Jacobi Iteration for Spatially Constrained Independent Component Analysis

By Maarten De Vos[1], Lieven De Lathauwer[2], Anneleen Vergult[1], Wim De Clercq[1], Wim Van Paesschen[3] and Sabine Van Huffel[1]

[1]Department of Electrical Engineering, ESAT-SCD, Katholieke Universiteit Leuven, Leuven, Belgium
[2]ETIS (CNRS, ENSEA, UCP), Cergy-Pontoise, France
[3]Neurology Department, University Hospital Gasthuisberg, Katholieke Universiteit Leuven, Leuven, Belgium

Abstract

Electroencephalography (EEG) measures the potentials generated by the brain. Unfortunately, eye artifacts often contaminate EEG recordings. Independent Component Analysis (ICA) has been proposed to remove these artifacts. Standard implementations of ICA extract independent components from a set of measured signals in a fully blind way. In the case of eye artifacts, there is strong *a priori* information about the mixing vector of this artifact component. We provide an algorithm which incorporates this knowledge into the calculation of the independent sources and use this constrained ICA technique to remove eye artifacts from the EEG. The advantage of this method is that the components have not to be inspected afterwards in order to detect the artifactual ones.

1. Introduction

The electroencephalography (EEG) is a non-invasive tool that offers millisecond temporal resolution for the study of neural mechanisms underlying mental activity. Unfortunately, the EEG is frequently and often severely contaminated by artifacts originating from different non-brain sources, mainly scalp muscles, eye blinks and eye movements (Niedermeyer and da Silva 2005). These artifacts complicate the interpretation of the EEG. For eye-artifacts, several techniques have already been proposed. However, most of them require the availability of peri-ocular electro-oculographic (EOG) electrodes. These EOG signals are not always recorded simultaneously with the EEG in an Epilepsy Monitoring Unit because it is too cumbersome for the patient.

Recently, Independent Component Analysis (ICA) has been introduced in order to blindly separate interesting brain signals from artifacts in EEG recordings (Makeig et al. 1996). ICA is a technique that extracts statistically independent components from a set of measured signals. The main disadvantage of ICA is that the extracted components are not ordered. When someone is only looking for the presence of one component, all the extracted components have to be inspected. However, in some applications, there is *prior* knowledge about the shape of the source signals being sought. This has led to the introduction of temporally constrained ICA (James and Gibson 2003). Another interesting way to constrain ICA is to impose constraints on the columns of the mixing

vector. In this paper, we show how prior knowledge can be incorporated about the spatial projections of a source to be estimated. We also use this spatially constrained ICA in order to remove eye artifacts from the EEG.

2. Independent Component Analysis (ICA)

2.1. *The concept of ICA*

Assume the basic linear statistical model

$$Y = \mathbf{M} \cdot X + N \tag{2.1}$$

where $Y \in \mathbb{R}^I$ is called the observation vector, $X \in \mathbb{R}^J$ the source vector and $N \in \mathbb{R}^I$ additive noise. $\mathbf{M} \in \mathbb{R}^{I \times J}$ is the mixing matrix.

The goal of Independent Component Analysis is to estimate the mixing matrix $\mathbf{M}$, and/or the source vector X, given only realizations of Y. In this study, we assume that $I \geqslant J$.

Blind identification of $\mathbf{M}$ in (2.1) is only possible when some assumptions about the sources are made. One assumption is that the sources are mutually statistically independent, as well as independent from the noise components and that at most one source is gaussian (Cardoso and Souloumiac 1994).

An other assumption is that the sources are mutually uncorrelated but individually correlated in time. Under these assumptions, the sources can be estimated with the SOBI algorithm (Belouchrani et al. 1997). This algorithm has shown to be useful in several biomedical applications.

Most algorithms start with a pre-whitening step, in which the observation vector Y is transformed into another stochastic vector with unit variance. During this step, also a projection on the signal subspace can be carried out. For more details, we refer to (De Lathauwer et al. 2000).

