

# Recursive smooth noise reduction using subspace approach: application to evoked potentials

P. Celka<sup>1</sup>, Tim Cutmore<sup>2</sup>

<sup>1</sup> Griffith University, School of Engineering, Gold Coast, Australia

<sup>2</sup> Griffith University, School of Psychology, Brisbane, Australia

**Abstract**— Event related potentials are corrupted by randomness due to internal and external sources such as biological neural noise, electromagnetic noise, subject movement, and other electromagnetic origins from the subject's body. Averaging many trials is the common way to reduce these perturbation on the bases that the actual signal is kept unchanged from trial to trial and noise sources are uncorrelated with each other and with the signal of interest. It is the aim of this paper to present a method for reducing the undesirable noise sources using an iterative subspace projection methodology. This approach is well suited for online real-time processing such as biofeedback, neurofeedback and brain machine interfaces for rehabilitation purposes.

**Keywords**— Event related potentials, noise reduction, randomness, subspace, brain machine interface, neurofeedback

## I. INTRODUCTION

For several decades, scientists and engineers have developed strategies to reduce the noise from signals. Most of these noise reduction techniques use a particular subspace decomposition operator that maps the *signal+noise* space into a particular set of subspaces, where noise and signal components are expected to be as separated as possible in a given space [1-5]. Then, in each subspace, a weighting operator reduces the noise components. The weighting operation is often performed by thresholding of the components or by Wiener filtering. Finally, by inverse transform, the signal is reconstructed.

This paper presents a generic approach for reducing noise in signals using a subspace decomposition and a recursive computation of a posterior estimates [1]. By choosing the largest subspace coefficients from the posterior estimate, we can select the corresponding coefficients from the noisy signal and finally perform a smooth inverse reconstruction mapping that reduces the noise. The subspace decomposition method can be chosen in the most appropriate way depending on the statistical nature of the signals. The proposed method is particularly suited for low signal to noise ratio, rhythmical signals or signals acquired during

many-trials. This method has the advantages to be softly adjustable using a restricted number of parameters (2 or 3), and being of low-complexity.

In evoked brain potential studies, the many-trials averaging method for reducing the random fluctuations has been used for a long time. Many drawbacks of this methodology have been pointed out, thus urging for developing single-trial noise reduction methods. We present results of our method on visual brain evoked potentials and show how with very few trials we can obtain a good estimate of the information. Eventually, using fewer trials, neuroscientists and neuroclinicians will reach new understandings of the brain function dynamics.

## II. METHOD

The EEG laboratory in the School of Psychology conducted the experiment. A Neuroscan system with 32 EEG recording channels and Electrocap has been used to acquire the neural response measures. The standard 10-20 system placement was used. A shielded room (Faraday cage) houses a screen and a comfortable chair for the participant. The sampling frequency was set to 1000 Hz. Data were down-sampled at 128 Hz. 15 subjects have participated to the experiment. The experimental protocol was the oddball test with visual stimuli. The non-target stimuli were chosen to be the letters G, T, and A. The target (oddball) was the letter X. 15% of the total stimuli was the oddball. The number of trials  $L$  per subject was between 255 and 265 stimuli. The subject was asked to press the left mouse button for the oddballs, and he right for the non-oddballs. Both target and non-target were displayed randomly. Each stimuli lasted 2000 ms with 1000 ms of pre-stimulus. Each subject has provided and signed an informed consent sheet, and the experimental protocol has been approved by the ethical committee.

### III. SUBSPACE PROJECTION

Most of these noise reduction techniques use a particular subspace decomposition operator  $\mathbf{H}=\{H_k\}$  for  $k=1\dots M$ , that maps the *signal+noise* space  $S$  into a particular set of subspaces  $St_k$ , where noise and signal components are expected to be as separated as possible in a given space. Then, in each subspace  $St_k$ , a weighting operator  $P_k$  reduces the noise components. We will call the entire space weighting operator  $\mathbf{P}=\{P_k\}$ . The weighting operation is often performed by thresholding of the components in the space  $St_k$  or by Wiener filtering. The weighted components now span the spaces  $St_{c(k)}$ . Finally, by inverse transform  $\mathbf{H}^{-1}=\{H_k^{-1}\}$ , the signal is reconstructed in the *signal space*  $S_c$  in  $S$ .

