Greedy KPCA in Biomedical Signal Processing

Ana Rita Teixeira[1], Ana Maria Tomé[1], and Elmar W. Lang[2]

[1] DETI/IEETA, Universidade de Aveiro, 3810-193 Aveiro, Portugal
ana@ieeta.pt
[2] Institute of Biophysics, University of Regensburg, D-93040 Regensburg, Germany
elmar.lang@biologie.uni-regensburg.de

Abstract. Biomedical signals are generally contaminated with artifacts and noise. In case artifacts dominate, the useful signal can easily be extracted with projective subspace techniques. Then, biomedical signals which often represent one dimensional time series, need to be transformed to multi-dimensional signal vectors for the latter techniques to be applicable. In this work we propose the application of a greedy kernel Principal Component Analysis(KPCA) which allows to decompose the multidimensional vectors into components, and we will show that the one related with the largest eigenvalues correspond to an high-amplitude artifact that can be subtracted from the original.

1 Introduction

In many biomedical signal processing applications a sensor signal is contaminated with artifact related signals as well as with noise signals of substantial amplitude. The former sometimes can be the most prominent signal component registered, while the latter is often assumed to be additive, white, normally distributed and non-correlated with the sensor signals. Often signal to noise ratios (SNR) are quite low. Hence to recover the signals of interest, the task is to remove both the artifact related signal components as well as the superimposed noise contributions. Projective subspace techniques can then be used favorably to get rid of most of the noise contributions to multidimensional signals. But many biomedical signals represent one dimensional time series. Clearly projective subspace techniques are not available for one dimensional time series to suppress noise contributions, hence time series analysis techniques often rely on embedding a one dimensional sensor signal in a high-dimensional space of time-delayed coordinates [6], [16]. As embedding is a non-linear signal manipulation [11] Kernel Principal Component Analysis (KPCA) should be a suitable technique. KPCA can simultaneously retain the non-linear structure of the data while denoising is achieved with better performance because the projections are accomplished in the higher-dimensional feature space. The KPCA method represents a projective subspace technique applied in feature space which is created by a non-linear transformation of the original data. In the feature space a linear principal component analysis is performed. Denoising is achieved by considering the projections related to the largest eigenvalues of the covariance matrix.

The mapping into feature space is avoided by using kernel functions which implicitly define a dot product in feature space computed using data in input space [10]. The kernel matrix (a dot product matrix) of the mapped data is easily computed and naturally its dimension depends on the size of the data set and the transformation. The

J. Marques de Sá et al. (Eds.): ICANN 2007, Part II, LNCS 4669, pp. 486–495, 2007.
© Springer-Verlag Berlin Heidelberg 2007

entries $k(i, j)$ of the matrix depend on the corresponding data points and are computed according to the defined kernel function. The size of the kernel matrix represents a computational burden once its eigendecomposition must be achieved. In practice, the goal of projective subspace techniques is to describe the data with reduced dimensionality by extracting meaningful components while still retaining the structure of the raw data. Then only the projections on the directions corresponding to the most significant eigenvalues of the kernel (or covariance matrix) need to be computed. The exploitation of methods like the well-known Nyström method to achieve a low rank eigendecomposition is a strategy that has been considered in [4],[15]. Furthermore these techniques can also achieve a solution without the manipulation of the full matrix. We show how Nyström's method can be applied to KPCA leading to what is usually known as greedy KPCA. In this work we reformulate both KPCA and greedy KPCA under a unifying algebraic notation underlying the differences between both approaches. And we exploit such greedy techniques to extract high-amplitude artifacts from EEG recordings. The extracted artifact is then subtracted from the original recording thus obtaining a corrected version of the original signal.

