THE ICA APPROACH FOR REMOVAL OF UNDESIRED COMPONENTS FROM EEG DATA

Joanna Górecka
Institute of Control Engineering, Technical University of Szczecin, 26 kwietnia, Szczecin, Poland

Keywords: Independent Component Analysis, Blind Signal Separation, EEG data, Artifacts.

Abstract: The aim of this results of research is to detect and remove selected undesired signals by means of ICA approach. In this paper have been presented the following algorithms BSS: HJ, Infomax and FastICA for separation and removal of selected group of artifacts (eye blinks, muscle activity) from EEG recordings. As it has been proven in experiments, the proposed algorithms can effectively detect and remove these artifacts from EEG recordings.

1 INTRODUCTION

The Electroencephalogram is a biological signal that represents the electrical activity of the brain. EEG signals recorded at the scalp are mixtures of the signals from multiple intra - and extracranial sources. The assumption of independence of the sources was justified through the successful application of the ICA technique to the identification and extraction of selected artifacts in EEG recordings, as presented in (Cichocki, Amari, 2002). In many cases a linear model is usually inappropriate for EEG signals (Girolami, 2000).

An important application of ICA is in Blind Signal Separation. The block diagram illustrating blind separation problem is presented in Figure 1.

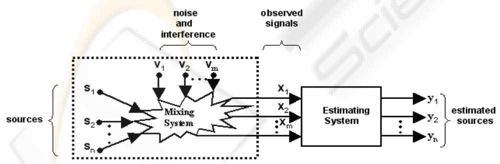

Figure 1: General scheme of the blind separation process (Cichocki, Amari, 2002).

Usually, in blind signal processing m mixed signals $x_i(t)$ for $i = 1,2,...,m$ are linear combinations of n mutually unknown and statistically independent and zero-mean source signals $s_j(t)$ for $j = 1,2,...,n$ and are noise - contaminated (Cichocki, Amari, 2002). It can be written in the matrix notation:

$$\mathrm{x}(t) = \mathbf{H}\mathrm{s}(t) + \mathrm{v}(t) \tag{1}$$

where: $\mathrm{x}(t) = [\mathrm{x}_1(t), \mathrm{x}_2(t), ..., \mathrm{x}_m(t)]^{\mathrm{T}}$ - is a vector of observed signals, $\mathbf{H} \in \mathbf{R}^{m \times n}$ - is an unknown mixing matrix, $\mathrm{s}(t) = [\mathrm{s}_1(t), \mathrm{s}_2(t), ..., \mathrm{s}_n(t)]^{\mathrm{T}}$ - is a vector of signal sources and $\mathrm{v}(t) = [\mathrm{v}_1(t), \mathrm{v}_2(t), ..., \mathrm{v}_m(t)]^{\mathrm{T}}$ - is a vector of additive noise. On the other hand, the demixing model is a linear transformation in the following form:

$$\mathrm{y}(t) = \mathbf{W}\mathrm{x}(t) \tag{2}$$

where: $\mathrm{y}(t) = [\mathrm{y}_1(t), \mathrm{y}_2(t), ..., \mathrm{y}_n(t)]^{\mathrm{T}}$ - is an estimate of source signals $\mathrm{s}(t)$ and $\mathbf{W} \in \mathbf{R}^{nxm}$ - is a separating matrix to be determined.

The aim of BSS using ICA is to estimate an unmixing (separating) matrix such that $\mathbf{Y} = \mathbf{WX}$ approximates the independent source signals as good as possible (Roberts, Everson, 2001). In this paper, the unmixing matrix for the instantaneous case is equal to the inverse of the mixing matrix, i.e. $\mathbf{W} = \mathbf{H}^{-1}$.

2 ADAPTIVE ALGORITHMS FOR NON-STATIONARY SIGNALS

The choice of adaptive algorithm depends on the statistical properties of sources and the specific applications. For separation of independent and non-gaussian signals, for example: EEG recordings, the

Górecka J. (2009).
THE ICA APPROACH FOR REMOVAL OF UNDESIRED COMPONENTS FROM EEG DATA.
In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*, pages 539-542
DOI: 10.5220/0001777605390542

best performance can by achieved by using the higher-order statistics (HOS) approach. In many applications, separation algorithms are combinations of two approaches: HOS and SOS. The second-order statistics (SOS) are useful for blind signal separation, when the source signals are statistically non-stationary.