A measurement signal  $x(n)$  ( $n$  is the discrete time index) in  $S$  can be written as:

$$x(n)=s(n)+e(n) \quad (1)$$

with  $n=0,\dots,N-1$  and  $s(n)$  is the information signal component we want to retrieve, and  $e(n)$  is the noise signal component. The following signals are generated through the noise reduction procedure

$$xt(n)=\mathbf{H}(x(n)); st(n)=\mathbf{P}(xt(n)); \text{ and } sh(n)=\mathbf{H}^{-1}(st(n)) \quad (2)$$

where  $sh(n)$  is the estimated information signal. The choice of the subspace operator  $\mathbf{H}$  is most often an open question and depends on the application, the underlying generating system linearity, the stationarity, the available computing and memory resources and the signal statistics. Generally, orthogonal filter banks, such as principal components filter banks or discrete orthogonal wavelet transforms, are preferred because of their numerous advantages such as perfect reconstruction of the signals, and fast implementation.

The use of the wavelet decomposition allows for a natural subband analysis of the signals under inspection and a good joint time-frequency decomposition. This implies that noise reduction using wavelet decomposition will lead to the best performance for non-stationary signals. The choice of the wavelet basis depends largely on the time and frequency behavior of the information we want to extract. Noise reduction using wavelets usually requires soft or hard thresholding on the wavelet coefficients. Determining the best threshold value is a matter of looking more closely at the signals at hand and the purpose of the noise reduction. We will use the discrete time wavelet subspace operator  $\mathbf{H}=\mathbf{W}$  in the following, but our approach can be implemented with any subspace decomposition.

Let us assume that we have at disposal an *a priori* signal from a model, or that we can build an *a posteriori* signal  $s^a$  which contains most of the time and frequency information of the signal  $s$ , and a measurement signal  $x$ . We transform both signals  $x$  and  $s^a$  through the subspace operator  $\mathbf{W}$

$$xt(n)=\mathbf{W}(x(n)) \text{ and } st(n)=\mathbf{W}(s^a(n)) \quad (3)$$

where the bold variables stand for the column vectors of samples from the corresponding signals' spaces. The coefficients' amplitude and positions in the vector  $st^a$  are very close to the one we should find in  $st$ , the wavelet transform of the signal  $s(n)$ . Quiroga *et al.* [2] has proposed to manually locate the most significant coefficients in  $st$ . The positions  $n_i$  of the selected coefficients  $st(n_i)$  are stored in a set  $I_M=\{n_i\}$  for  $i=1,\dots,K$ . Using this set of positions  $I_M$ , he proposed to locate the same coefficients in  $xt$  and and subsequently zeroed the other ones. This corresponds to our previously defined weighting matrix operator  $\mathbf{P}$  which can be formulated as:

$$\mathbf{P}=\text{Diag}(p(k)); k=1,\dots, N \quad (4)$$

with

$$p(k)=1 \text{ if } k \text{ is in } I_M; p(k)=0 \text{ else} \quad (5)$$

where  $\text{Diag}(\cdot)$  is a diagonal matrix. Inverse transforming  $\mathbf{P}xt$  results in the noise reduced vector:

$$sh=\mathbf{W}^{-1}\mathbf{P}xt \quad (6)$$

Note that this weighting matrix is different from a simple thresholding technique because of the user manual selection. On the one hand, manually selecting the positions  $I_M$  results in a supervised noise reduction and leads to optimal solutions in terms of expert knowledge. On the other hand, it is a drawback because of the subjectivity of the selection, the potential human mistakes from the 'getting bored' effect, and thus the lack of reproducibility. This method cannot be applied when a fully automatic noise reduction procedure is required. Also, this approach assumes that the signal  $s^a$  is very similar to  $s$  in the time and frequency domain, which may not always be the case.

