Independent Component Analysis Applied in Biomedical Signal Processing

M. Ungureanu, C. Bigan, R. Strungaru, V. Lazarescu

Politehnica University of Bucharest, Iuliu Maniu 1-3, Bucharest, Romania,
Email:mihaela@nspg.pub.ro

Abstract. *The EEG is composed of electrical potentials arising from several sources. Each source (including separate neural clusters, blink artifact, or pulse artifact) projects a unique topography onto the scalp-'scalp maps'. These maps are mixed according to the principle of linear superposition. Independent component analysis (ICA) attempts to reverse the superposition by separating the EEG into mutually independent scalp maps, or components so that the method used is the study is independent component analysis (ICA). The Polysomnogram (PSG), a complex of more signals recorded over night for the patients with respiratory problems contains also redundant information that can be removed using the ICA.*

Keywords: Independent Component Analysis, Electroencephalogram, Polysomnogram

1. Introduction

The paper is studying the applicability of ICA to EEG and PSG.
The EEG signal contains among the useful information, which allow scientists to view the cerebral activity, redundant or noise information. In order to conclude that something is wrong or that the patients have a disease further processing is necessary.

EEG obtained from scalp electrodes is a sum of the large number of neurons potentials. The interest is in studying the potentials in the sources inside the brain and not only the potentials on the scalp, which globally describe the brain activity. Direct measurements from the different centers in the brain require placing electrodes inside the head, which means surgery. This is not acceptable because of the risk for the subject. Another possibility is to calculate the signals of interest from the EEG obtained on the scalp. These signals are weighed sums of the neurons activity, the weights depending on the signal path from the brain cell to the electrodes. Because the same potential is recorded from more than one electrode, the signals from the electrodes are supposed to be highly correlated. If the weights were known, the potentials in the sources could be computed from a sufficient number of electrode signals. Independent component analysis (ICA), sometimes referred to as blind signal separation or blind source separation, is a mathematical tool that can help solving the problem.

This is an extension to principal components analysis (PCA), which has had a place in EEG research for years [1, 2]. While little is known about the actual distributions of brain activity, ICA practitioners claim that the sources comprising EEG are not of a Gaussian nature [3].

The polysomnogram has been recently very used in diagnosing the respiratory diseases, but the problem is that the polysomnographic analysis is very expansive. In order to reduce the cost the researcher of important sleeping laboratory have tried to use only some signal to get the information about the respiratory problems [4].

The complete analysis, over a long period of time and including a lot of signals (respiratory signal, ECG, EEG, EOG and EMG) is defined as polysomnogram (PSG).

Obstructive Sleep Apnea (OSA) is a disease in which airways involuntarily collapse during sleep, leading to serious consequences. An OSA attack is characterized by repeated episodes of upper airway closure during sleep and is defined as the total cessation of respiratory airflow that lasts at least for 10s. The researchers have proved that the patients affected by OSA are generally exposed to the hypertension, ischemic heart diseases and stroke. The consequences can be extended on the industrial accidents, driving problems and decreasing of the work quality due to daytime sleepiness. That is why usually in the clinics the patients are under the observation all over the night in order to analyze and estimate the respiratory problems they have.

2. Measurement technique

For the first study, on the applicability of ICA on electroencephalogram, the EEG consists of 22 signals, recorded by unipolar electrodes at a sampling rate of 256Hz.

For the second part of the study the polysomnographic data used in the study were collected at the Asklepios Clinic, Gauting, Munich, Germany. The sampling rates of the channels are presented in Tab. 1.

Table 1

Signal	Measurement Unit	Sampling Frequency
EEG --- C4/A1	µV	125 Hz
EEG --- C3/A2	µV	125 Hz
EOG RIGHT F4/A1	µV	125 Hz
BODYPOS HANDY F3/A2	"	125 Hz
EMG CHIN P4/P3	µV	250 Hz
ECG --- Ecg	µV	125 Hz
RESP FLOW Fp2/Cz	"	25 Hz
RESP FLOW Fp1/Cz	"	25 Hz
RESP FLOW Fz/Cz	"	25 Hz
RESP THORAX Thrx	µV	25 Hz
RESP ABDOMEN Abdm	µV	25 Hz
SOUND RAW Snd	µV	250 Hz
SAO2 SCHWARZ Oxi	%	25 Hz
EMG TIBIALR T4/T6	µV	250 Hz
EMG TIBIALL T3/T5	µV	250 Hz
BODYPOS SCHWARZ Pos	"	25 Hz
CPAP SCHWARZ Prss	mBar	25 Hz

3. Independent Component Analysis Procedure

We assume that we observe *n* linear mixtures $x_1,..., x_n$ of n independent components:

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ... + a_{jn}s_n,\ j = \overline{1,n} \tag{1}$$

In this equation the time has been ignored. Instead, it was assumed that each mixture x_j as well as each independent component s_i are random variables and $x_j(t)$ and $s_i(t)$ are samples of these random variables. It is also assumed that both the mixture variables and the independent components have zero mean [5].