EEGs are generally distorted by signals generated by eye movements, eye blinking, muscle activity, head movements, heart beats and line noise. Particularly, eye movements and blinking are major sources of EEG contamination. These ocular movement based signals (EOG) are of larger amplitude than cortical signals (EEG). As they propagate over the scalp, they are recorded in most EEG derivations. Especially the frontal channels often show prominent EOG artifacts which obscure the underlying EEG signals. The availability of digitalized EEG signals makes possible the application of more sophisticated techniques than simple linear filtering. More recently independent component analysis (ICA) [14], [9], blind source separation [8] or adaptive filtering techniques [7] have been discussed. The most recent works use independent component analysis: [9] used the INFOMAX algorithm [13], [17] applied the joint approximative diagonalization of eigen-matrices algorithm (JADE), in [8] an approximate joint diagonalization of time-delayed correlation matrices (SOBI) was used while in [14] the fast fixed point algorithm (FastICA) has been applied. In all the works except [13], the EOG channels are included in the processed data set. In this work we propose the application of a kernel technique to the multidimensional signal obtained by embedding a single univariate EEG signal in its time-delayed coordinates. And after reconstruction and reverting the embedding, the artifact is isolated. The corrected EEG is obtained subtracting the artifact from the original signal. Naturally, this method can be applied also in parallel to a subset of channels.

2 The Kernel Greedy Approaches to Decompose Univariate Signals

Time series analysis techniques often rely on embedding one dimensional sensor signals in the space of their time-delayed coordinates [11]. Embedding can be regarded as a mapping that transforms a one-dimensional time series $x = (x[0], x[1], ..., x[N-1])$ to a multidimensional sequence of $K = N - M + 1$ lagged vectors

$$\mathbf{x}_k = [x[k-1+M-1], \ldots, x[k-1]]^T, \; k = 1, \ldots, K \tag{1}$$

The lagged vectors $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_K]$ lie in a space of dimension M.

Kernel Principal Component Analysis (KPCA) relies on a non-linear mapping of given data to a higher dimensional space, called feature space. Without loosing generality, let's assume that the data set is centered and split into two parts yielding the mapped data set

$$\mathbf{\Phi} = [\phi(\mathbf{x}_1)\phi(\mathbf{x}_2)\ldots\phi(\mathbf{x}_R), \phi(\mathbf{x}_{R+1})\ldots\phi(\mathbf{x}_K)] = [\,\mathbf{\Phi}_R\ \mathbf{\Phi}_S] \tag{2}$$

In denoising applications, the first step of KPCA is to compute the projections of a mapped data set onto a feature subspace. Considering L eigenvectors (columns of $\mathbf{U}$) of a covariance matrix (a correlation matrix if the data is centered) corresponding to the L largest eigenvalues, the projections of the mapped data set $\mathbf{\Phi}$ are

$$\mathbf{Z} = \mathbf{U}^T\mathbf{\Phi} \tag{3}$$

The columns of the matrix $\mathbf{U}$ form a basis in feature space onto which the data set is projected. This basis can be written as a linear combination of the mapped data

$$\mathbf{U} = \mathbf{\Phi}_B\mathbf{A} \tag{4}$$

The matrix $\mathbf{A}$ is a matrix of coefficients and either $\mathbf{\Phi}_B = \mathbf{\Phi}$ (KPCA) or $\mathbf{\Phi}_B = \mathbf{\Phi}_R$ (greedy KPCA), representing a subset of the data set only. Note that the column j of $\mathbf{Z}$ depends on the dot products $\mathbf{\Phi}_B^T\phi(\mathbf{x}_j)$. However to avoid an explicit mapping into feature space, all data manipulations are achieved by dot products [10] and the kernel trick is applied. For instance, using an RBF kernel, the dot product between a vector $\phi(\mathbf{x}_i)$, with $i \in B$, and $\phi(\mathbf{x}_j)$ is computed using a kernel function that only depends on the input data

$$k(\mathbf{x}_i, \mathbf{x}_j) = exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}) \tag{5}$$

Finally, to recover the noise-reduced signal after denoising in feature space, the non-linear mapping must be reverted, i.e. the pre-image in input space of every signal, denoised and reconstructed in feature space, must be estimated. Denoising using KPCA thus comprises two steps after the computation of the projections in the feature space: a) the reconstruction in feature space and b) the estimation of the pre-image of the reconstructed point $\hat{\phi}(\mathbf{x}_j) = \mathbf{U}\mathbf{z}_j$, where $\mathbf{z}_j$ represents the projections of a noisy point $\mathbf{x}_j$. These two steps can be joined together by minimizing the Euclidian distance of the image $\phi(\mathbf{p})$ of a yet unknown point $\mathbf{p}$ from $\hat{\phi}(\mathbf{x}_j)$

$$\begin{aligned}\tilde{d}^{(2)} &= \|\phi(\mathbf{p}) - \hat{\phi}(\mathbf{x}_j)\|^2 \\ &= (\phi(\mathbf{p}) - \hat{\phi}(\mathbf{x}_j))^T(\phi(\mathbf{p}) - \hat{\phi}(\mathbf{x}_j))\end{aligned} \tag{6}$$

