COMPARISON OF ICA ALGORITHMS FOR ECG ARTIFACT REMOVAL FROM EEG SIGNALS

V. Matic[1], W. Deburchgraeve[2], S. Van Huffel[2]

[1]University of Belgrade, Department of Signals and Systems, Serbia

[2]Katholieke Universiteit Leuven, Department of Electrical Engineering ESAT-SCD,Leuven, Belgium

1 Introduction

In order to obtain accurate information from the Electroencephalogram (EEG) it is necessary to preprocess these signals and remove all artifacts. In this paper we used Independent Component Analysis (ICA) to remove Electrocardiogram (ECG) artifacts from EEG signals. Our goal was to quantify which ICA algorithm [1] would be the most effective. More specifically, we were interested to compare the performance of RobustICA, SOBI, JADE, and CCA algorithms. These algorithms were applied to real-life neonatal EEG data with added artificial artifacts.

2 Methods

The main approach in extracting an ECG artifact is to apply an ICA algorithm to the real-life EEG signal. Outputs of all algorithms will be mutually independent components and some of them will be estimates of the ECG artifact. After we identify them, the EEG can be reconstructed without these corrupting components. In order to validate effectiveness of the ICA algorithms we need to compare the ICA components with the ECG artifact. Unfortunately, we do not have much a priori knowledge about the artifact. In this work, we overcome this problem by creating an artificial ECG artifact that was added to an uncorrupted neonatal EEG signal.

2.1 ECG artifact modeling

After studying properties of ECG artifact signal we decided to model it as a combination of two signals. Those signals should approximate the main characteristics of ECG artifacts that occur on the EEG channels. The first simulated artifact is a spike train signal simulating corrupting QRS complexes of the ECG. The occurrence of these spikes is not strictly periodical but correlated with heart rate. The other simulated artifact is a sine wave with a frequency of 2 Hz. This sine wave corresponds to the pulsation artifact of an electrode close to a blood vessel. The frequency value is chosen as a frequency that is close to the heart rate of neonates. According to ICA standards both of these signals are created to be zero–mean and standardized to a unit variance. In this way ECG artifact is modeled as a combination of a highly dynamic source (spike train) and a slowly varying time–correlated source (sine).

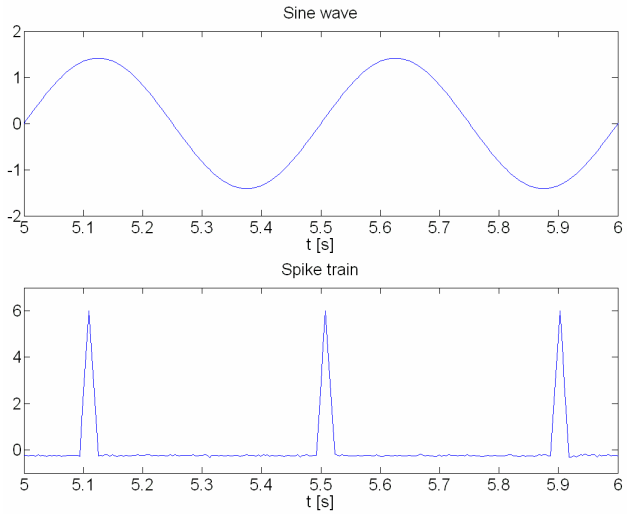


Figure 1. Artificially created ECG artifact sources

In our simulations we used a clean real–life 30 seconds, 20 channel neonatal EEG signal that was recorded with sampling frequency of 256 Hz. We added the spike train artifact source on two randomly selected EEG channels and on the other two channels we added the sine wave. In that way only one artifact was present on a single EEG channel.

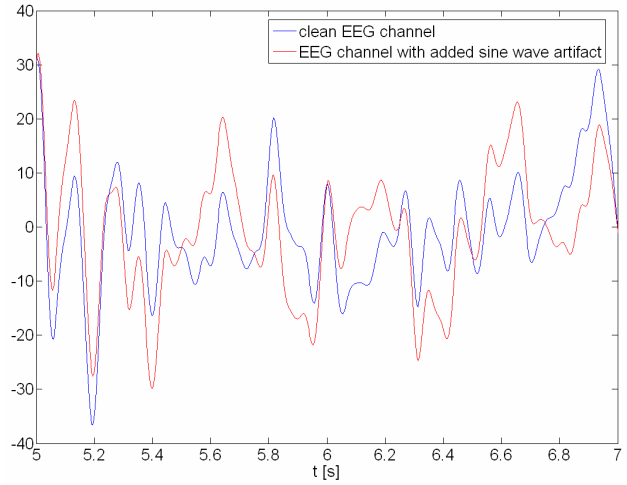


Figure 2. The original signal with added artificial sine wave artifact source

Mathematically this can be formulated as follows:

$$x(t) = EEG(t) + \mathbf{B} \cdot s_{ECG}(t) \quad (1)$$

Thus two new independent sources were added to the EEG to create mixture signal *x*.