2.2. *SOBI*

When the sources are individually correlated in time, but mutually uncorrelated, an ICA algorithm based on second order statistics can be derived (Belouchrani et al. 1997).
Mathematically, this means that for all time lags τ the source correlation matrices are diagonal.

$$\mathbf{R}_y(\tau) = E\{Y(t)Y(t+\tau)^T\} \tag{2.2}$$

$$= \mathbf{A} \cdot \mathbf{R}_x(\tau) \cdot \mathbf{A}^T \qquad \forall\tau \tag{2.3}$$

where $\mathbf{R}_x$ is the correlation matrix of the source signals. Considering that this equation holds for all values of τ, the mixing matrix $\mathbf{M}$ is the matrix $\mathbf{A}$ that jointly diagonalizes all the correlation matrices.

This simultaneous diagonalization can be computed by a Jacobi iteration (Cardoso and Souloumiac 1994).

2.3. *Jacobi angles for simultaneous diagonalization*

In (Cardoso and Souloumiac 1994), it is shown that the optimal Jacobi angles for simultaneous diagonalization are derived from an eigenvalue problem. In order to maximise the diagonal elements of $\mathbf{J}^H\mathbf{R}_y(\tau)\mathbf{J}$ with $\mathbf{J}$ a unitary rotation matrix, we write the cost

function as

$$f = \sum_p |\tilde{u}_{iip}|^2 \tag{2.4}$$

where
$$\tilde{u}_{iip} = (c \quad -s) \begin{pmatrix} u_{iip} & u_{ijp} \\ u_{jip} & u_{jjp} \end{pmatrix} \begin{pmatrix} c \\ -s^* \end{pmatrix}. \tag{2.5}$$

u_{ijp} denotes the (i,j)th entry of the pth correlation matrix after rotation, $i \leqslant N < j$ and $c = cos(\theta), s = sin(\theta)\mathrm{e}^{j\alpha}$. Note that we write the equations for Jacobi angles in a general, complex notation. The rotation angles θ and α can then be computed from the dominant eigenvector W of the following eigenvalue problem:

$$f = W^T \cdot \mathbf{G} \cdot W \tag{2.6}$$

with
$$W = (cos(2\theta) \qquad sin(2\theta)cos(\alpha) \qquad sin(2\theta)sin(\alpha))^T, \tag{2.7}$$

$$\mathbf{G} = \sum_p \mathrm{Re}(Z_p Z_p^H) \tag{2.8}$$

and
$$Z_p = \frac{1}{2} \begin{pmatrix} u_{iip} - u_{jjp} \\ -(u_{ijp} + u_{jip}) \\ i(u_{ijp} - u_{jip}) \end{pmatrix} \tag{2.9}$$

3. Imposing spatial constraints

3.1. *Weighted constraint*

In order to introduce the constraint that one, say the last, mixing vector should be similar to a known vector $M = (m_1, m_2, \ldots, m_I)$, a second term is added to the cost function (2.6). This second term is the squared projection of the mixing vector $(\mathbf{V}_{new})_J$ at the current iteration onto M: $|M^T(\mathbf{V}_{new})_J|^2$. We have shown that the resulting cost function for a Jacobi iteration can be written as:

$$f = W^T \cdot \mathbf{G} \cdot W + w \cdot g^T W \ (+C) \tag{3.1}$$

w is a scalar that allows one to weight the two terms in the cost function according to the confidence one has in M. C is an irrelevant constant.