The central idea of the fixed-point method [10] consists in computing the unknown pre-image of a reconstructed point in the projected feature subspace by finding a $\mathbf{p}$ which minimizes that distance (eq.6). If an RBF kernel is considered, the iterative procedure is described by the following equation

$$\mathbf{p}_{t+1} = \frac{\mathbf{X}_B(\mathbf{g}\diamondsuit\mathbf{k}_{\mathbf{p_t}})}{\mathbf{g}^T\mathbf{k}_{\mathbf{p_t}}} \tag{7}$$

where $\diamondsuit$ represents a Hadamard product, $\mathbf{g} = \mathbf{A}\mathbf{z}_j$. The components of the vector $\mathbf{k}_{\mathbf{p_t}} = \mathbf{k}(\mathbf{X}_B, \mathbf{p}_t)$ are given by the dot products between $\phi(\mathbf{p}_t)$ and the images $\mathbf{\Phi}_B$ of the training subset $\mathbf{X}_B$. The algorithm must be initialized and $\mathbf{p}_0 \equiv \mathbf{x}_i$ is a valid choice [11]. The points $\mathbf{p}_k$ then form the columns of $\hat{\mathbf{X}}$, the noise-free multidimensional signal matrix in input space. The one-dimensional signal, $\hat{x}[n]$, is then obtained by reverting the embedding, i.e. by forming the signal with the mean of the values along each descendent diagonal of $\hat{\mathbf{X}}$ [12]. Note that if $\hat{x}[n]$ corresponds to the high amplitude artifact, then the corrected signal is computed as $y[n] = x[n] - \hat{x}[n]$.

2.1 Computing the Basis

The projections $\mathbf{Z}$ of the training set are also related with the eigenvectors of a matrix computed using only dot products ($\mathbf{K}$), the kernel matrix. Naturally the entries of the latter matrix can be easily computed using the kernel trick thus avoiding an explicit mapping of the data set. Furthermore, considering a singular value decomposition of the data set

$$\mathbf{\Phi} = \mathbf{U}\mathbf{D}^{1/2}\mathbf{V}^T \tag{8}$$

where D is a diagonal matrix with its non-zero eigenvalues of the kernel matrix (or of the covariance matrix) ordered according to $\lambda_1 > \lambda_2 > ... > \lambda_L ... > \lambda_R$ and $\mathbf{V}$ and $\mathbf{U}$ are the R eigenvectors of the kernel and covariance matrices, respectively. The data set can be approximated using an SVD decomposition with the L most significant singular values and the corresponding eigenvectors. Then substituting the SVD decomposition (eq.8) in eqn. (3) the L projections for each element of the data set read

$$\mathbf{Z} = \mathbf{D}^{1/2}\mathbf{V}^T \tag{9}$$

The two approaches, KPCA and greedy KPCA, respectively, arise from two distinct strategies to deal with the eigendecomposition of the kernel matrix ($\mathbf{K}$) of the data set. In KPCA, where the whole data set is used to compute the kernel matrix, matrix $\mathbf{A}$ is computed using the largest eigenvalues $(\mathbf{D})$ and corresponding eigenvectors $(\mathbf{V})$. And by manipulation of eqns. (3) and (9), the basis vector matrix is obtained as

$$\mathbf{U} = \mathbf{\Phi}\mathbf{V}\mathbf{D}^{-1/2} \tag{10}$$

In greedy KPCA a low-rank approximation of the kernel matrix is considered. This leads to the eigendecomposition of matrices with reduced size. Considering that the training set was divided into two subsets, the kernel matrix can be written in block notation [15],[4]

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_r & \mathbf{K}_{rs} \\ \mathbf{K}_{rs}^T & \mathbf{K}_s \end{bmatrix} \tag{11}$$

where $\mathbf{K}_r$ is the kernel matrix within subset $\mathbf{\Phi}_R$, $\mathbf{K}_{rs}$ is the kernel matrix between subset $\mathbf{\Phi}_R$ and $\mathbf{\Phi}_S$ and $\mathbf{K}_s$ is the kernel matrix within the subset $\mathbf{\Phi}_S$. The approximation is written using the upper blocks of the original matrix [15], [4]

