PRINCIPAL AND INDEPENDENT COMPONENT ANALYSIS IN IMAGE PROCESSING

S.Vaseghi [1], *H.Jetelova* [1,2]

[1] Brunel University, Department of Electronics and Computer Engineering
[2] Institute of Chemical Technology, Department of Computing and Control Engineering

Abstract

This paper is devoted to practical utilization of Principal Component Analysis (PCA) and its extension Independent Component Analysis (ICA). Our intension is to demonstrate different applications of the above mentioned methods in biomedical image and signal processing. The concept of ICA in terms of blind source separation is illustrated on EEG signals, whereas the approach of sparse coding is explained using fMRI images.

1 Introduction

Both Principal (PCA) and Independent Component analysis (ICA) are transformations that rely on statistics of the given data set. PCA is based on the information given by the second order statistics, whereas ICA goes up to high order statistics. Therefore the result obtained by ICA is assumed to be more meaningful than the one gained by PCA. However ICA better works on the data that have been already preprocessed by PCA. Thus ICA is often perceived as an extension of PCA. PCA and especially ICA have recently become popular tools in various fields, e.g. blind source separation, feature extraction, telecommunication, finance, text document analysis, seismic monitoring and many others. Two different approaches applied on EEG signals and images will be considered in this paper. All successive ICA experiments were designed in MATLAB environment using FastICA package proposed by Aapo Hyvärinen et al.

2 Principle of Independent Component Analysis

For better understanding of how ICA works consider the following problem. Imagine two or more people talking at the same time. Let us denote their speech as sources $\mathbf{s}$. Furthermore let us record the mixtures $\mathbf{x}$ of their speech by appropriate number of microphones. This may be mathematically formulated as

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

where the elements a_{ij} of the matrix $\mathbf{A}$ are unknown mixing factors [2, 4].

Now the task is to find the original sources (independent components) with no prior information about the mixing matrix $\mathbf{A}$. ICA is a transformation that enables us to find such sources. Herein it has to be stated that beside other limitations ICA does not reflect time delays that occur in the signal. Thus ICA is not suitable for speech. However due to favourable properties of biomedical signals it seems to be promising tool for data analysis in radiology and other related fields.

In order to further simplify the noise-free model of ICA assume that the number of sources equals the number of observed mixtures. Therefore in such case the gathered mixing matrix $\mathbf{A}$ is always squared. Consequently it can be shown that the matrix $\mathbf{A}$ is invertible and the inverse matrix - unmixing matrix $\mathbf{W}$ may be computed. Thus with the knowledge of matrix $\mathbf{A}$ respectively $\mathbf{W}$ the sources can be computed as follows

$$\mathbf{s} = \mathbf{W}\mathbf{x} \tag{2}$$

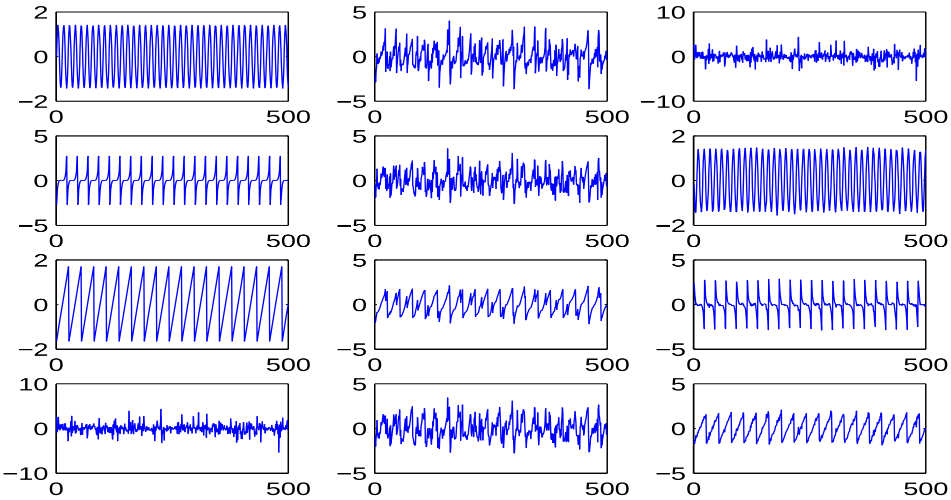

Figure 1: The original sources are depicted in the left column, their random linear mixtures in the middle, the obtained sources via ICA in the right column

Likewise other methods ICA embodies some inconveniences as well. Firstly the variances - powers of the sources can not be identified, since multiplying certain column of the matrix **A** by any scalar α can be always cancelled by dividing the appropriate column by the same scalar α. Secondly another problem arises because of the natural property of matrix computation - permutation. Thus it is impossible to determine the sequence of sources. Last but not least it should be remarked that ICA operates only with sources that posses non-gaussian distribution. Optionally only one source having gaussian distribution is allowed [2]. The illustration of ICA model on artificial signals is depicted Figure 1.

