using a Nihon Kohden model EEG-F/G amplifier and Neuroscan Acquire 4.0 software. EEG activity was recorded following the international 10-20 system from 15 electrodes. The reference electrodes were linked to the earlobes. The impedance for all the electrodes was below 5kΩ, sampling frequency Fs=2kHz and the data were subsequently bandpass filtered (0.1-70Hz). This frequency range was chosen to be compatible with [14].

[2] the third term should be $2K(\mathbf{w}(k) - \mathbf{w}_{opt})$ but the 2 is absorbed by the penalty parameter so it can be omitted

[3] some trials do not produce a P300 component

Subjects were required to sit alert and still with their eyes closed to avoid any interference. Also, to avoid any muscle artefact the neck was firmly supported by the back of the chair. The feet were rested on a footstep. The stimuli were presented through ear plugs inserted in the ear. Forty rare tones (1kHz) were randomly distributed amongst 160 frequent tones (2kHz). Their intensity was 65dB with 10ms and 50ms duration for rare and frequent tones respectively. The subject was asked to press a button as soon as they heard a low tone (1kHz). The ability to distinguish between low and high tones was confirmed before the start of the experiment. The task is designed to assess basic memory processes. ERP components measured in this task included N100, P200, N200 and P3a and P3b.

After obtaining the data, they were temporally averaged for all event related trials (40 events). The algorithm automatically obtains the reference signal, $\mathbf{w}_{opt}$ and hence $\mathbf{y}_{LS}$. The algorithm was applied to many sets of data and it extracted the desired components successfully. The obtained P3a and P3b were more highlighted compared to the reference signal and their shape was more in agreement to typical P3a and P3b shapes. Some typical obtained signals can be seen in Figures 2 and 4 while the corresponding $\mathbf{y}_{LS}$ signals are shown in Figures 3 and 5 and the spatio-temporal averaged reference signal[4] is shown in Figure 6. As can be seen from the figures the constrained method obtains good representations of P3a and P3b. Their respective latencies are in agreement with prior physiological research and their shapes are more smooth than those of the $\mathbf{y}_{LS}$ signals. This is expected since the algorithm tries to obtain an output close to the $\mathbf{y}_{LS}$ but also as less mixed as possible.

By comparing the resulting signals with the unconstrained case useful insights can be obtained as to the selection of the appropriate parameters for the algorithm (such as the K penalty parameter, the learning rate n and the n_0 paratemter). If K is set to zero the algorithm is unconstrained and the resulting output is not the desired one. Gradually increasing the K parameter starting from a small value (about 10^{-5}) the algorithm's behaviour can be observed. At really small values the influence of the constraint is minimal and does not produce valuable results. High values tend to make the algorithm crash. A practical value that produces good results while the signal is not very close to the reference is 10^{-4}. In fact, the value of K can be adapted and updated iteratively according to the changes in the gradients of J_c and J_m. The learning rate was set to 10^{-3} and it was reduced every iteration by 1%. The n_0 parameter was set to 0.5.

4. CONCLUSIONS

In this paper a robust constrained BSE method has been developed to extract the P3a and P3b signals from within an

[4]This signal is used to obtain the $\mathbf{y}_{ref}$ signals for P3a and P3b.

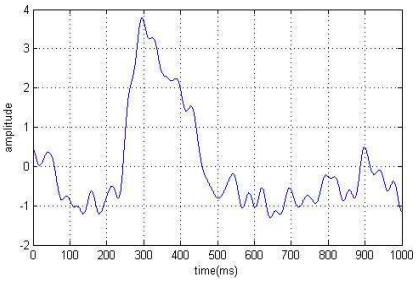

Fig. 2. The P3a signal obtained with the constrained BSE algorithm.

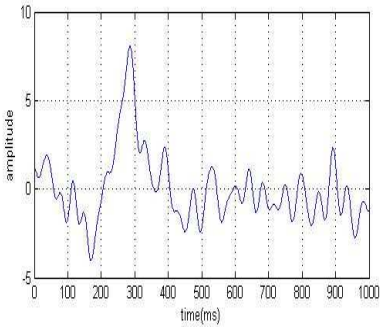

Fig. 3. The $\mathbf{y}_{LS}$ signal used to obtain the signal of Figure 2.

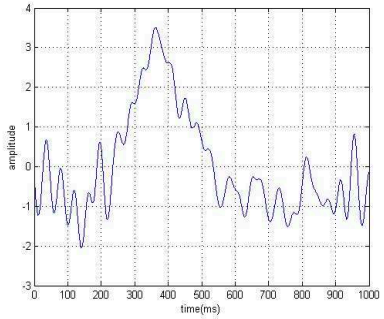

Fig. 4. The P3b signal obtained with the constrained BSE algorithm.