If not subtracting the sample mean can always center the observable variables x_i. This procedure reduces the problem to the model zero-mean:

$$\hat{\mathbf{x}} = \mathbf{x} - \mathrm{E}(\mathbf{x}) \tag{2}$$

Let x be the random vectors whose elements are the mixtures $x_1,..., x_n$ and let s be the random vector with the components $s_1,\ldots, s_n$. Let **A** be the matrix containing the elements a_{ij}. The model can now be written:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \qquad \text{or} \qquad \mathbf{x} = \sum_{i=1}^{n} \mathbf{a}_i s_i \tag{3}$$

The above equation is called independent component analysis or ICA. The problem is to determine both the matrix **A** and the independent components s, knowing only the measured variables x. The only assumption the methods take is that the components s_i are independent. It has also been proved that the components must have nongaussian distribution [6, 7]. ICA looks a lot like the "blind source separation" (BSS) problem or blind signal separation: a source is in the ICA problem an original signal, so an independent component. In ICA case it is also no information about the independent components, like in BSS problem.

Whitening can be performed via eigenvalue decomposition of the covariance matrix:

$$\mathbf{V}\mathbf{D}\mathbf{V}^{\mathrm{T}} = \mathrm{E}\left[\hat{\mathbf{x}}\hat{\mathbf{x}}^{\mathrm{T}}\right] \tag{4}$$

where **V** is the matrix of orthogonal eigenvectors and **D** is a diagonal matrix with the corresponding eigenvalues. The whitening is done by multiplication with the transformation matrix **P**:

$$\begin{aligned} \tilde{\mathbf{x}} &= \mathbf{P}\hat{\mathbf{x}} \\ \mathbf{P} &= \mathbf{V}\mathbf{D}^{-\frac{1}{2}}\mathbf{V}^{\mathrm{T}} \end{aligned} \tag{5}$$

The matrix for extracting the independent components from $\tilde{\mathbf{x}}$ is $\tilde{\mathbf{W}}$, where $\mathbf{W} = \tilde{\mathbf{W}}\mathbf{P}$.

4. Fast ICA for *n* units

A unit represents a processing element, for example an artificial neuron with its weights **W**. To estimate several independent components, the weights $\mathbf{w}_1,...,\mathbf{w}_n$ must be determined. The problem is that the outputs $\mathbf{w}_1^T\mathbf{x},...,\mathbf{w}_n^T\mathbf{x}$ must be done as independent as possible after each iteration in order to avoid the convergence to the same maxima. One method is to estimate the independent components one by one [1, 8].

Algorithms:
i) Initialize w_i
ii) Newton phase

$$\mathbf{w}_i = \mathrm{E}\left\{\tilde{\mathrm{x}}\mathrm{g}\left(\mathbf{w}_\mathrm{i}^\mathrm{T}\tilde{\mathbf{x}}\right)\right\} - \mathrm{E}\left\{g'\left(\mathbf{w}_i^\mathrm{T}\tilde{\mathbf{x}}\right)\right\}\mathbf{w}_i \tag{6}$$

where *g* is a function with one of the following form:

$$g_1(y) = \tanh(a_1 y),\ g_2(y) = y\exp\left(-\frac{1}{2}y^2\right),\ g_3(y) = 4y^3 \tag{7}$$

iii) Normalization

$$\mathbf{w}_i = \frac{1}{\|\mathbf{w}_i\|}\mathbf{w}_i \tag{8}$$

iv) Decorrelation

$$\mathbf{w}_i = \mathbf{w}_i - \sum_{j=1}^{i-1}\mathbf{w}_i^T\mathbf{w}_j\mathbf{w}_j \tag{9}$$

v) Normalization (like in the step iii))
vi) Go to step ii) if not converged.

5. Results

All the computations have been made in Matlab.

5.1 ICA applied for EEG

In the first study 120 samples of EEG were recorded at a sampling frequency of 256Hz. At the beginning, 22 components have been acquired. Analyzing the result, 7 independent components were identified, the remaining components being correlated with the other ones. Further studies will try to remove the artifacts based on these independent components and the recorded samples.

