

Estimation of Auditory Event-Related Potentials using a Combination of Principal Component Analysis and Kalman Filtering

Kevin Paulson¹, Othman Alfahad^{1,2}

¹School of Engineering and Computer Science, University of Hull
Hull, HU6 7RX, UK

k.paulson@hull.ac.uk; O.A.Aalfahad@2014.hull.ac.uk

²Biomedical Technology Department, Prince Sultan Military College of Health science
Dhahran, KSA

Abstract - This paper presents a method that aims to separate auditory event-related potentials (ERP) from noise. In practice, ERPs would be approximated by weighted sums of Principal Component (PCA) basis signals calculated from clean data. Projection of measured signals onto the PCA subspace greatly reduces noise. A second step uses Kalman filtering to optimally combine the PCA filtered signal with the ERP expected before measurement. Much of the power of the proposed algorithm comes from exploiting apriori cross-channel information in the form of a PCA weight covariance matrix. In this paper the performance of the method is quantified using synthetic multi-channel ERP signals to which known amounts of synthetic noise is added to all the channels. The use of synthetic data means and signal and noise are known and so signal-to-noise improvement can be determined. For a wide range of initial SNRs, PCA filtering increases SNR by 10 dB while Kalman filtering yields another 10 dB improvement.

Keywords: EEG, ERP Signal, Synthetic ERPs, PCA, Kaman Filter, SNR

1. Introduction

Auditory event-related potentials (ERPs) are contaminated by a variety of artefacts and noises, making it difficult to separate ERPs from other brain signals, biological signals such as muscle (EMG) and eye movements, artefacts due to electrode and equipment movement, and interference from other systems. In measurements from ambulatory or mobile EEG systems (MEEG), the artefacts are often even larger and more frequent than for a static system. However, there is increasing interest in the measurement of ERPs while individuals engage in the physical activities associated with normal living [1].

Typically, individual ERP measurements are too noisy to allow ERPs to be observed. ERP signals may have peaks of several micro-Volts (μV), while noise artefacts can have peaks of tens or hundreds of μV . Often, ERPs from many repeated measurements, on single or multiple individuals, are averaged before clinically important features can be observed. Important features include the positive and negative voltage peaks, often labelled N100, P200, P300 etc. [2]. A 20-electrode EEG system provides 19 channels of differential ERP measurement from different standard locations on the scalp [2], [3]. Each electrode measures the same brain reaction, but filtered due to the electrode position and the intervening tissue, plus noise. This document introduces a method that allows averaging across channels. Although each channel sees a different view of the ERP response, apriori knowledge of the signal correlations allows the channels to be combined to yield the best estimate of the underlying ERP.

Principal Component Analysis (PCA) has been applied to ERP signals before. PCA was used for reducing ocular artefacts in ERP signal by subtracting the principal component related to eye artefacts such as eye blinks, horizontal eye movements and vertical eye movements from raw EEG data [4], [5]. The Electrooculogram (EOG) signal is a large noise artefact in measured EEG data and for frontal electrodes has an amplitude much larger than brain signals. Reducing the EOG component of the measured EEG signal is vital when observing ERPs. The Casarotto et al. [5] PCA method yielded efficient and effective reduction of EOG artefacts. Similarly, Kobayashi and Kuriki [6] employed PCA to increase the signal-to-noise ratio (SNR) in evoked neuromagnetic signals applied to subjects. Using simulated evoked fields they demonstrated SNR improvement compared to the common averaging method.

Kalman filtering was developed by Kalman in 1960 for parameter estimation and has been widely applied to parameter tracking applications in many fields. It yields the maximum likelihood estimator given a priori and posteriori estimates of

parameter vectors, [7]. Kalman Filters (KF) have been applied to EEG time series. Kalman Smoothing has been applied to EEG signals to identify spikes associated with psychological diseases, [8]. Oikonomou [9] found that there is a significant enhancement in EEG SNR when using time-varying coefficients for an autoregressive model signal, estimated using Kalman filtering.