The fundamental restriction in ICA methods is that independent components must be non-gaussian for ICA to be as much as possible. The classical measure of nongaussianity is kurtosis or the fourth-order cumulant (Cichocki, Amari, 2002). A second measure is given by negentropy, which is based on the information - theoretic quantity of (differential) entropy. Next approach for ICA separation is based on information theory - minimization of mutual information (Van Hulle, 2008).

Usually, algorithms for Independent Component Analysis can be divided into two categories (Cichocki, Amari, 2002). In the first category algorithms rely on batch computations minimizing or maximizing some relevant criterion functions, for example: FOBI (Fourth Order Blind Identification), FOBI-E (Fourth Order Blind Identification with Transformation matrix E), JADE (Joint Approximate Diagonalization of Eigen matrices), JADE TD (Joint Approximate Diagonalization of Eigen matrices with Time Delays), FPICA (Fixed-Point ICA). It was a problem with these algorithms, because they require very complex matrix or tensorial operations.

In the second category adaptive algorithms often based on stochastic gradient methods, for example: NG-FICA (Natural Gradient - Flexible ICA), ERICA (Equivariant Robust ICA - based on Cumulants), SANG (Self Adaptive Natural Gradient algorithm with nonholonomic constraints). The main problem of these algorithms is the slow convergence and dependence on the correct choice of the learning rate parameters (in neutral networks). It has been proven (Cichocki, Amari, 2002) that the Natural Gradient algorithms improves greatly the learning efficiency in blind separation process.

Generally, the adaptive learning algorithms can by written in the general form by using estimating functions (Cichocki, Amari, 2002):

$$\mathbf{W}(t+1) = \mathbf{W}(t) + \mathbf{\Delta W} \tag{3}$$

where: $\mathbf{W}(t)$ - is a separating matrix; $\mathbf{\Delta W} = \mu(t)(\mathbf{I} - \mathbf{R}_{fy})\mathbf{W}(t)$ for that: $\mu(t)$ - is a learning rate at time, $\mathbf{I}$ - is an identity matrix and $\mathbf{R}_{fy}$ - is a covariance matrix.

Many methods have been proposed to remove eye blinks and muscle activity from EEG recordings (Rangayyan, 2002; Sanei, Chambers, 2007). Applications of ICA approach to EEG data have concentrated on source localization and on artifacts removal. Usually, the EEG recordings can be first decomposed into useful signal and undesired subspace of components using standard techniques like local and robust PCA, SVD or nonlinear adaptive filtering (Rangayyan, 2002). In the following step, the ICA algorithms decomposed the observed signals (signal subspace) into independent components. It is worth to noting, that some useful sources are not necessarily statistically independent. Therefore, the perfect separation of primary sources by using any ICA procedure cannot be achieved (Roberts, Everson, 2001). However, in this experiment the separation of the EEGs is not important, but only the removal of independent undesired components.

3 METHODS AND MATERIALS

The performance of three chosen adaptive algorithms presented in this paper have been implemented in MATLAB software.

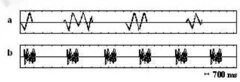

Figure 2: Artifacts: a) eye blinks (1÷2,5) Hz; b) muscle (20÷60) Hz.

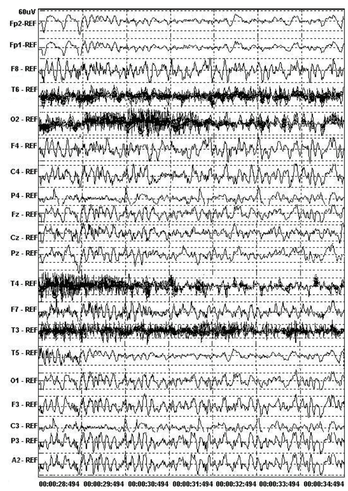

Figure 3: An example of EEG data with eye blinks and muscle artifacts.

The EEG signals have been prepared using BIOSIG (http://biosig.sourceforge.net/index.html), EEGLAB (http://sccn.ucsd.edu/eeglab/) and ICALAB software (Cichocki et al., 2007).

For verification of the quality of separation and removal of selected artifacts from EEGs, well-known source signals have been artificially mixed using well-known full rank mixing matrix (BSS problem). Furthermore it has been assumed, that the number of sources is equal to the number of sensors. In the following step, two types of artifacts have been added to appropriate channels: (T3, T4, T6, O2) - muscle artifacts; (F3, F4) - eye blinks (Majkowski, 1986). All signals were mixed using the mixing matrix H ($\det(H) = 68{,}7546$).