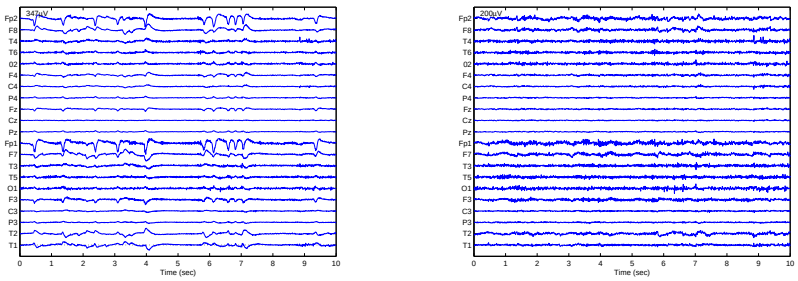
Maximization of this function under the constraint $\parallel W \parallel = 1$, using a Lagrange multiplier, leads to (Belouchrani et al. 1998)

$$W = -\frac{1}{2}(\mathbf{G} + \lambda \mathbf{I})^{-1}\mathbf{g} \tag{3.2}$$

where λ is a real scalar chosen in such a way that

$$\parallel W \parallel = 1 \Leftrightarrow \frac{1}{4} \sum_{i=1}^{3} \frac{(E_i^T \mathbf{g})^2}{(\lambda_i + \lambda)^2} = 1, \tag{3.3}$$

E_i and λ_i being the eigenvectors and -values of $\mathbf{G}$. As can be seen from (3.3), the problem amounts to rooting a polynomial of degree 6 in the complex case and degree 4 in the real case and selecting the root of which the corresponding W maximizes f in (3.1).



(a) (b)

FIGURE 1. (a) EEG containing eye artifacts; (b) the same EEG with the artifacts removed by spatially constrained ICA

4. Constrained ICA for eye-artifact removal

Eye-blinks are characterized by sharp positive deflections in the most anterior electrodes with a rapid amplitude fall-off at more posteriorly located electrodes (Iwasaki et al. 2005). This *prior* knowledge is used to constrain the ICA computation for eye artifact removal from the EEG. Figure 1a shows background EEG contaminated with eye atifacts. The underlying sources are computed with constrained ICA and the artifactual component is removed. The cleaned EEG is shown in 1b.

REFERENCES

A. BELOUCHRANI, K. ABED-MERAIM, J.-F. CARDOSO AND E. MOULINES 1997 A blind Source Separation Technique using Second Order Statistics. *IEEE Trans. Signal Processing* **45**, 434-444.

A. BELOUCHRANI, K. ABED-MERAIM, Y. HUA 1998 Jacobi-like Algorithms for Joint Block Diagonalization: Applications to Source Localization. *Proc. IEEE Int Work on Intelligent Sig Proc and Comm Sys (ISPACS)*

J.-F. CARDOSO AND A. SOULOUMIAC 1994 Blind beamforming for non-Gaussian signals. *IEE Proc. F* **140**, 362-370.

M. IWASAKI, CH. KELLINGHAUS, A.V. ALEXOPOULOS, R.C. BURGESS, A.N. KUMAR, Y.H. HAND AND H.O. LÜDERS 2005 Effects of eyelid closure, blinks and eye movements on the electroencephalogram. *Clin. Neurophysiol* **116**, 878-885.

CH. J. JAMES AND O. J. GIBSON 2003 Temporally Constrained ICA: An application to Artifact Rejection in Electromagnetic Brain Signal Analysis. *IEEE Trans. Biomed. Eng.* **50**, 1108-1116.

L. DE LATHAUWER, B. DE MOOR AND J. VANDEWALLE 2003 An introduction to independent component analysis. *J Chemometrics* **14**, 123-149.

S. MAKEIG, A.J. BELL, T.-P. JUNG AND T. J. SEJNOWSKI 1996 Independent component analysis of electroencephalographic data. *In M. Mozer and M. Hasselmo, editors. Advances in Neural Information Processing Systems* 145-151.

E. NIEDERMEYER AND F. LOPES DA SILVA 2005 Electroencephalography: basic principles, clinical applications and related fields. *Fifth Edition Philadelphia, PA: Lippincott/Williams & Wilkins*