The paper begins by introducing a method to produce synthetic multi-channel ERP signals. These synthetic signals are used to quantify the performance of the proposed ERP estimation method. Section 3 describes the calculation of PCA bases for each channel ERP signal and the calculation of the covariance of channel PCA basis weights. The projection of ERP signals onto PCA bases, and the filtering it provides, is described in Section 4. Section 5 introduces Kalman filtering to provide the optimal combination of PCA filtered channel ERPs and the apriori expected multi-channel ERP. Finally, Sections 6 test the algorithms with synthetic signals and noises and this is followed by a conclusions section.

2. Multi-Channel ERP Simulation

When the brain responds to a stimulus, electrical signals move around the part of the brain involved in processing that stimulus. This could be modelled as a movement of charge or the movement, rotation and evolution in strength, of one or more electric dipoles. Electrodes attached to the scalp measure the electric potential on the surface due to this brain activity. The multi-channel simulator assumes that all channels measure the same underlying brain process, but filtered by the bone and tissue between the activity and the electrode on the scalp. Therefore, many features of the ERP response will be consistent across channels. For example, the P300 response is assumed to be measured at the same time on all channels. Similarly, the response amplitude is assumed to vary proportionately across all channels. It is the resulting correlation across channels that will be exploited in the proposed algorithm introduced below.

The ERP signal on each channel is approximated by a sum of Gaussian pulses:

$$E_i(t) = \sum_{j=1}^{N_G} A_{ij} f(t; t_j, \sigma_j) \quad (1a)$$

$$\text{where } f(t; t_j, \sigma_j) = \exp\left(-\frac{1}{2\sigma_j^2}(t - t_j)^2\right) \quad (1b)$$

E_i is the pure ERP signal measured on the i th channel, for a particular trial. For this project, three pulses were used i.e. $N_G = 3$; corresponding to the N100, P200 and P300 responses. The three parameters for each pulse : A_{ij} , t_j , and σ_j ; specify the amplitude, centre time and width of each of the Gaussian pulses respectively. For each trial, these parameters are determined from six independent samples from a Standard Normal distribution: $z_k \leftarrow N(0,1), k = 1, \dots, 6$.

$$A_{ij} \leftarrow \bar{A}_{ij} \times (1 + 0.1z_j), j=1,2,3 \quad (2a)$$

$$t_j \leftarrow \bar{t}_j + 10z_{j+3}, j=1,2,3 \quad (2b)$$

The 2D array \bar{A}_{ij} is preset with amplitudes that reflect the spatial-channel dependence of the responses. Each of the N100, P200 and P300 response amplitudes are assumed to vary independently, from trial to trial and from each other. However, for any trial, the amplitudes are consistent across channels. Similarly, the response peak times are assumed to vary independently but are the same for all channels. The mean peak times are: $\bar{t}_j = \{100, 200, 350\} ms$ and the peak widths are $\sigma_j = \{50, 30, 75\} ms$. These parameters have been chosen to match the ERP responses in [2]. For 3 Gaussians and 19 differential measurements, the mean peak amplitude array used is:

$$\bar{A}_{ij} = \begin{bmatrix} 18 & 17 & 53 & 85 & 110 & 90 & 55 & 85 & 170 & 190 & 180 & 60 & 54 & 160 & 100 & 95 & 52 & 12 & 11 \\ 18 & 17 & 53 & 85 & 110 & 90 & 55 & 85 & 170 & 210 & 180 & 60 & 54 & 160 & 120 & 95 & 52 & 12 & 11 \\ 30 & 40 & 53 & 85 & 160 & 130 & 55 & 85 & 190 & 220 & 180 & 60 & 54 & 160 & 200 & 155 & 52 & 20 & 18 \end{bmatrix} \times 10^{-2} \mu V$$

Figure 1 illustrates a typical synthetic multi-channel ERP response.

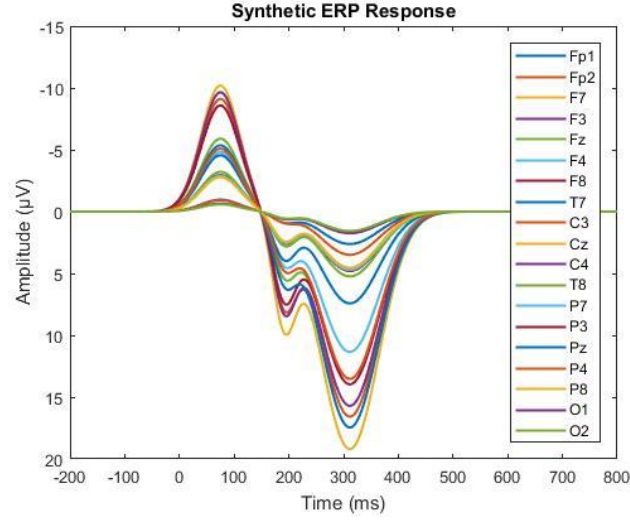


Fig. 1: Synthetic ERP Signal for 19 channels.