3 Whitening

When estimating the ICA model the transformation called whitening is often employed. For better illustration suppose a random vector **z** that posses following property

$$E\{z_i z_j\} = \delta_{ij} \tag{3}$$

In other words the Eq. (3) limits the elements of vector **z** to be uncorrelated and of unit variance. It may be generalized that any vector **z** that posses the above mentioned properties is called white (sphered) [2]. Now let us find a linear transformation **V** that projects random vector **x** into whitened vector **z**

$$\mathbf{z} = \mathbf{V}\mathbf{x} \tag{4}$$

Such transformation may be estimated via PCA due to the uncorrelatedness property of vector **z**. Thus PCA is often utilized in ICA as a preprocessing step. The outline is as follows. First the input data has to be centred. This is achieved by subtracting the sample mean value from given data set. Next the correlation/covarinace matrix $\mathbf{C}_{\mathrm{XX}}$ is evaluated and its eigen decomposition $\mathbf{C}_{\mathrm{XX}} = \mathbf{E}\mathbf{D}\mathbf{E}^T$ is performed. Consequently the whitening matrix **V** is designed as $\mathbf{V} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T$ by simply extracting the square root of each element of **D**. Now the set of **x** can be whitened as

$$\mathbf{z} = \mathbf{V}\mathbf{x} = \mathbf{V}\tilde{\mathbf{A}}\mathbf{s} \tag{5}$$

However whitening (sphering) still does not give us independent components. When transforming the gathered data **z** by any orthogonal matrix **M** the obtained vector $\mathbf{y} = \mathbf{M}\mathbf{z}$ is white as well. Thus whitening is not sufficient operation for estimating the independent components. In

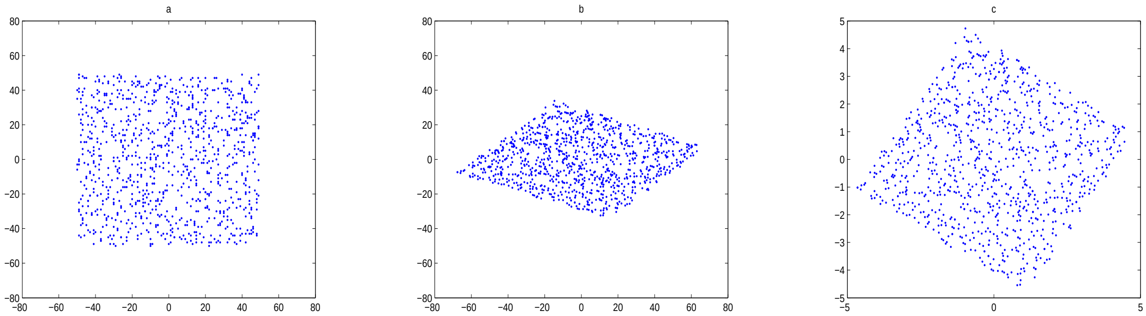

Figure 2: Whitening presenting **(a)** the joint distribution of the independent components s_1 and s_2 with uniform distributions, **(b)** the joint distribution of the observed mixtures x_1 and x_2, and **(c)** the joint distribution of the whitened mixtures z_1 and z_2

other words it can not be said if the independent components are addressed by $\mathbf{z}$ or $\mathbf{y}$. Nevertheless whitening is still very important step when preprocessing the data. It limits the space of mixing matrices to the orthogonal ones. The effect of whitening is illustrated in Figure 2.

The last step when performing ICA is choice of the matrix that is already assumed to be orthogonal. This choice is done by searching for a matrix $\mathbf{B}$ that maximizes some given measure of non-gaussianity of the sources $\mathbf{s} \approx \mathbf{Bz}$.

4 EEG Signal Analysis

It is well known that the activity of active nerve cells in the brain produces current that may be detected on the scalp using electrodes. The obtained signals are then traced in order to obtain electroencephalography (EEG) recording. However EEG recording reveals number of inconveniences. Briefly, the line noise at 50Hz plus another contaminations (artifacts) caused by factors such as eye movements, blinks, cardiac signals and muscle noise [3]. In this section the effective de-noising and possible artifacts rejecting will be shown.