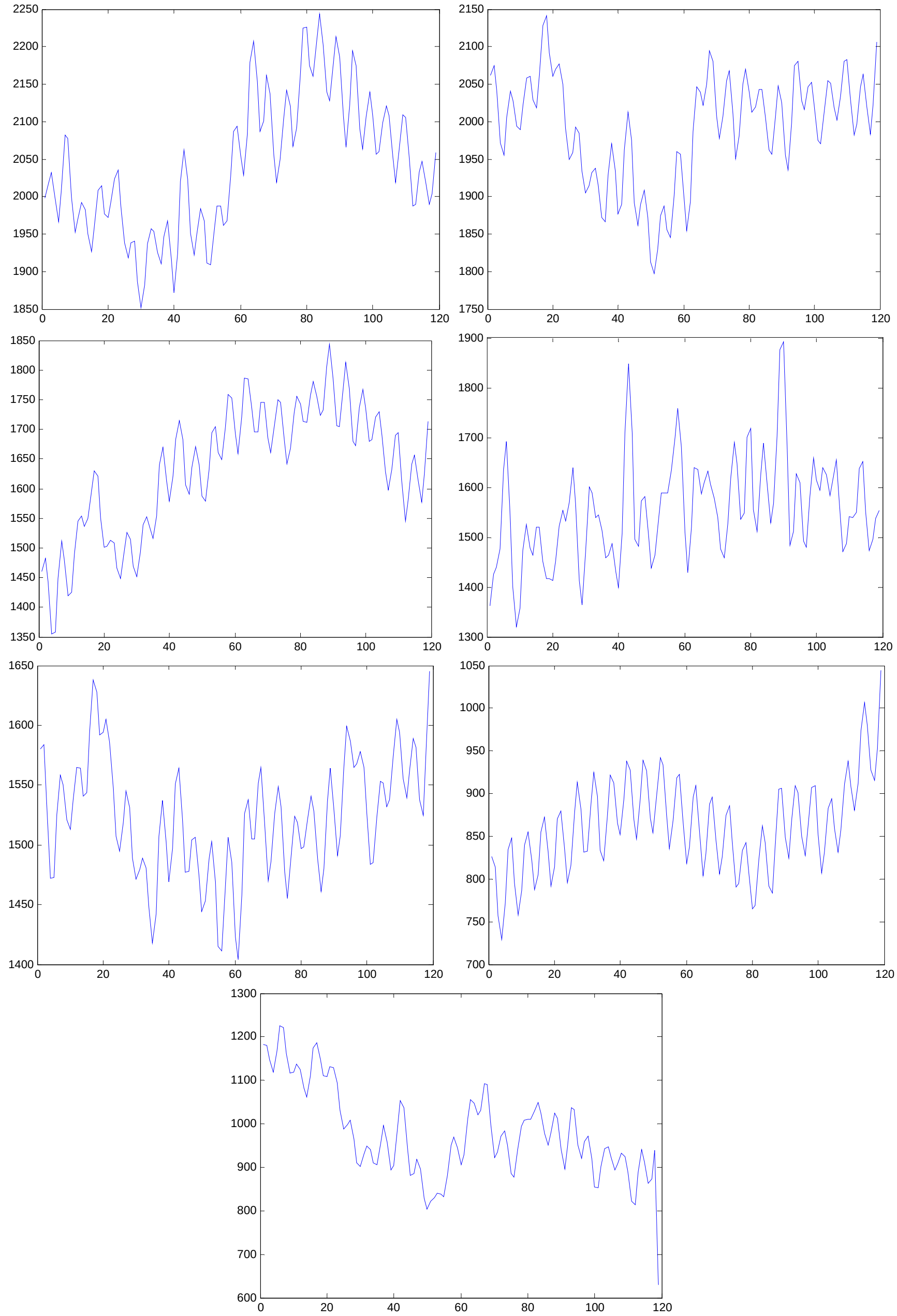

Fig. 1. The 7 identified independent components from the EEG (initial 22 signals were recorded)

In order to compare more easily the extracted components, an averaging filter of 30th order is applied. The study proves that independent component analysis can be usefully applied for EEG signals processing in order to get the most important information contained by this signals. That is important both in diagnosis and research because the amount of data to be processed is reduced.