3. Principle Component Basis for Individual Channels

An ERP measurement on a single channel yields a discrete Voltage time series $\mathbf{E} \in \mathbb{R}^{N_s}$, of N_s samples. If a set of N_t ERP measurements is available, then an uncentred covariance matrix $\mathbf{C} \in \mathbb{R}^{N_s \times N_s}$ may be estimated. This matrix provides information on the joint probabilities of sample values. The Eigen vectors of the covariance matrix provide the principal component (PCA) basis signals. The Eigen values indicate how much of the variation between signals comes from components along PCA basis directions. Typically, the first few PCA basis signals span a large majority of the variation in ERP responses. Later basis signals span the noise. Projecting the ERP signal on the subspace spanned by the first few PCA basis signals will keep many of the features of the ERP signal while greatly reducing the noise. Let $\mathbf{B} \in \mathbb{R}^{N_s \times N_b}$ be a matrix whose N_b columns are the first N_b PCA basis vectors of length N_s samples. The first PCA basis vector is the mean ERP signal: $\bar{\mathbf{E}}$. Projecting the measurement vector onto the subspace spanned by the PCA basis vectors to get a vector of PCA basis weights $\mathbf{A} \in \mathbb{R}^{N_b}$ can be written in matrix notation as:

$$\mathbf{A} = \mathbf{B}^t \mathbf{E} \quad (3)$$

and the projected signal is:

$$\mathbf{E}_{pca} = \mathbf{B} \mathbf{A} = \mathbf{B} \mathbf{B}^t \mathbf{E}. \quad (4)$$

When the PCA weights are to be filtered, then a diagonal filter matrix $\mathbf{F} \in \mathbb{R}^{N_b \times N_b}$ may be introduced: $\mathbf{F} \equiv \text{diag}(f_i)$. The filter weights can be chosen to yield a smooth truncation to avoid Gibbs ringing e.g. $f_i = \exp\left(-\frac{(i - T_{pca})^2}{2}\right)$ where T_{pca} is the truncation basis number. The PCA filtered ERP signal may be written:

$$\mathbf{E}_{pca}^f = \mathbf{B} \mathbf{F} \mathbf{A} = \mathbf{B} \mathbf{F} \mathbf{B}^t \mathbf{E}. \quad (5)$$

If $\mathbf{E} \in \mathbb{R}^{N_s \times N_t}$ is a matrix whose N_t columns are ERP measures from N_t trials, then (5) yields $\mathbf{E}_{pca}^f \in \mathbb{R}^{N_s \times N_t}$ a matrix of the results after PCA filtering of each trial ERP measurement. A weighted sum of ERP measurements may be calculated by introducing a weight vector $\mathbf{W} \in \mathbb{R}^{N_t}$ where: $\mathbf{W} \equiv (w_i)$. The weighted mean ERP measurement may be written: $\bar{\mathbf{E}} = \mathbf{E} \mathbf{W}$, and the PCA filtered mean measurement is:

$$\bar{\mathbf{E}}_{pca}^f = \mathbf{B} \mathbf{F} \bar{\mathbf{A}} = \mathbf{B} \mathbf{F} \mathbf{B}^t \mathbf{E} \mathbf{W}. \quad (6)$$

Due to the associative property of matrix multiplication, $\mathbf{BFB}^t(\mathbf{E}\mathbf{W}) = (\mathbf{BFB}^t\mathbf{E})\mathbf{W}$, and so the weighted mean of the PCA filtered ERP responses is the same as the PCA filtered weighted mean ERP response.