4.1 De-noising

In our attempt EEG signals obtained from 19 channels were sampled at rate of 200Hz. Furthermore stop band filtering (because of the electricity supply at 50Hz) was performed in order to denoise the signals. For such purposes MATLAB function `fir1` $[y(m) = \sum_{k=0}^{M} b_k x^{(m-k)}]$ was implemented. The original and filtered EEG signals are depicted in Figure 3. Consequently the stop band filtered EEG signals were used as an input for ICA.

4.2 Independent Component Analysis

Two basic assumption about EEG signals have to be taken into account in order to proceed ICA. Firstly it is supposed that the data gathered from sensors (electrodes) are linear sums of signals evoked by spatially distributed sources in the brain. Secondly due to the nature of EEG signals the time delays are omitted. When considering the general model of ICA each EEG signal may be perceived as one observation $\mathbf{x}_i$. Furthermore the matrix $\mathbf{x}$ of such observations may be designed by simply taking each observation $\mathbf{x}_i$ as one row. The sources are then evaluated using the Eq. (2) and their biological feasibility is analyzed. Consequently sources that reveal the artifacts are substituted by zero vectors and the inverse ICA is proceeded in order to obtain corrected EEG recording.

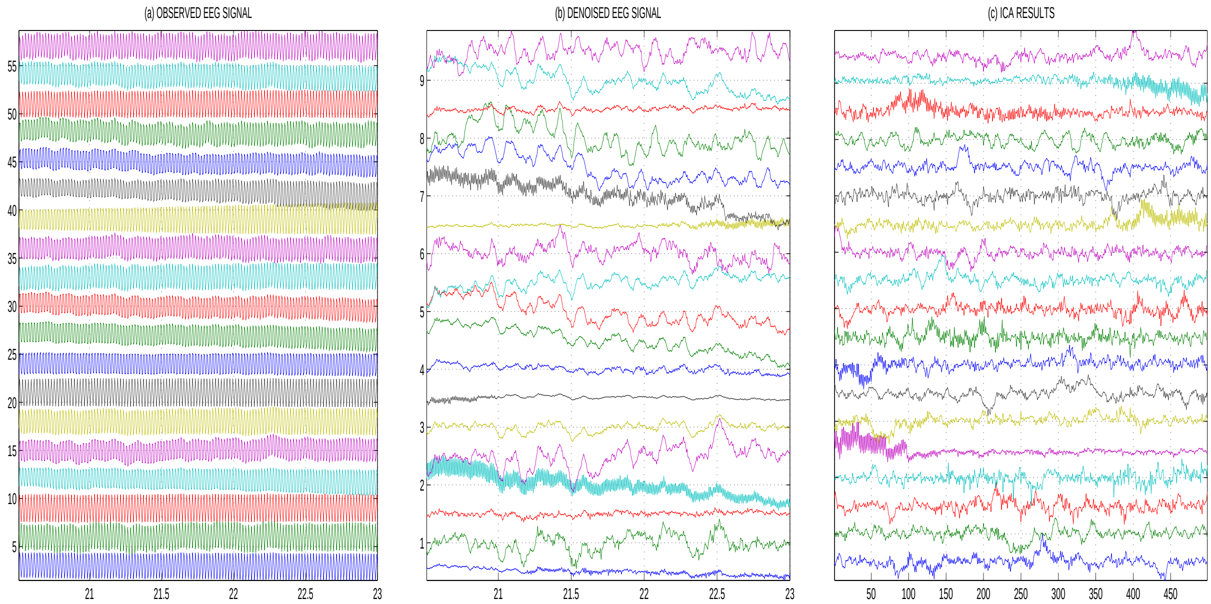

Figure 3: EEG Signals processing presenting **(a)** original signals sampled at rate of 200Hz, **(b)** signals obtained after the stop band filtering, and **(c)** signals (sources) obtained by ICA

5 ICA in Image Processing

So far we have concerned ourselves by application of ICA in terms of blind source separation. In this section utilization in image processing will be explained [1, 5]. The main concept of ICA applied to images insists on the idea that each image (subimage) may be perceived as linear superposition of features $a_i(x, y)$ weighted by coefficients s_i. In case of ICA, features are represented by columns of mixing matrix A and s_i are elements of appropriate sources. In addition ICA features are localized and oriented and sensitive to lines and edges of varying thickness of images (see Figures 4 and 5). Furthermore the sparsity of ICA coefficients should be pointed out. It is expected that suitable soft-thresholding on the ICA coefficients leads to efficient reducing of Gaussian noise [1].