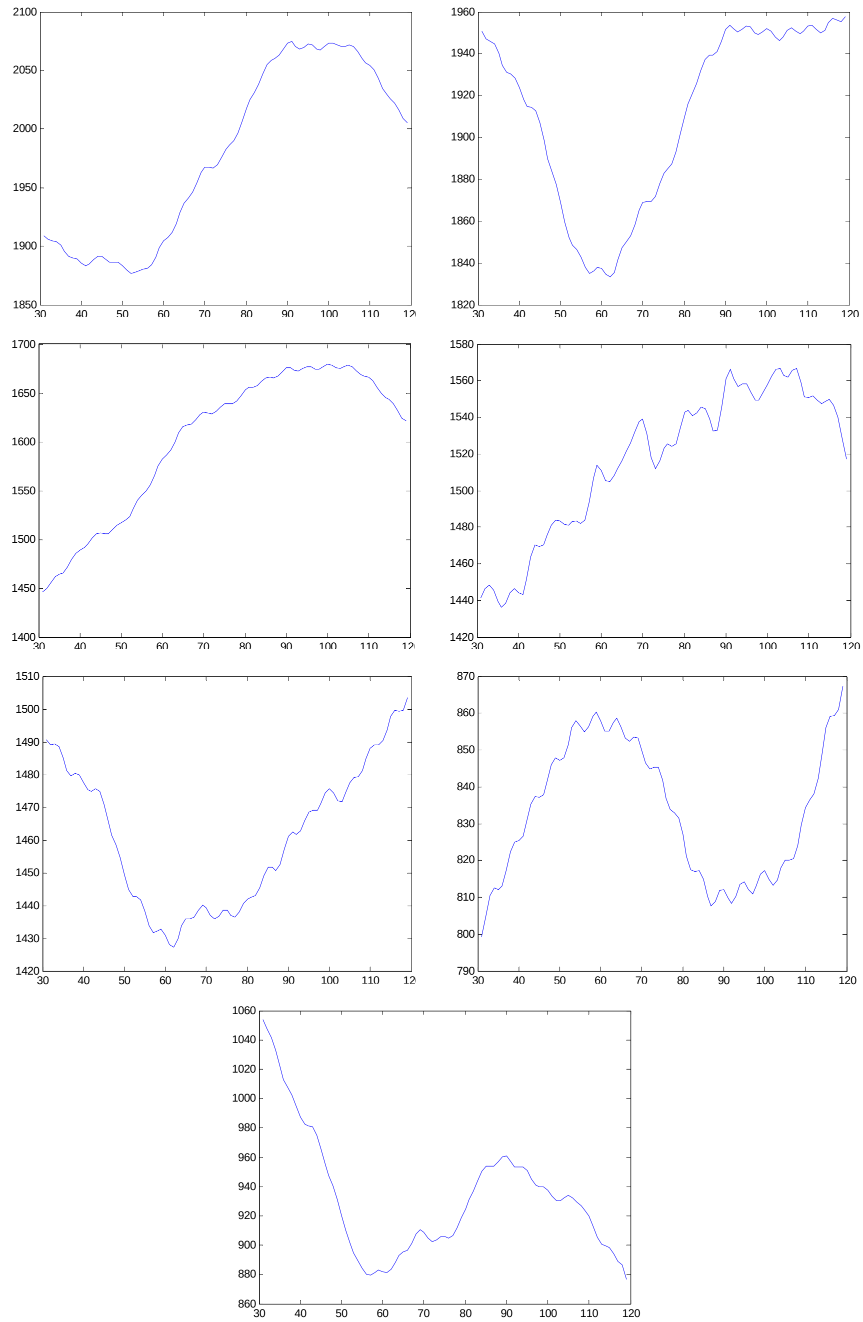

Fig. 2. The averaged independent components

5.2 ICA applied for EEG

In the second study segments of 4s was considered in the analysis. All the signals contained in the PSG were interpolated, when necessary, in order to have a common sampling frequency of 250 Hz.

First the ICA was applied for the EEG signals contained in the PSG and then for the EOG signals. As it can be seen in Fig. 3, one signal is dependent on the others. When considering the respiratory signal, represented by the signal collected from the nose, mouth, thorax and abdomen, the results of applying the ICA is more encouraging. In this case only three signals are identified as independent, Fig. 4.

