32nd Annual International Conference of the IEEE EMBS
Buenos Aires, Argentina, August 31 - September 4, 2010

Clustering Evoked Potential Signals Using Subspace Methods

A. M. Tomé, A.R. Teixeira, N. Figueiredo, P. Georgieva, I.M. Santos and E. Lang

Abstract— **This work proposes a clustering technique to analyze evoked potential signals. The proposed method uses an orthogonal subspace model to enhance the single-trial signals of a session and simultaneously a subspace measure to group the trials into clusters. The ensemble averages of the signals of the different clusters are compared with ensemble averages of visually selected trials which are free of any artifact. Preliminary results consider recordings from an occipital channel where evoked response $P100$ wave is most pronounced.**

I. Evoked potentials

Evoked potentials (EPs) represent transient components in the electroencephalogram (EEG) generated in response to a stimulus, e.g. a visual or auditory stimulus. Due to the small amplitude of evoked potentials, they are typically obscured by the the spontaneous activity of the brain (EEG). Although EP waves occur at lower frequencies than the background electroencephalogram (EEG), they are difficult to visualize in the overall single-trial scalp signal. To render the EP signal visible, a large number of single-trial responses are required to perform an ensemble average over all trials. However, this simple methodology has some drawbacks. The EP averaging may not cancel some artifacts induced by eye movements or blinks if they are time-locked to the experimental events. Thus it is common practice in EP studies to reject all EEG epochs contaminated with artifacts. But when the amount of data is limited, this constitutes an unacceptable procedure [1]. The averaging process has the further drawback of masking the single-trial variability of the task-related responses, e.g., in amplitude or latency. The fact that the same stimulus can elicit somewhat different signals has been evident since a few decades [2]. Despite these drawbacks, cognitive brain studies even consider a grand average of the ensemble averages obtained in several sessions with different patients to characterize main peaks. Early works, reporting signal processing manipulations applied to single trial signals, attempted to improve the entire waveform of an average EP. Artifact reduction in single-trials was another motivation to manipulate the single-trial signals [1]. But it was argued that the trial-to-trial variability in event-related activity implies that the simple ensemble average is not the optimal estimator for event-related potentials (ERP). For instance, Woody filtering is a matched filter technique for estimating latency of single-trials and aligning the signals to estimate an ensemble averaged signal [2]. Despite not being optimal, the ensemble average waveform can still be the goal or it can constitute a reference signal [3], [4] for further data manipulations. Different signal processing techniques like adaptive filtering, wavelets [5], subspace techniques like Principal Component Analysis (PCA) [1], Independent Component Analysis (ICA) [1], [3] and Blind Source Separation (BSS) [4], are known to be efficient for denoising or signal enhancement. Both wavelet and subspace techniques perform a decomposition of the signals into components that need to be selected and eventually combined to obtain the evoked response. PCA or ICA based methods can also differ in the methodology used to compute the subspace model. Some use the spacial information (e.g. the multichannel records)[1], [4] while others use temporal correlations like [3] which uses blocks of trials or temporal correlations within the single-trial signal [6]. The latter approaches as well as wavelet based methods [3] are then applied in parallel to each of the channels of the multichannel recording.

This work proposes a method that combines a subspace distance measure [7] with single spectrum analysis to obtain an enhanced version of the evoked potential signal. The subspace measure is used to cluster the trials so that the signals in the same clusters have similar characteristics while SSA provides signals with frequency profiles that represent different frequency bands of the input signal. The ensemble averages of the signals of the different clusters are compared with ensemble averages of visually selected trials which are free of any artifact.