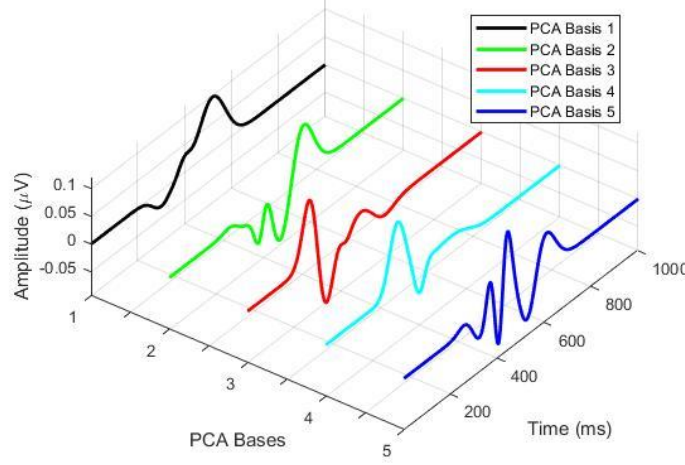


Fig. 2: Principle component Basis for channels.

3.1. Multi-Channel Information

Using the multi-channel ERP simulator introduced in Section 2, PCA bases can be calculated for each channel using the methods described in Section 3.1. As the underlying ERP signals have so much in common, the PCA bases and PCA projection weights \mathbf{A} will also have similarities. For a given trial j , the PCA projection weights for each channel i can be collected into a vector: $\mathbf{X}_j \equiv (\mathbf{A}_{ij})$, where \mathbf{A}_{ij} are the PCA weights for channel i in trial j . The vector $\mathbf{X}_j \in \mathbb{R}^{N_c N_b \times 1}$ where N_c is the number of channels. The collection of \mathbf{X} vectors for the N_t trials allows a mean and covariance matrix to be estimated:

$$\bar{\mathbf{X}} = \text{mean}(\{\mathbf{X}_j\}, j = 1, \dots, N_t) \quad (7)$$

$$\mathbf{C}_\mathbf{X} = \text{cov}(\{\mathbf{X}_j\}, j = 1, \dots, N_t) \quad (8)$$

The covariance matrix encapsulates all knowledge of the correlation between ERP components on all channels. The mean state vector is our *a priori* best estimate of the multi-channel ERP PCA basis weight vector.

4. ERP Estimation

This section follows the conventional notation of Kalman Filtering, applied to an individual PCA filtered ERP measurement i.e. an individual subject and trial. All the channels, as described by the vector \mathbf{X} , are estimated at the same time using an algorithm we call KFPCA. The estimation forms an optimal weighted sum of two sources of information: the *a priori* $\bar{\mathbf{X}}$ and the measured \mathbf{X} .

A priori, an ERP state vector is a sample from the multi-dimensional Normal distribution with mean $\bar{\mathbf{X}}$ and covariance $\mathbf{C}_\mathbf{X}$. In the absence of any further information, the maximum likelihood estimate of the ERP signal is that given by the mean vector $\bar{\mathbf{X}}$. In terms of Kalman Filters the $\bar{\mathbf{X}}$ and $\mathbf{C}_\mathbf{X}$ correspond to $\mathbf{X}_{k/k-1}$ and $\mathbf{P}_{k/k-1}$.

Given an ERP measurement on all channels, for a given subject and trial, a measured vector $\hat{\mathbf{X}}_j$ can be calculated using the processes in (3) and Section 3.1. The measurement has an uncertainty covariance matrix \mathbf{R}_j that is estimated from the difference between the truncated PCA estimate and the measurements, see Section 4.1. Given the *a priori* knowledge of

ERP and the information provided by the measurement process, an estimate of the particular ERP for this subject and trial \mathbf{X}_j , can be calculated by solving (9) for \mathbf{K} and then using (10):

$$\mathbf{K}(\mathbf{C}_X - \mathbf{R}_j) = \mathbf{C}_X, \quad (9)$$

$$\mathbf{X}_j = \hat{\mathbf{X}}_j + \mathbf{K}(\mathbf{X}_j - \bar{\mathbf{X}}) = \mathbf{C}_X \quad (10)$$