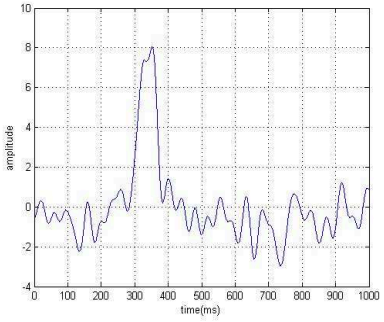

Fig. 5. The $\mathbf{y}_{LS}$ signal used to obtain the signal of Figure 4.

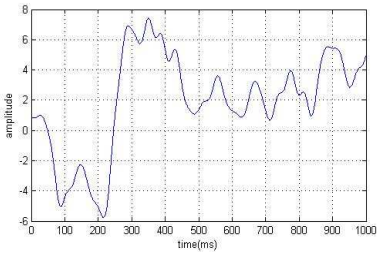

Fig. 6. Spatio-temporal averaged EEG reference signal.

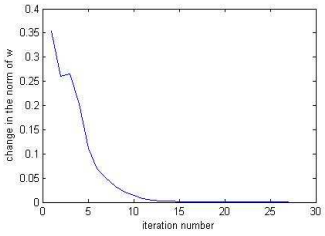

Fig. 7. Convergence of the algorithm for the signal of Figure 2.

EEG. The algorithm minimises the distance between a reference signal and the estimated output while trying to make the output as less mixed as possible. The effect of the constraint can be adjusted via the penalty parameter and in effect the algorithm points the solution towards the desired signal. The reference signal is obtained automatically by the algorithm and the convergence is very fast (less than 30 iterations). The algorithm was applied to real EEG data and P3a and P3b signals were separated successfully. Future work entails incorporating reference signals for more ERP components (P50, N100, P200, N200), the adaptation of the optimisation parameters (such as K, n and n_0) and the application to single-trial EEGs.

5. REFERENCES

[1] J. Dien, K. Spencer and E. Donchin . "Localization of the event-related potential novelty response as defined by principal components analysis." Brain Res. Cogn. Brain Res., Vol. 17, No. 3, pp.637-650, 2003.

[2] T. Frodl-Bauch, R. Bottlender and U. Hegerl,"Neurochemical substrates and Neuro anatomical Generators of the Event-Related P300", Neuropsychobiology, Vol. 40, pp. 86-94, 1999.

[3] A. Kok, J. Ramautar , M. De Ruiter , G. Band and K. Ridderinkhof, "ERP components associated with successful and unsuccessful stopping in a stop-signal task", Psychophysiology, Vol. 41, No. 1 , pp. 9-20, 2004.

[4] J. Polich, "Clinical application of the P300 event-related brain potential", Phys. Med. Rehabil. Clin. N. Am., Vol. 15, No. 1, pp. 133-161, 2004.

[5] D. Friedman, Y. Cycowicz and H. Gaeta, "The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty", Neurosci. Biobehav. Rev., Vol. 25, No. 4, pp. 355-373, 2001.

[6] M. Comerchero and J. Polich, "P3a and P3b from typical auditory and visual stimuli", Clin Neurophysiol., Vol. 110, No.1, pp. 24-30, 1999.

[7] E. Niedermeyer, F. L. da Silva, *Electroencephalography, Basic Problems, Clinical Applications, and Related Fields*, Ed. 4, LW&W, 1999, pp. 623.

[8] S. Makeig, A. J. Bell, T. Jung and T. J. Sejnowski, "Independent Component Analysis of Electroencephalographic data" Advances in Neural Information Processing Systems 8, MIT Press, Cambridge MA, pp. 145-151, 1996.

[9] S. Makeig, T. Jung, A. J. Bell, D. Ghahremani and T. J. Sejnowski "Blind separation of auditory event-related brain responses into independent components" Proc. National Acadamy of Science, USA, Vol. 94, pp. 10979-10984, September 1997.

[10] A. J. Bell and T. J. Sejnowski, "An information-maximisation approach to blind separation and blind deconvolution" Neural Computation, Vol. 7, No. 6, 1004-1034, 1995.

[11] J. C. Mosher, P. S. Lewis and R. M. Leahy "Multiple Dipole Modelling and Localisation from Spatio-Temporal MEG Data" IEEE Transactions on Biomedical Engineering, Vol. 39, No. 6, pp. 541-557, 1992.

[12] J. C. Mosher and R. M. Leahy "EEG and MEG Source Localisation using Recursively Applied (RAP) Music" Proc. Thirtieth Annual Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, Nov 3-6, 1996.

[13] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing, Learning Algorithms and Applications* John Wiley & Sons, Ltd, 2002.

[14] B. Karoumi, A. Laurent, F. Rosenfeld, T. Rochet, A. M. Brunon and J. Dalery, "Alteration of event related potentials in siblings discordant for schizophrenia", Schizophr. Res., Vol 41, No. 2, pp. 325-334, 2000.