II. Linear Subspace Models

Orthogonal subspace techniques can be applied to univariate sensor signals by forming a data matrix with shifted versions of the signal. These methods can be found in literature under distinct names depending on the domain of application: Singular Spectrum Analysis(for instance in climate time series analysis) [8] and SVD (for instance in speech enhancement) [9]. Considering a segment of an univariate signal $(x[0],x[1],\ldots,x[N-1])$, its multivariate counterpart is obtained by embedding the signal into its time-delayed coordinates forming, e.g., $\mathbf{x}_k = (x[k-1+M-1],\ldots,x[k-1])^T, k = 1,\ldots,K = N-1-M$. The lagged vectors lie in a space of dimension M, and constitute the columns of the data matrix $\mathbf{X}$, usually called *trajectory matrix*[8]. The non-normalized correlation matrix $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ allows the estimation of the orthogonal subspace model of the data via its eigen-decomposition $\mathbf{S} = \mathbf{U}\mathbf{D}\mathbf{U}^T$. The eigenvectors $\mathbf{U} = (\mathbf{u}_1 \cdots \mathbf{u}_M)$ form an orthonormal basis of the multidimensional space of the time-delayed coordinates. The data $\mathbf{X} = (\mathbf{x}_1 \cdots \mathbf{x}_K)$

A. M. Tomé, A.R. Teixeira, N. Figueiredo, P. Georgieva are with IEETA/DETI, Universidade de Aveiro, 3810-193 Aveiro, Portugal (email: ana@ua.pt).

E. W. Lang is with CIML Group, Biophysics, University of Regensburg, 93040 Regensburg, Germany.

I.M. Santos ,Ciências da Educação, Universidade de Aveiro, 3810-193 Aveiro, Portugal (email: isabel.santos@ua.pt)

can be projected onto selected eigenvectors and afterwards reconstructed to recover the original dimension

$$\hat{\mathbf{X}} = \mathbf{U}\mathbf{P}\mathbf{U}^T\mathbf{X} = \mathbf{U}\mathbf{P}\mathbf{Y} \quad (1)$$

The matrix $\mathbf{P}$ is a diagonal matrix that represents the selection process: with the m-th diagonal entry equal $p_{mm} = 1$ if the m-th row of $\mathbf{Y}$ is to be selected or equal to $p_{mm} = 0$ if it is to be removed. Inserting the subspace model $\mathbf{U} = (\mathbf{u}_1 \cdots \mathbf{u}_M)$ into eqn. (1) and using block manipulation operations yields

$$\hat{\mathbf{X}} = \mathbf{u}_1 p_{11}\mathbf{u}_1^T\mathbf{X} + \mathbf{u}_2 p_{22}\mathbf{u}_2^T\mathbf{X} + \ldots + \mathbf{u}_M p_{MM}\mathbf{u}_M^T\mathbf{X} \quad (2)$$

Each term on the right side is a rank-one matrix. And $\mathbf{u}_m^T\mathbf{X}$ is the m-th row of $\mathbf{Y}$. If all components are selected, i.e, $p_{mm} = 1 \,\forall\, m = 1,\ldots,M$ the reconstructed data matrix becomes $\hat{\mathbf{X}} = \mathbf{X}$.

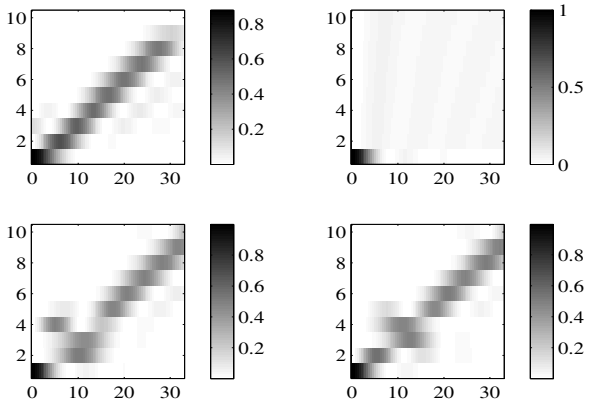

Fig. 1. Frequency response of filters computed with $M = 31$ of 4 trials. Ten filters cover the relevant frequency range $[0-30]HZ$ of EEG signals.