4.1. Estimation of Measurement Uncertainty

In order to implement Kalman Filtering it is necessary to develop an estimate of the PCA filtered measurement uncertainty in a single ERP measurement. If $\mathbf{A}_{\text{pca}}^f = \mathbf{F}\mathbf{B}^t\mathbf{E}$ are the filtered PCA measurement weights, then we need to estimate the measurement uncertainty matrix $\mathbf{C}_A \in \mathbb{R}^{N_b \times N_b}$ such that the probability density function for the actual ERP PCA weight vector, given the measurement, is multi-variate Normal with mean $\mathbf{A}_{\text{pca}}^f$ and covariance \mathbf{C}_A^f . It will be assumed that the noise in the ERP measurement is uniformly distributed across the unfiltered PCA basis weights, assuming a full PCA basis of rank N_s . Let n_i^2 be the noise power or variance in the i th PCA weight. Then, by Parseval's Theorem:

$$N^2 = \sum_{i=1}^{N_s} n_i^2 \quad (11)$$

where N^2 is the noise power in the measured ERP signal. This can be estimated by assuming the noise is close to the difference between the measured ERP signal and the PCA filtered signal:

$$N^2 \approx \|\mathbf{E}_{\text{pca}}^f - \mathbf{E}\|^2. \quad (12)$$

Given (11) and the assumption of uniform distribution of noise, then for all i :

$$n_i^2 = \frac{N^2}{N_s} \approx \frac{\|\mathbf{E}_{\text{pca}}^f - \mathbf{E}\|^2}{N_s}. \quad (13)$$

The covariance in the unfiltered PCA weights is then:

$$\mathbf{C}_A = n_i^2 \mathbf{I}_{N_b}. \quad (14)$$

where \mathbf{I}_{N_b} is the identity matrix of rank N_b , and the covariance of the PCA filtered ERP is:

$$\mathbf{C}_A^f = n_i^2 \text{diag}(w_i^2). \quad (15)$$

The multi-channel uncertainty matrix \mathbf{R}_j that is the block diagonal matrix formed from the individual channel uncertainty matrices i.e. $\mathbf{R}_j = \text{diag}(\mathbf{C}_{A_1}^f, \mathbf{C}_{A_2}^f, \dots, \mathbf{C}_{A_{N_c}}^f)$

5. SNR Performance

A simulated 20 channel MEEG system, with a 1 kHz sample rate, has been used to test the performance of the PCA and KFPCA filters. The channel PCA bases, mean PCA weights vector $\bar{\mathbf{X}}$ and weights covariance matrix \mathbf{C}_X were calculated using 100 simulated multichannel ERP signals. Each channel signal is samples from 200 ms before the stimulus to 800 ms after. In Monte Carlo tests, synthetic noise was added to the synthetic signals to simulate 100 ERP measurement trials for a range of noise scenarios. The ERP signal and synthetic noise was different for every trial. Three noise scenarios have been tested: additive white Gaussian noise (AWGN), additive pink Gaussian noise (APGN), and AWGN noise varying with channel. All noise was band limited to a maximum frequency of 15 Hz before SNR calculations. Signal and noise powers were defined to be the mean square amplitude of the 1000 sample signal sequences.

Figure 3 illustrates a single channel with band limited AWGN with an initial SNR=0 dB. The black curve is the synthetic ERP signal produced by the multi-channel simulator, and for these tests can be taken as the true, noise-free ERP signal. The red curve is the measured signal including synthetic noise. PCA filtering reduces the noise in the signal by removing PCA components that are largely noise. In this example the PCA components are smoothly truncated around the 10th basis signal using Gaussian weights. This signal is further combined with the expected signal using a Kalman Factor to yield the green curve. Although only one channel is illustrated, the algorithm mixes information from all channels. All channels had the same initial SNR although, as the channel ERP signals have different amplitudes, the noise amplitudes will also change between channels. Figure 4 shows the SNR improvement with filtering, starting from a range of initial SNRs. PCA filtering (solid lines) yield about a 10 dB SNR improvement just by rejecting components that are predominantly noise. KFPCA (dashed lines) yields a further improvement of about 7 dB for relatively clean signals and the improvement grows linearly for noisy signals where the measured signal is largely disregarded and the a priori expected signal is returned as the most likely response. Note that some of the noise reduction comes from the projection method forcing the ERP signal to be zero before the stimulus and after the expected ERP response. A shorter time span, say from -100 ms to 500 ms, would yield a smaller SNR improvement.