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{K}_r \\ \mathbf{K}_{rs}^T \end{bmatrix} \mathbf{K}_r^{-1} \begin{bmatrix} \mathbf{K}_r \; \mathbf{K}_{rs} \end{bmatrix} \tag{12}$$

It can be shown that the lower block is approximated by $\mathbf{K}_s \approx \mathbf{K}_{rs}^T \mathbf{K}_r^{-1} \mathbf{K}_{rs}$. The R eigenvectors $\mathbf{V}$ corresponding to the R largest eigenvalues are then computed as

$$\mathbf{V}^T = \mathbf{H}^T \left[\mathbf{K}_r \; \mathbf{K}_{rs} \right] = \mathbf{H}^T \mathbf{\Phi}_R^T \left[\mathbf{\Phi}_R \; \mathbf{\Phi}_S \right] \tag{13}$$

where $\mathbf{H}$ can be computed following different strategies [15], [4] and [1]. The latter [1] refers to an incomplete Cholesky decomposition $\tilde{\mathbf{K}} = \mathbf{C}^T \mathbf{C}$, with

$$\mathbf{C} = \left[\mathbf{L} \; \mathbf{L}^{-T} \mathbf{K}_{rs} \right] \tag{14}$$

The matrix $\mathbf{L}$ is a triangular matrix corresponding to the complete Cholesky decomposition of $\mathbf{K}_r = \mathbf{L}^T \mathbf{L}$.

The $R \times R$ matrix $\mathbf{Q} = \mathbf{C}\mathbf{C}^T$ and its eigendecomposition $\mathbf{V}_q \mathbf{D} \mathbf{V}_q^T$ are used to obtain the low rank approximation of the kernel matrix. In particular the eigenvectors (see eq. 13) are computed with $\mathbf{H} = \mathbf{L}^{-1} \mathbf{V}_q \mathbf{D}^{-1/2}$. The eigenvectors are orthogonal ($\mathbf{V}^T \mathbf{V} = \mathbf{I}$) and consequently the data, in the feature space, is represented by non-correlated projections ($\mathbf{Z}$). The manipulation of equations (13), (9) and (3) gives the basis vector matrix

$$\mathbf{U} = \mathbf{\Phi}_R \mathbf{L}^{-1} \mathbf{V}_q \tag{15}$$

The eigenvectors in the matrix $\mathbf{V}_q$ should be placed according to their corresponding eigenvalues. The first column should have the eigenvector corresponding to the largest eigenvalue and so on. Furthermore the matrix should have $L < R$ columns to enable projections of the data onto the directions related to the L largest eigenvalues. Note that the pivoting index of the incomplete Cholesky decomposition [1] leads to the selection of $\mathbf{\Phi}_R$ within the training set. Further note that in any practical implementation the data is not centered in the feature space. However their centering can be achieved in matrix $\mathbf{C}$, i.e. , after the incomplete Cholesky decomposition of the kernel matrix.

2.2 Choosing Training and Testing Subsets

In the last section it was discussed that with an incomplete Cholesky decomposition it is possible to compute the parameters of the model using a training data set divided into two subsets. But the parameters usually depend on both subsets. However, there are approaches, where the parameters depend on only one of the subsets [15] whose elements are chosen randomly. In this application, we consider an hybrid approach which leads to the choice of three subsets of data. We start by splitting the data (with J vectors) into two data sets: the training set with K vectors and the testing set which contains the remaining data to be processed. In this application, we consider to form the training set with two strategies

- choosing the K vectors randomly.
- choosing the K vectors that correspond to a subsegment (with artifact) of the segment to be processed.

The subset R of the training set is chosen using a Cholesky decomposition performed with the symmetric pivoting algorithm [1]. The methodology is based on the minimization of the trace $tr(\mathbf{K}_s - \mathbf{K}_{rs}^T \mathbf{K}_r^{-1} \mathbf{K}_{rs})$ applied to iteratively update the subset R. So by identifying the maximum value of the trace operator (the pivot), an element of subset S is moved to the subset R and the matrix $\mathbf{K}_r$ increases its size while the others decrease. The process stops when the trace of the matrix corresponding to the actual approximation is less than a threshold [1], [5] and/or the matrix $\mathbf{K}_r$ is not well conditioned [2].