2.2 Independent Component Analysis (ICA)

The separation effectiveness of ICA algorithms are related to statistical properties of the mixing signals on which they are applied. Algorithms based on the time structure of the data set have advantages in separating sine shaped signals. In our work we used SOBI and CCA as time structure based algorithms. On the other hand, for separation of the spike train signal, it is more suitable to use algorithms which are based on maximizing non–Gaussianity. Therefore, we selected RobustICA and JADE. It is good to remark that although the main characteristic of a spike train signal is its "spiky" nature and super–Gaussian distribution, this signal is also time correlated to some extent.

2.2.1 ICA Concept

In our work we have used a standard linear ICA model:

$$x(t) = \mathbf{A} \cdot s(t) \quad (2)$$

Here *x* represents a multi channel signal mixture of mutually independent sources *s*. It is necessary that the number of signals (sensor observations) in *x* is not less than the number of sources in *s*. Mixing matrix **A** is unknown and the goal of an independent component analysis is to find an estimate of its inverse matrix **W** such that:

$$y(t) = \mathbf{W} \cdot x(t) \quad (3)$$

Signal *y* represents independent components that are actually estimates of sources *s*. There is one limitation in the ICA method in the sense that an estimated signal y_i cannot determine the variance of a source s_i. That is, there exists an infinite number of factors α_i:

$$y_i(t) = \frac{1}{\alpha_i} \cdot s_i(t) \quad (4)$$

Fortunately, we can always choose α_i in that way that we create a unit variance signal, but this still leaves the ambiguity of the sign.

2.2.2 ICA algorithms

In order to calculate the de–mixing matrix **W**, numerous ICA algorithms have been developed with various approaches. We have used four of them in our comparison and we will shortly overview them:

A. RobustICA

FastICA is arguably one of the most widespread deflation–based ICA algorithms. In the deflation approach, an extracting vector w_i is sought so that the estimate

$$y_i \overset{def}{=} w_i x \quad (5)$$

maximizes the criterion function. FastICA uses the normalized kurtosis as its criterion function

$$K(w) = \frac{E\left\{|y|^4\right\} - 2E^2\left\{|y|^2\right\} - \left|E\left\{y^2\right\}\right|^2}{E^2\left\{|y|^2\right\}} \quad (6)$$

whereas independent components are extracted one by one. RobustICA algorithm represents a simple modification of FastICA which optimizes the kurtosis contrast:

$$\mu_{OPT} = \arg\max_{\mu} K(w + \mu g) \quad (7)$$

The search direction *g* is typically the gradient $g = \nabla_w K(w)$ as explained in [2]. RobustICA is more robust than FastICA and has a very high convergence speed.

B. Second Order Blind Identification (SOBI)

When the sources are individually correlated in time, but mutually uncorrelated, an ICA algorithm based on second order statistics can be derived [3]. Mathematically, this means that for all time lags τ the source correlation matrices are diagonal:

$$\begin{aligned} \mathbf{R}_x(\tau) &= E\left\{x(t)\,x(t+\tau)^T\right\} \\ &= \mathbf{A}\mathbf{R}_s(\tau)\mathbf{A}^T, \qquad \forall \tau \end{aligned} \quad (8)$$

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

C. JADE

Another signal source separation technique is the Joint Approximation Diagonalisation of Eigen matrices (JADE) algorithm [4]. This approach exploits the fourth order moments in order to separate the source signals from mixed signals.

At the beginning, the whitening matrix **P** and the signal z = **P***x* are estimated. After that, the cumulants of the whitened mixtures $\hat{\mathbf{Q}}_i^Z$ are computed. An estimate of the unitary matrix **R** is obtained by

maximizing the criteria $\lambda_i V_i$ by means of the joint digitalization. If $\lambda_i V_i$ cannot be exactly jointly digitalized, the maximization of the criteria defines a joint approximate digitalization. An orthogonal contrast is optimized by finding the rotation matrix **R** such that the cumulant matrices are as diagonal as possible:

$$\mathbf{R} = \arg\min_{\mathbf{R}} \sum_i Off\left(\mathbf{R}^T \hat{\mathbf{Q}}_i^Z \mathbf{R}\right) \tag{9}$$

The mixing matrix **A** is calculated as $\hat{\mathbf{A}} = \mathbf{R}\mathbf{P}^{-1}$ and the independent components are estimated as $y = \hat{\mathbf{A}}^{-1}x = \hat{\mathbf{W}}x$.