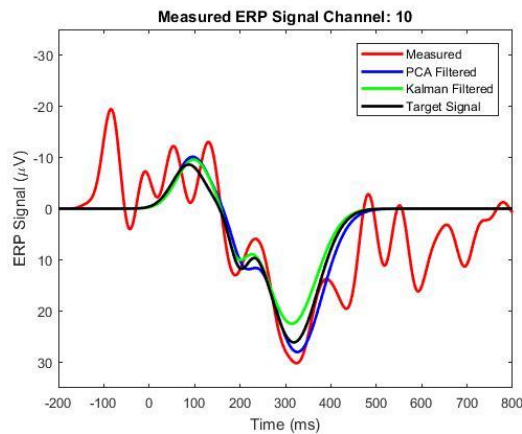


Fig. 3: Measured, filtered and true channel 10 ERP signal for band limited AWGN, SNR=0 dB.

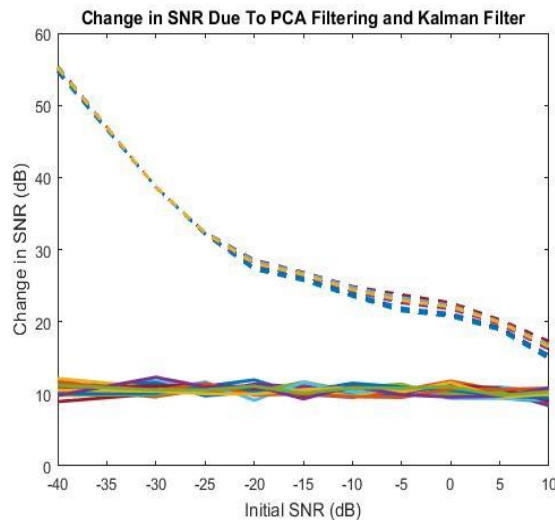


Fig. 4: Change in SNR Due to PCA and Kalman Filtering for AWGN, for all channels.

Figure 5 illustrates the results for the same experiment but using pink noise across all channels. Pink noise is much more challenging as it has more power at the lower frequencies which overlap more with the expected ERP signal. However, both PCA projection and Kalman filtering offer similar SNR improvements as with AWGN.

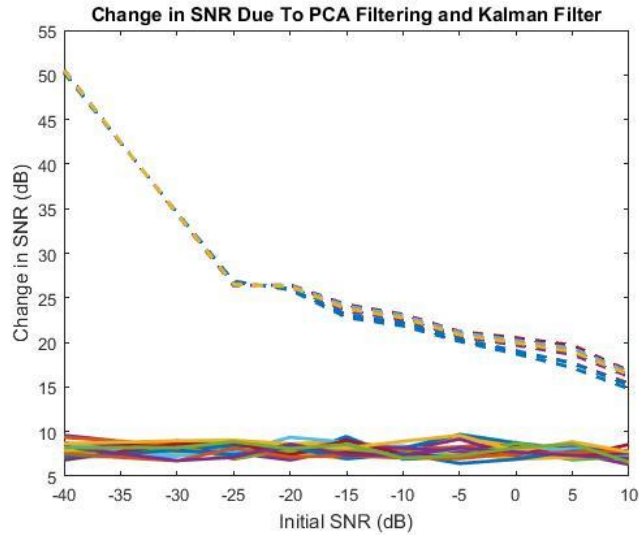


Fig. 5: Change in SNR Due to PCA and Kalman Filtering for APGN.

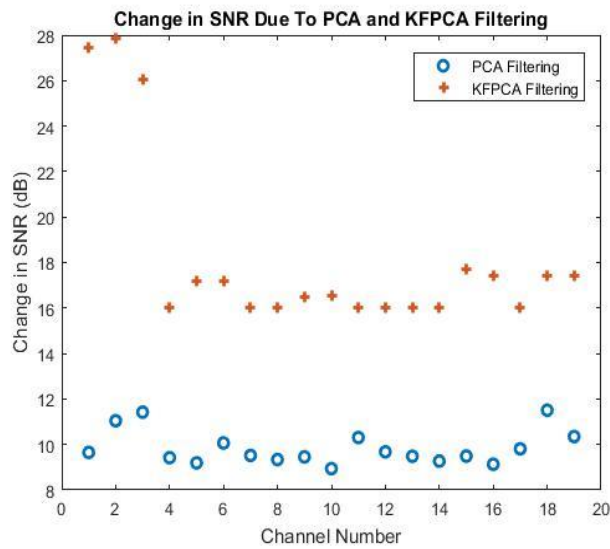


Fig. 6: Change in SNR Due To PCA and Kalman Filtering where channels 1, 2 and 3 have an initial SNR=0 while the other channels have an SNR=10.