We propose a flexible approach which enables the user to adjust the noise reduction to his needs by adjusting very few parameters. First we locate the most significant coefficients in  $st^a$  by a wavelet soft thresholding operator  $\mathbf{P}^{W_s}$ , resulting in the set of positions  $I_S$  for non zero coefficients. Second, we locate the coefficients in  $xt$  which correspond to the positions in  $I_S$  and we apply a FIR Gaussian filter  $G(k)$ , which is summarized in the  $\mathbf{P}^{W_x}$  operator. The Gaussian filtering allows the user to account for a possible time shift of the signal  $st$  compared to  $st^a$ , and is performed using a forward-backward approach for canceling its phase effect. This results in the following two-step nonlinear weightings, coefficient selection and filtering operations:

**First step:** The signal  $st^a$  is softly thresholded:

$$st^a_s=\mathbf{P}^{W_x}_s st^a \quad (7)$$

where the weighting operator is given by:

$$\mathbf{P}^{wx}_S = \text{Diag}(p^{ws}_S(k)) \quad (8)$$

and the soft thresholding nonlinear function expressed by

$$p^{ws}_S(k) = \text{Heav}(|st^a(k)| - T) (1 - T/|st^a(k)|) \quad (9)$$

where Heav is the Heaviside function,  $T$  is the threshold.

**Second step:** The resulting non zero coefficients' positions in  $st^a_S$  are stored in the set  $I_S$ . The set  $I_S$  is used for the selection of the coefficients in  $xt_S$  as finally expressed by the following equation:

$$xt_S = \mathbf{P}^{wx}_S xt \quad (10)$$

where the selection and filtering operator is given by:

$$\mathbf{P}^{wx}_S = G(k) X_c \text{Diag}(p^{wx}_S(k)) \quad (11)$$

with  $X_c$  the convolution operator, and

$$p^{wx}_S(k) = 1 \text{ if } k \text{ is in } I_S; \text{ and } 0 \text{ otherwise} \quad (12)$$

The selection operation expressed in (15) allows for the extraction of small information bearing signal components embedded in  $\mathbf{x}$ . Note that this selection procedure is different from hard or soft thresholding, and is nonlinear by nature. Finally, we apply a *soft mixing* of the coefficients  $st^a_S$  and  $xt_S$ :

$$xt_{Smooth} = a xt_S + (1-a) st^a_S \quad (13)$$

where  $a$  is the mixing parameter. The value of  $a$  reflects our confidence in the amplitude of the *a priori/posteriori* signal  $s^a$ . Determining  $a$  is the main issue in this approach and is left to the expertise of the user, which render this method flexible and easy to use. Yet, in the result section we present some insights about the performance of the method for different values of  $a$  which could help the choice of the user. The estimated information signal is recovered as follows:

$$sh = W^{-1} xt_{Smooth} \quad (14)$$

#### IV. RECURSIVE A POSTERIORI METHOD

In some applications, users have some knowledge about the nature of the system or signals under study. They may then use this knowledge to construct the signal  $s^a(n)$ . In evoked potential studies, the experiment is using several trials or repetitions. The way to build the signal  $s^a$  is to use a recursive approach where the user is making repetitive measurements  $s^a(n;i)$  on the system, with  $i=1, \dots, L$  being the repetition index [3]. The recursive *a posteriori* signal approach is performed by an autoregressive moving average procedure:

$$s^a_R(n;i) = d s^a_R(n;i-1) + (1-d) x(n;i) \quad (15)$$

The parameter  $0.5 < d < 1$  controls the recursive averaging memory. If  $d(i) = 1 - 1/i$ , (15) is a recursive estimate of the total average over all the  $L$  trials.

#### V. RESULTS

We have used the biorthonormal B-Spline wavelet of order (Nh=4, Ng=20) to perform the subspace projection  $W$ . This wavelet has been chosen for its close matching with the actual VEP waves and has also been used in previous studies [2].

##### A. Synthetic data

In order to validate our method, we have used simulated VEP. The synthetic VEP are based on a simple additive Gaussians [2] to represent the main VEP waves: P100, N200 and P300. The maximum peak values of each waves were randomly chosen in the following intervals IP1=[90,125] ms, IN2=[126,155] ms; IP3=[400,650] ms. In order to measure the quality of the noise reduction, we have used a normalized mean square error (see [3] for the definition).

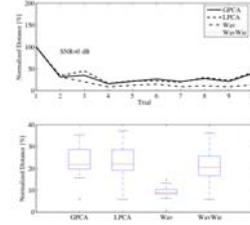


Fig. 1 Statistical results of the various noise reduction techniques

A background noise was added to the synthetic signal and constructed using surrogate method from real electroencephalogram signal. We have used the same sampling frequency as in the VEP measurements.