A. Extraction of Signal Components

In SSA or SVD methods, a reconstructed univariate time series $\hat{x}[n]$ is formed with the mean of the values along each descendent diagonal of $\hat{\mathbf{X}}$, i.e. by reverting the embedding in time-delayed coordinates. However, considering that this operation can be applied to every term on the right hand side of eqn. (2), the reconstructed time series can be written as

$$\hat{x}[n] = \sum_{m=1}^{M} p_{mm} a_m[n] \quad (3)$$

where p_{mm} can be also a scaling factor that changes the amplitude of components $a_m[n]$ [9] instead of being a simple selection term. Each sequence $a_m[n]$, $m = 1,\ldots,M$ is obtained via

- Computing $\mathbf{u}_m^T\mathbf{X}$ which is equivalent to the following convolution sum

$$y_m[n] = \sum_{i=1}^{M} u_{im} x[n-i+1] \quad (4)$$

where $(M-1) \le n \le N$ and $y_m[n]$ are the elements of the $m-th$ row of $\mathbf{Y}$ taken by their natural order. The components of the eigenvector $\mathbf{u}_m$ are the coefficients of the finite impulse response (FIR) filter.

- The product $\mathbf{u}_m\mathbf{Y}_m$ with diagonal averaging can also be written as a convolution sum

$$a_m[n] = \frac{1}{M_d}\sum_{i=l}^{s} u_{im} y_m[n+i-1] \quad (5)$$

where the values M_d, l and s have values according to the number of elements in the diagonals to be averaged and according to the time index. However in case of a steady state response, i.e. $(M-1) \le n \le (N-M)$ it holds $l = 1, s = M, M_d = M$.

Notice that the global system forms a cascade of a causal filter (eqn.4) with an anti-causal filter (5). The transfer functions $H_m(z), F_M(z)$ can be computed by substituting every delay step by a corresponding delay operator $z^{\pm d}$. For instance in eqn. (4) $x[n \pm d]$ is substituted by $z^{\pm d}X(z)$ and $y_m[n]$ by $Y_m(z)$ [10].

$$H_m(z) = \frac{Y_m(z)}{X(z)} = (u_{1m} + u_{2m}z^{-1} + \ldots u_{Mm}z^{-(M-1)}) \quad (6)$$

And the transfer function of the second filter of the cascade reads

$$F_m(z) = \frac{A_m(z)}{Y_m(z)} = \frac{1}{M}(u_{1m} + u_{2m}z^1 + \ldots + u_{Mm}z^{(M-1)}) \quad (7)$$

Notice that the transfer functions differ by a scale factor $1/M$ and by the sign of the powers of z. Therefore the magnitudes of the frequency response of both filters are related by the same scale factor $1/M$ and their phases are symmetric. Therefore the zero-phase global transfer function reads

$$T_m(z) = \frac{A_m(z)}{X(z)} = F_m(z)H_m(z) = \sum_{k=-(M-1)}^{M-1} t_{km} z^k \quad (8)$$

where $t_k = t_{-k}$, $k = 1,\ldots,(M-1)$ holds . Therefore, the frequency response $T_m(e^{j\omega})$ has the following expression [10]

$$T_m(e^{j\omega}) = t_{0m} + \sum_{k=1}^{M-1} 2t_{km}\cos(k\omega) \quad (9)$$

where $j = \sqrt{-1}$. The frequency response is a real function, with period equal $\omega = 2\pi$ (the normalized sampling rate), so corresponds to a zero-phase filter. This means that each extracted component $a_m[n]$ is then in-phase with its related original component $x[n]$. Fig. 1 represents the frequency responses of 10 filters computed using signals belonging to different trials. The filters $\mathbf{u}_m$, $m = 1,\ldots,10$ are ordered according to the corresponding values of the eigenvalues: $\mathbf{u}_1$ corresponds to the largest eigenvalue, $\mathbf{u}_2$ to the second largest and so on. The filters are centered so that they reflect the importance of that band in the signal. It can be seen that the chosen trials have different profiles. In the the first on top the frequency profile decreases with increasing frequency (f); the second on top is a trial with no signal (electrode contact problem); the ones on the bottom have clear peaks around $10Hz$ (alpha burst).