Fig. 3. The 3 (right side) identified independent components in the signals corresponding to the brain activity (left side), when PSG is recorded

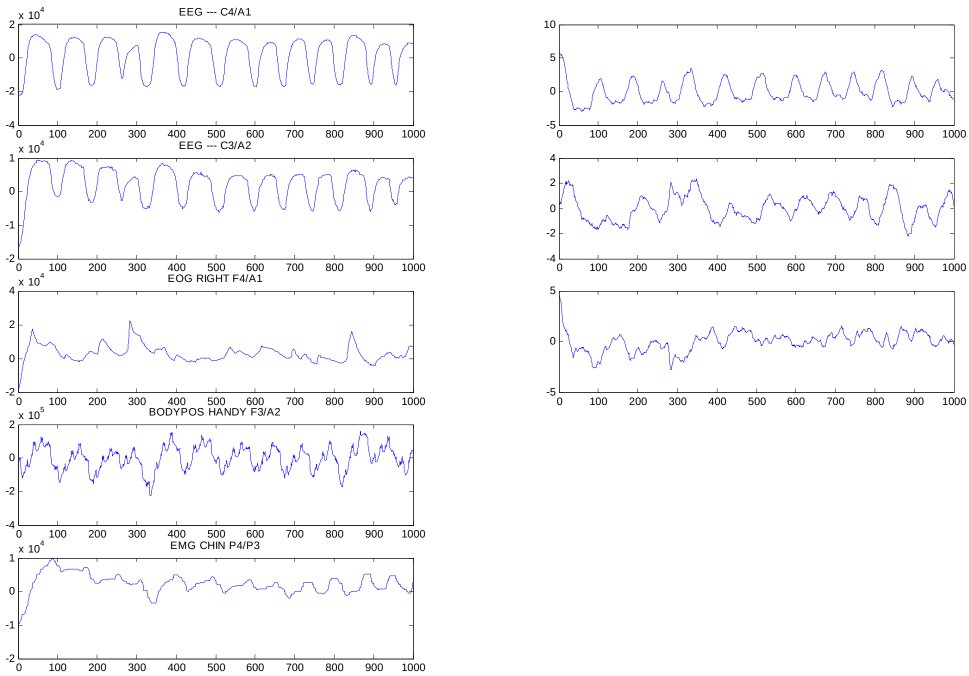

Fig. 4. The 2 (right side) identified independent components in the signals corresponding to the brain activity (left side), when PSG is recorded

6. Discussions and conclusions

The study proves that ICA is a powerful tool when the biomedical analysis involved more channels, which is the case of electroencephalogram and polysomnogram. In this case the important information can be obtained considering only the relevant signals, obtained after applying the Independent Component Analysis.

Acknowledgement

The authors would like to thank Werner Wolf and Maximilian Bauer, from for the German Armed Forces University, Munich, Germany and ASKLEPIOS Medical Center and Clinic for Lung diseases, Gauting, Munich, Germany respectively, for the polysomnographic data.

References

[1] L. Xu L., C.C. Cheung, H.H. Yang and S. Amari, Independent Component Analysis by the Information-theoretic Approach with Mixture of Densities", IEEE Trans. Neural Networks, Vol. BME-5 (1997), pp. 1821-1826
[2] E. Moulines, J.-F. Cardoso, and E. Gassiat, Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models, Proc. ICASSP'97 IEEE Int. Conf. Acoustics, Speech and Signal Processing, Munich, Germany, 1997, pp. 3617-3620
[3] Makeig S, Enghoff S, Jung T-P, and Sejnowski TJ, Independent Components of Event-related Electroencephalographic Data, Cognitive Neuroscience Society Abstracts, 2000, pp. 93.
[4] S.D. Pittman, N. Tal, G. Pillar, A. Malhotra, M.F. Hilton, R.B. Fogel, N. Ayas and D.P. White, Automated detection of obstructive sleep-disordered breathing events using peripheral arterial tonometry and oximetry, Computers in Cardiology, Vol. 27 (2000), pp. 485-488.
[5] J. Hyvarinen, Fast and Robust Fixed-Point Algorithms for Independent Component Analysis, IEEE Trans. Neural Networks, Vol. NN-10(3) (1999), pp. 626-634.
[6] E. Oja., Nonlinear PCA criterion and maximum likelihood in independent component analysis, Proc. ICA'99 Int. Workshop Independent Component Analysis and Signal Separation , Aussois, France, 1999, pp. 143-148.
[7] P. Pajunen and J. Karhunen, A maximum likelihood approach to nonlinear blind source separation, Proc. ICANN'97 Int. Conf. on Artificial Neural Networks, Lausanne, Switzerland, 1997, pp. 541-546.
[8] R. Vigário, V. Jousmäki, M. Hämäläinen, R. Hari, and E. Oja., Independent component analysis for identification of artifacts in magnetoencephalographic recordings, Proc. NIPS'97, Cambridge, MA, MIT Press, 1998, pp. 229-235.