3 Numerical Simulations

The method described in the previous section will be applied to the removal of prominent EOG artifacts from EEG recordings. The corrected EEG is obtaining by subtracting the artifact from the original channel.

Two experiments will be discussed: one dealing only with the details of the application of the method to a single channel; the other application illustrates the use of the algorithm to ease visual inspection of critical segments. The algorithm is always applied to segments of $12s$ or $10s$ (typical window size on displays) duration. The multidimensional signal is obtained with an embedding into an $M = 11$ feature space. An RBF

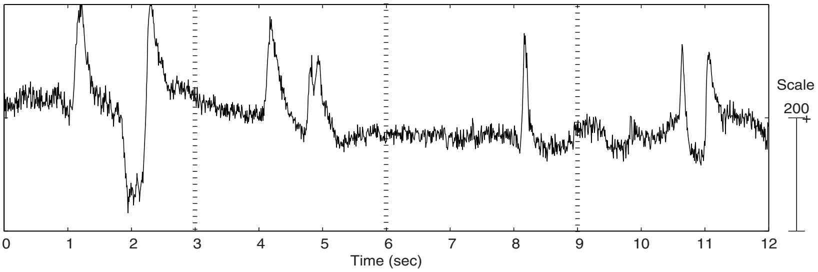


Fig. 1. Fp2 channel (with reference to Cz): all subsegments contain an EOG artifact

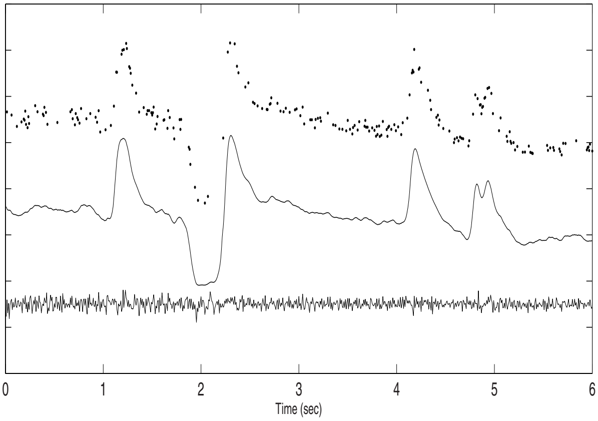


Fig. 2. Greedy KPCA algorithm in different steps: - *top* random selected training set, - *middle* extracted EOG signal and *bottom* - corrected EEG

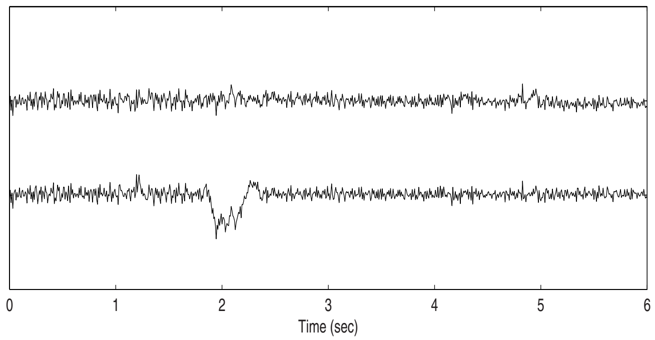


Fig. 3. Corrected EEG using as training set: *top* first subsegment of fig. 1, *bottom* the second segment of fig. 1