B. Comparing Subspaces

Signal enhancement is achieved by selecting a subspace of the M-dimensional space of time-delayed coordinates, i.e. by choosing a subset of the eigenvectors which span the M-dimensional data space. Suppose that a subspace $\mathscr{U}_A$ with $p < M$ eigenvectors is estimated for a single trial ERP and another $\mathscr{U}_B$ with $q < M$ eigenvectors is estimated for another single trial ERP. It will be interesting to know if there exists some similarity between these two subspaces. Considering two matrices ($\mathbf{U}_A$ and $\mathbf{U}_B$) with p and q eigenvectors spanning the two subspaces, then the following distance measure compares the two subspaces [7]

$$d(\mathscr{U}_A, \mathscr{U}_B) = \sqrt{max(p,q) - trace(\mathbf{U}_A^T \mathbf{U}_B \mathbf{U}_B^T \mathbf{U}_A)} \qquad (10)$$

The maximum distance corresponds to $\sqrt{max(p,q)}$ in case of orthogonal subspaces and the minimum is 0 when the subspaces coincide. So [7] considers the normalized distance defined as $d_{AB} = d(\mathscr{U}_A, \mathscr{U}_B)/\sqrt{max(p,q)}$ to establish a criterium to classify the models: if $d_{AB} \leq 1/2$, the two subspaces are similar otherwise not. For instance taking the subspaces formed with the $\mathbf{u}_m, m = 1,2,3$ of the previous example, the distance of the first model with the remaining models is: 0.76, 0.57 and 0.56 respectively.

C. Ensemble Average in Clusters

The subspace distance is used to cluster the single trial ERP signals and the ensemble average is performed for each cluster. Therefore adapting a subspace model for each single trial signal, the subspace distance is used to assign every single trial ERP to a particular cluster. The main steps of the algorithm are

- Initialization: compute a subspace model $\mathbf{U}_c$ randomly selecting a single trial EEG and assign $\mathbf{U}_c$ as the subspace model for cluster $c = 1$
- Repeat for all $t = 1, \ldots, T$ single trial ERPs $x_t[n]$.
  - Compute the subspace model for single trial ERP $x_t[n]$
  - Compute the normalized subspace distances d_{ct}, $\forall c$ clusters.
  - If
    - $\min(d_{ct}) < \alpha$, assign the single trial ERP to the closest cluster and update the ensemble average and the related subspace model.
    - $\min(d_{ct}) \geq \alpha$, create a new cluster $c \rightarrow c+1$ and its subspace model $\mathbf{U}_c = \mathbf{U}_t$
- Aggregate clusters into meta-clusters. Recursively group the clusters whose normalized subspace distances are $d_{AB} < \beta$.

The thresholds α and β as well as the dimension ($p = q$) of the subspace models must be assigned before the application of the algorithm.

III. Results and Discussion

The data set comprises 32 sessions (two sessions by participant) with around 250 trials per session. The experimental protocol and data acquisition is completely described in [11]. The stimulus consisted of overlapping pictures of faces and houses. The participant's task was to determine, during each trial, if the relevant stimulus (house or face, depending on the condition) had the same identity as the relevant stimulus presented on the previous trial, i.e., if it was the same house or the same person. Disregarding any eventual differences between conditions, early potentials like the P100 are clearly visible in ensemble averages , mostly in occipital derivations [11]. All subspace models used in the experimental section were computed using a similar strategy: the embedding dimension is $M = 31$, $N = 300$ ($B = 37$ before stimulus) samples used to estimate $\mathbf{S}$ and finally $p = 3$ filters/eigenvectors, corresponding to the largest eigenvalues, were used to form the subspace model.

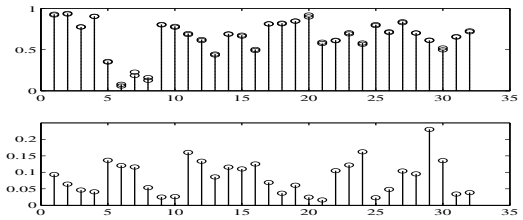

Fig. 2. Grand-average versus ensemble average: *top*: correlation coefficients (input and output); *bottom*: distance between subspace models.