Finally we consider the case where the initial SNR is not the same across all channels. This is typically the situation where noise is due to electrode connection problems, electrode movement or for biological signals such as eye blink artefacts. Eye blinks in particular effect the channels at the front of the head much more strongly than those further back. KFPCA filtering has the large advantage of utilizing expected correlations between all channel signals. Noisy channels are identified as part of the algorithm and the filtered signals on these channels are guided by the signals measured on the less noisy channels. This is illustrated in an example where channels 1, 2 and 3 (Fp1, Fp2 and F7) have initial SNR=0 dB while the other channels have SNR=10 dB. For this test, AWGN was used. PCA filtering yields the same SNR improvement on all channels as the same noise amplitude exists in all the PCA bases. However, Kalman Filtering uses the cross-channel information and yields much higher SNR improvement in the three noisy channels. In effect the algorithm reconstructs the signals on these channels from the more reliable information measured on the relatively clean channels.

6. Discussion and Conclusion

This paper presents a synthetic multichannel ERP simulator. The simulator has been used to demonstrate the SNR improvements produced by two denoising filter processes. Initially, PCA bases are calculated for each channel. In practice this would be done using a set of clean ERP signals derived from averages over many trials from many individuals. We have used synthetic ERP signals as these allow SNR improvements to be calculated. PCA filtering reduces artefact noise by projecting measured ERP signals onto low dimensional subspaces spanned by the first 10 principal component signals in each channel. A smooth truncation was used to reduce Gibbs ringing. In these tests, PCA projection increased SNR by about 10 dB for both white and pink Gaussian noise. In a second stage, the PCA projection weights were optimally combined with a priori weights using knowledge of the weights covariance and a Kalman factor. In effect, this uses information from all channels to reduce noise in each channel. The method was demonstrated in a scenario where three channels initially had 10 dB more noise than the others, to simulate eye blink artefacts. After KFPCA filtering, all channels had near the same SNR. This method provides a new way to interpolate missing channels that is much more sophisticated than methods based on weighted sums of adjacent channels. We propose to use KFPCA filtering on real data as a preprocessor, before combining trial outputs in a statistically optimised way to estimate ERP signals in the minimum number of trials.

Acknowledgements

This work was supported in part by the Saudi Arabia, Prince Sultan Military College of health sciences under Grant MD207.

References

- [1] A. D. Bateson et al, "Categorisation of Mobile EEG: A Researcher's Perspective," *BioMed Research International*, vol. 2017, 2017.
- [2] S. J. Luck, "An introduction to the event-related potential technique MIT press," *Cambridge, Ma*, pp. 45-64, 2005.
- [3] S. Debener et al, "How about taking a low-cost, small, and wireless EEG for a walk?," *Psychophysiology*, vol. 49, no. 11, pp. 1617-1621, 2012.
- [4] J. C. Costa et al, "Validation in principal components analysis applied to EEG data," *Comput. Math. Methods Med.*, vol. 2014, pp. 413801, 2014.
- [5] S. Casarotto et al, "Principal component analysis for reduction of ocular artefacts in event-related potentials of normal and dyslexic children," *Clin. Neurophysiol.*, vol. 115, no. 3, pp. 609-619, 2004.
- [6] T. Kobayashi and S. Kuriki, "Principal component elimination method for the improvement of S/N in evoked neuromagnetic field measurements," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 8, pp. 951-958, 1999.
- [7] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35-45, 1960.
- [8] V. P. Oikonomou et al, *The use of Kalman Filter in Biomedical Signal Processing*. INTECH Open Access Publisher, 2009.
- [9] V. P. Oikonomou, A. T. Tzallas and D. I. Fotiadis, "A Kalman filter based methodology for EEG spike enhancement," *Comput. Methods Programs Biomed.*, vol. 85, no. 2, pp. 101-108, 2007.