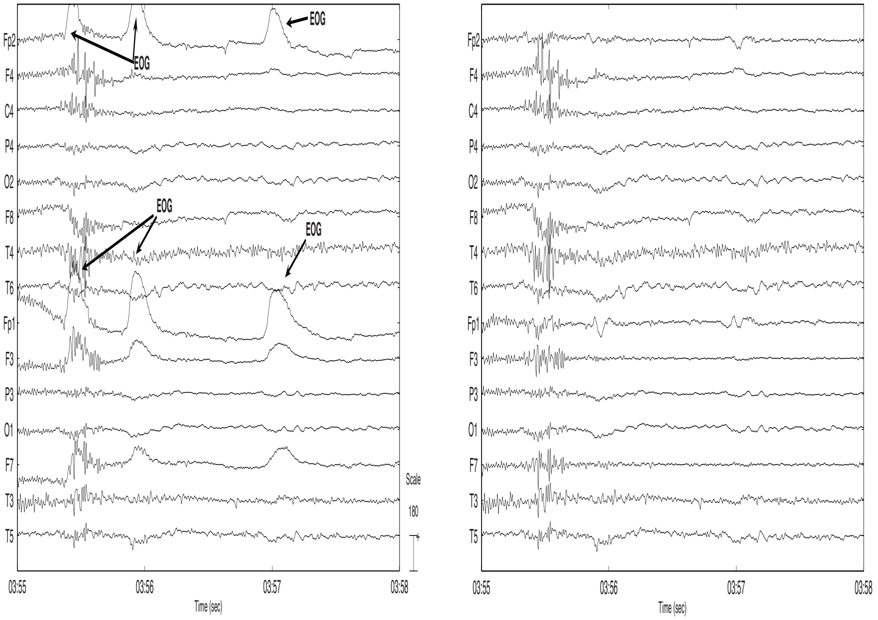
kernel is used with $\sigma = max_i(\|\mathbf{x}_i - \mathbf{x}_{mean}\|), i = 1, \ldots, J$, where $\mathbf{x}_{mean}$ is the mean of the data set.

Single Channel Analysis. A frontal (Fp2-Cz) EEG channel sampled at $128Hz$ was used. The segment of the signal containing high-amplitude EOG artifacts is shown in fig. 1). After embedding, the data set with $J = 1526$ vectors was split into testing and training sets. The training set is formed with roughly 25% of the data considering the following strategies

1. A random choice of $K = 381$ data vectors. Fig. 2 (plot on top) illustrates the random choice by plotting the first row of the training data vectors according to their time reference.
2. Using all vectors that correspond either to the first subsegment of the signal $(0-3s)$ or to the second subsegment $(3-6s)$.

The incomplete Cholesky decomposition algorithm is used with $R = 20$. For the three cases data were projected onto $L = 6$ basis vectors ($\mathbf{U}$) which correspond to the leveling off of the eigenspectrum of matrix $\mathbf{Q}$. Fig. 2 shows the result of the application of the algorithm when the training set is selected randomly. Fig. 3 shows the first $6s$ of the corrected EEG when the training set was formed by a subsegment of the signal. It is easily visible that when the training set do not have ”examples” of the features to extract the algorithm does not work well. The strongly negative features present in the first subsegment are not represented in the second subsegment. Then when the training set is formed with vectors belonging to the 2nd segment, the artifact is not removed (see fig. 3- plot on bottom).

Multichannel Analysis. The algorithm is applied in parallel to recordings from 4 different EEG electrodes all of which contain a high amplitude EOG interference. The segments each have a duration of $10s$ and fig. 4 (on the left) shows a subsegment of $3s$. After embedding, the training set was formed with roughly 25% of the multidimensional data set. The incomplete Cholesky decomposition with $R = 20$ was used. In



(a) Original EEG

(b) Corrected EEG

Fig. 4. EEG signals placed according $10 - 20$ system with reference to Cz. Only $3s$ of the $10s$ are plotted. Channels processed: Fp2, Fp1, F3 and F7.

the feature space, the mapped multidimensional signal is projected onto L directions chosen according to the eigenspectrum of matrix $\mathbf{Q}$ (see table 1).

Note that the number of directions of the Fp channels is higher because the artifacts have a higher energy in these recordings. The results of the processing are plotted in fig. 4 and firmly corroborate our previous conclusions. Note that the algorithm has only one parameter (L), so it is very easy to do an automatic tool to provide on-line help in any visual inspection and analysis of the recordings.