Finally, three adaptive learning algorithms have been chosen: HJ (Herault, Jutten, 1991), Infomax (Bell, Sejnowski, 1995) and FastICA (Hyvärinen, 1999).

In the HJ algorithm non-gaussian sources with similar number of independent sources and mixtures have been considered. A solution based on a recurrent artificial neural network for separation of these sources has been proposed. It can be written as:

$$\frac{d\mathbf{W}}{dt} = \eta(t)\mathbf{g}(\mathbf{y}(t))\mathbf{h}(\mathbf{y}(t)) \tag{4}$$

where: $\eta(t)$ - is a learning rate, $g(\cdot)$ and $h(\cdot)$ - are the different nonlinear odd functions, $\mathbf{W} = \begin{bmatrix} 0 & w_{12} & \dots & w_{1n} \\ w_{21} & 0 & \dots & w_{2n} \\ \dots & \dots & \dots & \dots \\ w_{n1} & w_{n2} & \dots & 0 \end{bmatrix}$. For the simulations the following parameters have been used: $\eta_0(t) = 2000$, $g(x) = x^3$ and $h(x) = arctg(x)$.

In the next algorithm Infomax it has been shown that maximizing the joint entropy $\mathbf{H}(\mathbf{Y})$, of the output of a neural processor minimizes the mutual information among the output components, $y_i = g(u_i)$, where $g(u_i)$ is an invertible bounded nonlinearity and $\mathbf{u} = \mathbf{W}\mathbf{x}$. For EEG recordings the learning rule can be represented in the following form:

$$\Delta\mathbf{W} = -\mu\left(\frac{\partial\mathbf{H}(\mathbf{Y})}{\partial\mathbf{W}}\right)\mathbf{W}^T\mathbf{W} = \mu(\mathbf{I} + \hat{\mathbf{y}}\mathbf{u}^T)\mathbf{W} \tag{5}$$

where: $\mu = 0{,}01$, $g(u_i) = \frac{1}{1 + \mathrm{e}^{-u_i}}$, $\hat{y} = \frac{\partial}{\partial u_i} ln\left(\frac{\partial y_i}{\partial u_i}\right)$

have been used for simulations.

The third adaptive algorithm - FastICA is a fixed-point algorithm that can be used for estimating the independent components one by one. This algorithm finds one of the columns of the separating matrix and so identifies one independent source within signal duration (Hyvärinen, 1999).

The corresponding independent source signal can be found using the following equation:

$$\hat{\mathbf{s}}(k) = \mathbf{W}^T\mathbf{v}(k) \tag{6}$$

where: $\mathbf{V}$ - is a whitening matrix given by $\mathbf{V} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{U}^T$, $\mathbf{\Lambda} = \mathrm{diag}[\lambda(1),...,\lambda(m)]$ - is a diagonal matrix with the eigenvalues of the data covariance matrix $E\{\mathbf{x}(k)\mathbf{x}(k)^T\}$, $\mathbf{U}$ - is a matrix with the corresponding eigenvectors as its columns. Each lth iteration of this adaptive algorithm is defined as:

$$\mathbf{w}^*_l = E\{\mathbf{v}(\mathbf{w}_{l-1}^T\mathbf{v})^3\} - 3\mathbf{w}_{l-1} \;;\; \mathbf{w}_l = \mathbf{w}^*_l / \|\mathbf{w}^*_l\| \tag{7}$$

4 RESULTS

The results of comparisons of three selected algorithms are presented below. Figure 4 shows a 2-sec interval of an EEG time series and ‘corrected’ EEG signals obtained by removing selected artifacts using different adaptive algorithms.