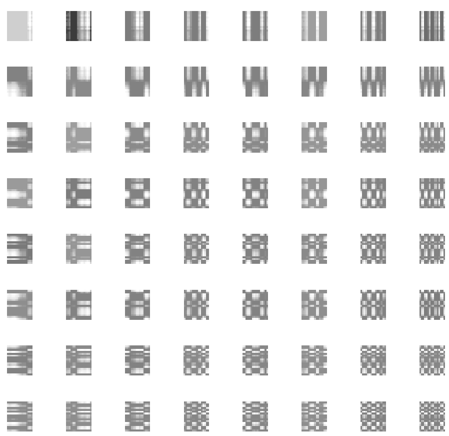

Figure 4: The set of eigen images extracted from fMRI scan of a brain tissue

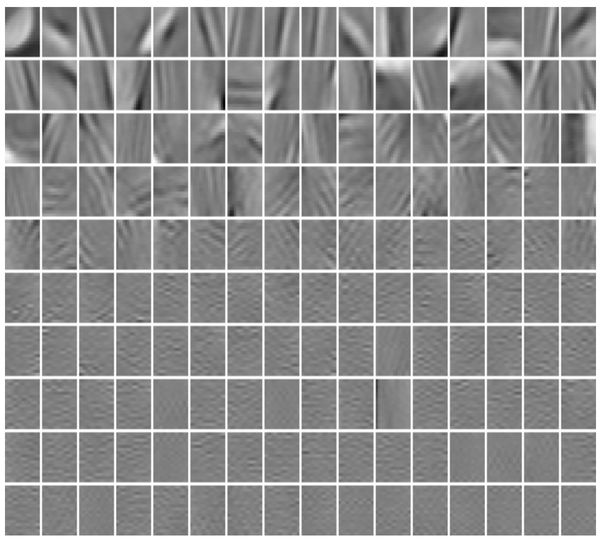

Figure 5: The basis vectors of fMRI scans gained by Fast ICA algorithm

6 Conclusion

In preceding sections multiple applications of ICA/PCA were outlined. It was also shown how much of the hidden information could be detected using ICA in consequence of PCA. Our future interest is to focus on the transformed data provided by ICA/PCA in order to avoid redundant information. Briefly, effective rejecting of artifacts from EEG signals and further de-noising of biomedical images.

7 Acknowledgments

The work has been supported by the research grant of the Faculty of Chemical Engineering of the Institute of Chemical Technology, Prague No. MSM 6046137306. Data were kindly provided by MUDr. Oldřich Vyšata from the Neurocenter Caregroup in Rychnov n. Kněžnou. Authors express their thanks also to Prof. Aleš Procházka for his valuable advice and suggestions.

References

[1] A. Hyvarinen, P. O. Hoyer, and E. Oja. Sparse Code Shrinkage: Denoising by Nonlinear Maximum Likelihood Estimation . *Advances in Neural Information Processing Systems*, (11):473–479, 1999.

[2] Aapo Hyvarinen, Juha Karhunen, and Erkki Oja. *Independent Component Analysis.* J. Wiley, 2001.

[3] Tzyy-Ping Jung, Colin Humphries, Te-Won Lee, Scott Makeig, Martin J. McKeown, Vicente Iragui, and Terrence J. Sejnowski. Removing Electroencephalographic Artifacts: Comparison between ICA and PCA. *Neural Networks for Signal Processing*, VIII:63–72, 1998.

[4] James V. Stone. *Independent Component Analysis.* The MIT Press, 2004.

[5] Daoqiang Zhang, Songcan Chen, and Jun Liu. Representing Image Matrices: Eigenimages Versus Eigenvectors. *Lecture Notes in Computer Science*, 3497/2005:659–664, 2005.

Prof. Saeed Vaseghi
Brunel University, London
Department of of Electronics and Computer Engineering
West London, UB8 3PH, UK
Phone: +44(0)1895 − 203186
E-mail: Saeed.Vaseghi@brunel.ac.uk

Helena Jetelová
Institute of Chemical Technology, Prague
Department of Computing and Control Engineering
Technická 1905, 166 28 Prague 6
E-mail: jetelova@email.cz