A. Raw data and model adaption

The EEG was visually inspected and segments with excessive EOG artifacts were marked. The averages $x_s[n]$ of a session were estimated

$$x_s[n] = \frac{1}{T} \sum_{t=1}^{T} x_t[n] \quad n = -B, -B+1, \ldots, 0, \ldots, \qquad (11)$$

with all T single trial ERPs considered free of artifacts. The signals $x_s[n]$ were averaged finally to form the Grand-average $x[n]$. A model was computed for x[n] and a subspace model with $p = 3$ provides a smoother version of the signal. Subspace models were also computed for each $x_s[n]$. The fig. 2 shows the result of using two different comparison measures: on top the correlation coefficient is shown (see [3]) and below the normalized distance (see 2). The former is a comparison between the $x_s[n], s = 1 \ldots 32$ and $x[n]$. The latter is a comparison between models computed with the signals.

And globally both comparisons give a similar indication, i.e. when the correlation coefficient is high, the subspace distance is low with the exception of the signal corresponding to session 29 that exhibits the largest subspace distance and a large correlation coefficient. Comparing the most strongly correlated signal with the Grand average signal $x[n]$ reveals that after the stimulus $n > 0$ both signals have a very similar shape except for the peak value occurring around $0.1s$, called the $P100$ (upper trace, see fig. 3). For the signal corresponding to the minimal correlation coefficient the positive peak is less relevant than the following negative peak, called $N170$ (lower trace, fig. 3). Concerning subspace

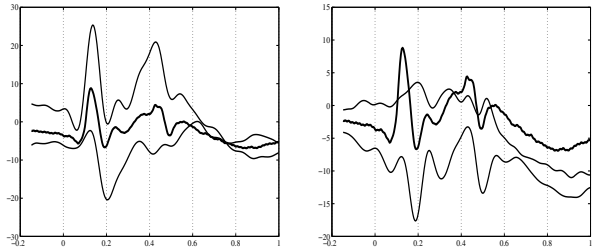

Fig. 3. Grand-average (thick trace) and ensemble average signals: *left*-trials with largest (top) and smallest(bottom) correlation coefficient ; *right* largest (top) and smallest (bottom) subspace distance.

distances, for the signal with the maximum distance the session signal corresponds to $29th$ and the waveform after the stimulus has a slower increase up to the maximum when compared to the Grand-average (upper trace, see fig. 3).

B. Ensemble average in clusters

The algorithm was applied to all the single trial ERPs of a session (including those which were rejected by visual inspection). The clustering was performed using $\alpha = 0.5$ and aggregation was done with $\beta = 0.1$. Comparing the ensemble average of the cluster which contains the largest number of single trial EEG signals with the ensemble average formed with the visually selected single trial EEG signals taken from the same session, results similar to fig. 2 are achieved. The largest correlations are for the same group of signals as well as the minimum normalized distances. Concerning the correlation coefficients of most of the signals, the minimal correlation coefficient corresponds to differences between the signals occurring after $200ms$ (see fig 4 - top left). In case of maximal correlation coefficients, the signals practically coincide (see fig 4 - bottom left). With the distance measure, the signals also show noticeable differences only after $200ms$ (see fig 4 - top right). Fig. 5 shows the ensemble averages of

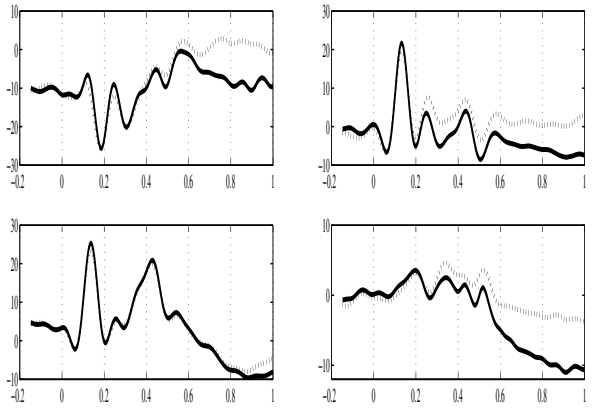

Fig. 4. Evoked potentials:selected trials (dot) versus cluster ensemble average (line).*left*:correlation coefficient; *right*: normalized distance