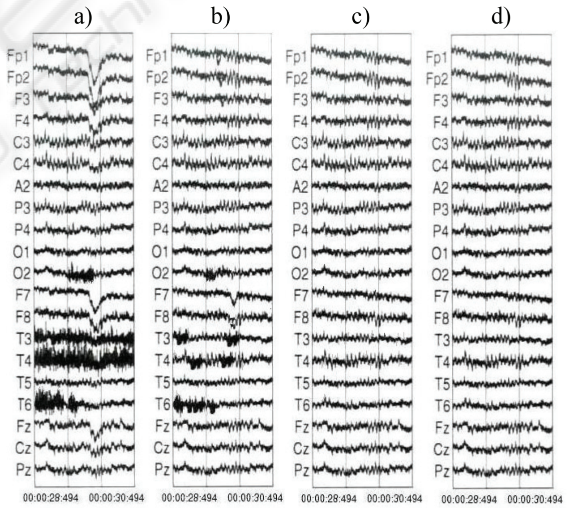

Figure 4: Plots illustrating EEG recordings: a) a set of normal EEG signals affected by the artifacts: eye blinks and muscle activity and the EEG signals after removal of these artifacts using the following algorithms: b) HJ, c) Infomax , d) FastICA.

The algorithms are compared using the coefficient ε - the difference between an estimate of source signals $y_n(t)$ and original EEG signals $s_n(t)$ (without artifacts) defined for different channels in the following form:

$$\varepsilon_n(t) = \sum_{n=1}^{m} \left[y_n(t) - s_n(t) \right] \qquad (8)$$

For ideal case, when the perfect removal is achieved, the coefficient ε is zero. In these simulations any of the presented adaptive algorithms cannot remove all artifacts, but only minimizes their influence on desired EEG signals.

Below, it is presented how error quantity depends on the type of adaptive algorithm and the type of channel.

a) b) c)

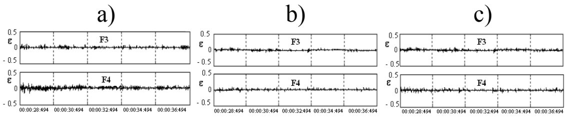

Figure 5: Plots illustrating error signals ε versus time function for eye blinks: a) HJ, b) Infomax, c) FastICA.

a) b) c)

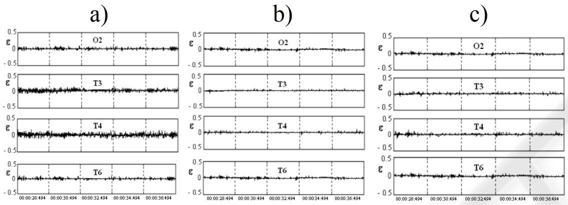

Figure 6: Plots illustrating error signals ε versus time function for muscle activity: a) HJ, b) Infomax, c) FastICA.

5 CONCLUSIONS

The paper presents selected adaptive algorithms and compares the performance of three separation algorithms of the EEG signals in the presence of two types of artifacts.

Biomedical source signals are usually distorted by different artifacts. Besides classical signal analysis tools (such as adaptive supervised filtering, parametric or non-parametric spectral estimation, time-frequency analysis) the proposed ICA approach can be used for detection and reduction of artifacts from EEG recordings.

During tests, it has been observed that the proposed adaptive algorithms can effectively detect and remove these selected artifacts, but their effectiveness depends on the type of artifact and on the type of channel (Figure 5, Figure 6).

ACKNOWLEDGEMENTS

This work is supported by MNiSW (of Government Administration in Poland) - Grant no. N N518 335035 (2008÷2010).

REFERENCES

Bell A. J., Sejnowski T. J., 1995. An information - maximization approach to blind separation and blind deconvolution. *Neural Computation, 7:1129-1159.*

Cichocki A., Amari S., 2002. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, Wiley, England.

Cichocki A., Amari S., Siwek K., Tanaka T., Phan H., 2007. http://www.bsp.brain.riken.jp/ICALAB.

Girolami M., 2000. *Advances In Independent Component Analysis*, Springer, London.

Hyvärinen A., 1999. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks 10(3):626-634.*

Jutten C., Herault J., 1991. Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing, 24:1-10.*

Majkowski J., 1986. *Elektroencefalografia kliniczna*, Państwowy Zakład Wydawnictw Lekarskich, Warszawa.

Rangayyan R. M., 2002. *Biomedical Signal Analysis*, Wiley, New York.

Roberts S., Everson R., 2001. *Independent Component Analysis: Principles and Practice*, Cambridge University Press.

Sanei S., Chambers J. A., 2007. *EEG Signal Processing,* Wiley, Cardiff.

Van Hulle M., 2008. Sequential Fixed-Point ICA Based on Mutual Information Minimization. *Neural Computation, Vol. 20, No. 5: 1344–1365.*