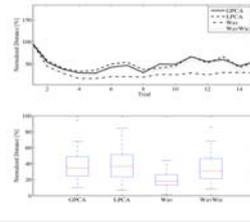


Fig. 2 Statistical results of the various noise reduction techniques on the first subject VEP

Fig. 1 (top panel) shows the results of our recursive approach (Wav) together with different noise reduction techniques: Local State Space Projection (LPCA) [3], Global State Space Projection (GPCA), a subband Wiener filtering (WavWie) and the presented smooth wavelet recursive scheme. For the last method, we have used  $a=0.5$  in (13) and a signal to noise ratio of 0 dB. For sake of comparison with the other noise reduction methods, we have performed a autoregressive averaging for each trials, such that the  $i$ th noise reduced trial is actually computed using a similar equation as (15) where  $s^a_R(n;i)$  is the  $i^{\text{th}}$  averaged noise reduced trial and  $x(n;i)$  is the  $i^{\text{th}}$  trial being processed. We can appreciate the effect of the a posteriori scheme (dotted-dashed line) by the constantly decreasing performance measure (expressed in %) as we have access to more trials. The other methods give a fluctuating behavior reflecting a 'non-learning' curve. Fig.1 (bottom panel) is showing the trial by trial performance measure (expressed in %) and shows a clear difference between the presented scheme and all the other ones.

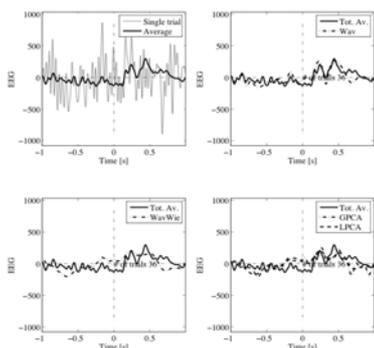


Fig. 3 Time series of the different noise reduction techniques on the first subject VEP (channel F3).

The noise reduced signals looks almost the same for all methods GPCA, LPC, Wav and WavWie. Actually the NMSE performance measure don't tells much about how the fine details of the VEP waves have been recovered: the NMSE is an average quantity across all the time span of the evoked potential. Specifically, it don't addresses the timing between the different waves, i.e. P100, N200, P300, neither their specific shapes. An other approach has been developed recently to address this issue [5].

### B. Visual evoked potential

During the oddball test, the visual pathway is excited with the visual external stimulus of the letters A, T, G and X appearing on the screen.

The oddball actually adds a surprise effect which tuns the VEP into more complex shapes both in the space and time domains. Fig. 2 shows the NMSE when we increase the number of trials. We have used  $a=0.5$  as in the simulations. These results show the same qualitative and quantitative behavior as our simulations.

In Fig.3, different waves are clearly present in the noise reduced signals, while impossible to detect in the single trial (top left panel).

## III. CONCLUSIONS

We have presented a recursive noise reduction scheme which make use of trials as they are recorded. This method make use of a subspace projection and nonlinear filtering in the dual space, together with the use of the a posteriori knowledge we have from the previously averaged trials. This method is particularly suited for online applications such as neurofeedback for rehabilitation of post-surgery patients or biofeedback.

## VI. ACKNOWLEDGMENT

We would like to thanks Griffith University through the GURG Grants for the financial support, and the Applied Cognitive Research Centre to have provided the research facilities to perform the recordings.

## REFERENCES

1. P. Celka (2005) Smoothly adjustable denoising using a priori knowledge. Signal Processing (in Press)
2. R. Q. Quiroga and H. Garcia (2003), "Single-trial event-related potentials with wavelet denoising", Clin. Neurophysiol., 114:376-390,
3. P. Celka et al. (2006), Dealing with randomness in biosignals, in Handbook of Time Series Analysis, eds Matthias Winterhalder, Bjorn Schelter, Jens Timmer, Wiley-VCH Verlag Berlin GmbH
4. A. Efferm et al.( 2000), "Nonlinear denoising of transient signals with application to event related potentials", Physica D, 140:257-266
5. M. Browne and T. Cutmore (2002), "Low-probability event-detection and separation via statistical wavelet thresholding: an application to psychophysiological denoising", Clin. Neurophysiol., 113:1403-1411
6. P. Celka and B. Kilner (2006), "Carmeli's S index assesses motion and muscle artefacts reduction in rower's electrocardiograms", Physiol. Meas. (in Press)

Address of the corresponding author:

Author: Patrick Celka  
 Institute: School of Engineering  
 Street: Parkland Drive  
 City: Southport, Queensland  
 Country: Australia  
 Email: p.celka@griffith.edu.au