the 3 meta-clusters for two sessions. Note that the subspace distance of the 3rd meta-cluster might be even larger than the threshold distance β as it collects all remaining signals. Remarkably, in session 32 (fig. 5, right) a clear $P100$ response is visible for all clusters. This seems to be observable only for strong $P100$ responses. The rhythmic activity visible in session 8 (fig. 5, left) seems to be related with the α-rythmn of the proband.

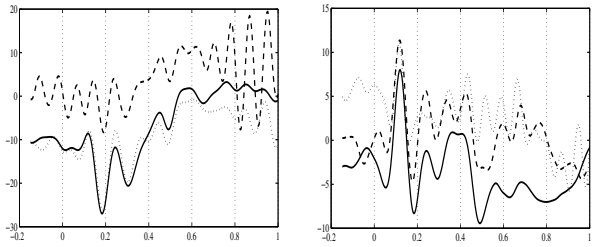

Fig. 5. Ensemble averages of the clusters: largest cluster (line), second largest (dash), and last (dot). The number of trials in each cluster :on the *left*- $\{137, 84, 10\}$; on the *right*-$\{211, 22, 12\}$.

IV. ACKNOWLEDGMENTS

A.R. Teixeira and N. Figueiredo are supported by the Portuguese Foundation for Science and Technology (FCT), PhD scholarships (SFRH/BD/28404/2006) and (SFRH/BD/48775/2008), respectively.

REFERENCES

[1] T.-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clinical Neurophysiology*, vol. 111, no. 10, pp. 1745–1758, 2000, doi: DOI: 10.1016/S1388-2457(00)00386-2.

[2] C. D. McGillem, J. I. Aunon, and C. A. Pomalaza, "Improved waveform estimation procedures for event-related potentials," *Biomedical Engineering, IEEE Transactions on*, vol. BME-32, no. 6, pp. 371–379, 1985.

[3] D. Iyer and G. Zouridakis, "Single-trial evoked potential estimation: Comparison between independent component analysis and wavelet denoising," *Clinical Neurophysiology*, vol. 118, p. 495504, 2007.

[4] S. Hajipour, M. B. Shamsollahil, H. Mamaghanian, and V. Abootalebi, "Extracting single trial visual evoked potentials using iterative generalized eigen value decomposition," in *IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2008.*, 2008, pp. 233–237.

[5] R. Q. Quiroga and H. Garcia, "Single-trial event-related potentials with wavelet denoising," *Clinical Neurophysiology*, vol. 114, no. 2, pp. 376–390, 2003, doi: DOI: 10.1016/S1388-2457(02)00365-6.

[6] M. Z. Yusoff, N. Kamel, and A. F. M. Hani, "Single-trial extraction of visual evoked potentials from the brain," in *16th European Signal Processing Conference (EUSIPCO-2008)*, Lausanne, 2008.

[7] X. Sun, LiweiWang, and J. Feng, "Further results on the subspace distance," *Pattern Recognition*, vol. 40, p. 328 329, 2007.

[8] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure: SSA and Related Techniques*. Chapman & HALL/CRC, 2001.

[9] P. C. Hansen and S. H. Jensen, "Subspace-based noise reduction for speech signals via diagonal and triangular matrix decompositions: Survey and analysis," *Eurasip Journal on Advances in Signal Processing*, vol. Vol 2007, 2007.

[10] L. B. Jackson, *Digital Filters and Signal Processing*. Kluwer Academic Publishers, 1996.

[11] I. M. Santos, J. Iglesias, E. I. Olivares, and A. W. Young, "Differential effects of object-based attention on evoked potentials to fearful and disgusted faces," *Neuropsychologia*, vol. 46, no. 5, pp. 1468–1479, 2008.