Table 1. Multichannel Analysis. Number of directions to project data

	$Fp1$	$Fp2$	$F3$	$F7$
L	6	6	4	3

4 Concluding Remarks

In this work we present a new variant of greedy KPCA algorithms where the data is split into training and testing sets. The data set is projected onto basis vectors computed with the training set only. This strategy is important to decrease the computational burden of kernel principal component analysis methods when large data sets are involved. In the proposed variant the complexity of the algorithm is related with the splitting of the training set into two subsets. This procedure is achieved by computing the trace of a matrix in each step which involves the manipulation of the whole training set. In our application, we conclude experimentally that with 20 steps we form a subset with $R = 20$ elements that allows to estimate the $L < R$ basis vectors needed to project the data. In what concerns the artifact elimination, the proposed method needs the information contained within a single channel only, hence can be applied to each channel separately. Thus only channels which contain such artifacts need to be processed. Our preliminary results show good performance in removing artifacts like eye or head movements. In summary, the method achieves the separation of EEG signal recordings into two components: artifacts and undistorted EEG. Although this is ongoing work, we present a method that is intended to ease a visual inspection of the EEG recordings by an experienced clinician, hence might be useful in some critical segment analysis like the onset of ictal seizures. It is also to be noticed that despite the variety of methods applied, it is not possible to conclude about their performance once they use distinct databases, different measures and goals. The proposed method needs to be evaluated in a more quantitative and systematic approach, concerning for instance the spectral distortion in the important frequency ranges of EEG. A plug-in to EEGLAB platform [3] has been developed to introduce the method in the clinical routine and facilitate the comparison with other methods.

Acknowledgments

A.R. Teixeira received a PhD Scholarship (SFRH/BD/28404/2006) supported by the Portuguese Foundation for Science and Technology (FCT). This work was also supported by grants from DAAD and CRUP which is gratefully acknowledged. We are grateful to A. Martins da Silva and Hospital Geral de Santo António for providing the signals and the helpful discussions.

References

1. Bach, F.R., Jordan, M.I.: Kernel independent component analysis. Journal of Machine Learning Research 3, 1–48 (2002)
2. Cawley, G.C., Talbot, N.L.C.: Efficient formation of a basis in a kernel induced feature space. In: Verleysen, M. (ed.) European Symposium on Artificial Neural Networks, Bruges, Belgium, pp. 1–6 (2002)
3. Delorme, A., Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. Journal of Neuroscience Methods 134, 9–21 (2004)
4. Fowlkes, C., Belongie, S., Chung, F., Malik, J.: Spectral grouping using the nyström method. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(2), 214–225 (2004)
5. Franc, V., Hlaváč, V.: Greedy algorithm for a training set reduction in the kernel methods. In: Petkov, N., Westenberg, M.A. (eds.) CAIP 2003. LNCS, vol. 2756, pp. 426–433. Springer, Heidelberg (2003)

6. Golyandina, N., Nekrutkin, V., Zhigljavsky, A.: Analysis of Time Series Structure: SSA and Related Techniques. Chapman & HALL/CRC (2001)
7. He, P., Wilson, G., Russel, C.: Removal of ocular artifacts from electroencephalogram by adpative filtering. Medical & Biological Engineering & Computing 42, 407–412 (2004)
8. Joyce, C.A., Gorodniysky, I.F., Kutas, M.: Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. Psychophysiology 41, 313–325 (2004)
9. Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M.J., Iragui, V., Sejnowski, T.j.: Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 163–178 (2000)
10. Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K., Schölkopf, B.: An introduction to kernel-based algorithms. IEEE Transactions on Neural Networks 12(2), 181–202 (2001)
11. Teixeira, A.R., Tomé, A.M., Lang, E.W., Stadlthanner, K.: Nonlinear projective techniques to extract artifacts in biomedical signals. In: EUSIPCO2006, Florence, Italy (2006)
12. Teixeira, A.R., Tomé, A.M., Lang, E.W., Gruber, P., da Silva, A.M.: Automatic removal of high-amplitude artifacts from single-channnel electroencephalograms. Computer Methods and Programs in Biomedicine 83(2), 125–138 (2006)
13. Urrestarazu, E., Iriarte, J., Alegre, M., Valencia, M., Viteri, C., Artieda, J.: Independent component analysis removing artifacts in ictal recordings. Epilepsia 45(9), 1071–1078 (2004)
14. Vigário, R.N.: Extraction of ocular artefacts from EEG using independent component analysis. Electroencephalography and Clinical Neurophysiology 103, 395–404 (1997)
15. Williams, C.K.I., Seeger, M.: Using the nyström method to speed up kernel machines. In: Advances in Neural Information Processing Systems, pp. 682–688. MIT Press, Cambridge (2000)
16. You, C.H., Koh, S.N., Rahardja, S.: Signal subspace speech enhancement for audible noise reduction. In: ICASPP 2005, Philadelphia, USA, vol. I, pp. 145–148. IEEE, Los Alamitos (2005)
17. Zhou, W., Gotman, J.: Removing eye-movement artifacts from the EEG during the intracarotid amobarbital procedure. Epilepsia 46(3), 409–414 (2005)