D. Canonical Correlation Analysis (CCA)

Canonical correlation analysis solves the ICA problem by forcing the sources to be mutually uncorrelated and maximally correlated with a predefined function [5]. We define a predefined function *z* as a delayed version of a signal *x* with K mixtures and N samples:

$$z(t) = x(t-1) \tag{10}$$

Canonical correlation analysis obtains two sets of basis vectors, one for *x* and the other for *z*, such that the correlations between the projections of the variables onto these basis vectors are mutually maximized. Linear combinations of components in *x* and *z* are:

$$\begin{aligned} x &= w_{x1}\cdot x_1 + \ldots + w_{xK}\cdot x_K = \mathbf{w}_X^T \cdot \mathbf{x} \\ z &= w_{z1}\cdot z_1 + \ldots + w_{zK}\cdot z_K = \mathbf{w}_Z^T \cdot \mathbf{z} \end{aligned} \tag{11}$$

CCA finds the vectors $\mathbf{w}_x$ and $\mathbf{w}_z$ that maximize the correlation ρ between the *x* and *z* by solving the following maximization problem:

$$\begin{aligned} \max_{\mathbf{w}_x,\mathbf{w}_z} \rho(x,z) &= \frac{E\{xz\}}{\sqrt{E\{x^2\}\cdot E\{z^2\}}} \\ &= \frac{\mathbf{w}_x^T \mathbf{C}_{xz} \mathbf{w}_z}{\sqrt{\left(\mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x\right)\left(\mathbf{w}_z^T \mathbf{C}_{zz} \mathbf{w}_z\right)}} \end{aligned} \tag{12}$$

This maximization problem can be solved by setting the derivatives of (12) to zero, which results in the following two eigenvalue problems:

$$\begin{cases} \mathbf{C}_{xx}^{-1}\mathbf{C}_{xz}\mathbf{C}_{zz}^{-1}\mathbf{C}_{zx}\hat{\mathbf{w}}_x = \rho^2 \hat{\mathbf{w}}_x \\ \mathbf{C}_{zz}^{-1}\mathbf{C}_{zx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xz}\hat{\mathbf{w}}_z = \rho^2 \hat{\mathbf{w}}_z \end{cases} \tag{13}$$

After solving this problem we can get the estimates of the sources:

$$y_i(t) = \hat{\mathbf{w}}_{\mathbf{x}_i}^T x(t) \tag{14}$$

2.3 Comparison criteria

In order to compare different ICA algorithms it is necessary to create a function that will match independent components with previously created ECG artifact sources. To accomplish this we had to estimate the correlation of the artifact sources and the independent components. Components with the highest correlation coefficients are selected for comparison as matching components. Independent components are created to be zero–mean and they are standardized to unit variance.

In our work we have varied the impact of the ECG artifact by changing non–zero values in matrix **B**. In that way, our EEG signal was corrupted by ECG artifact ranging from weak to severe. The quality of source separation was estimated with two parameters. We have calculated these parameters for all impact levels of the ECG artifact.

2.3.1 Correlation based criterion

The Spearman correlation coefficient *r* [1] proves to be a good choice to compare the original ECG artifact source and the independent component because it is dependent on the (relative) shape of the signal. It shows normalized values between 0 and 1 regardless of the signal sign. The Spearman correlation coefficient is calculated according to the following formula:

$$r = 1 - 6\sum \frac{d^2}{N\left(N^2-1\right)} \tag{15}$$

where *d* is the difference in statistical rank of the corresponding signal and *N* is the signal length. Correlation index *r* is calculated for every ICA algorithm and for both types of ECG artifact. Good separation quality of an ICA algorithm is indicated by a higher value of *r*.

2.3.2 Root–Mean Square Error (RMSE)

RMSE is used as a standard statistical method and was applied to the source signal *s* (ECG artifact) and to the appropriate matching component *y* (ICA component). Once again, it is important that all signals are zero–mean and standardized to unit variance. ICA algorithms determine components but their sign is unknown. Therefore, before calculating RMSE we must check that our components are not inversed estimates of ECG artifact sources

3 Results

In our work we randomly added 4 artifact signals to the clean 20 channel neonatal EEG signal.

In order to make a fair comparison of algorithms, the same mixing matrix **B** was used for each algorithm. Combination of these channels was varied and we calculated the same parameters which were later averaged. In Table 1 we showed averaged results for the Spearman correlation coefficient *r*:

	RobustICA	SOBI	JADE	CCA
Sine	0.872	0.999	0.723	0.995
Spike	0.689	0.397	0.321	0.399

Table 1. Average Spearman correlation coefficient

Another way to estimate the effectiveness of algorithms is to compare the RMSE for different signal to noise ratios (Fig. 3. and Fig 4.).

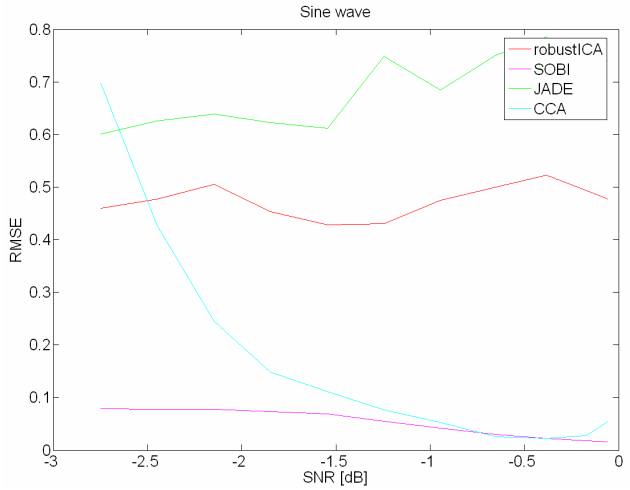


Figure 3. RMSE for sine wave signal

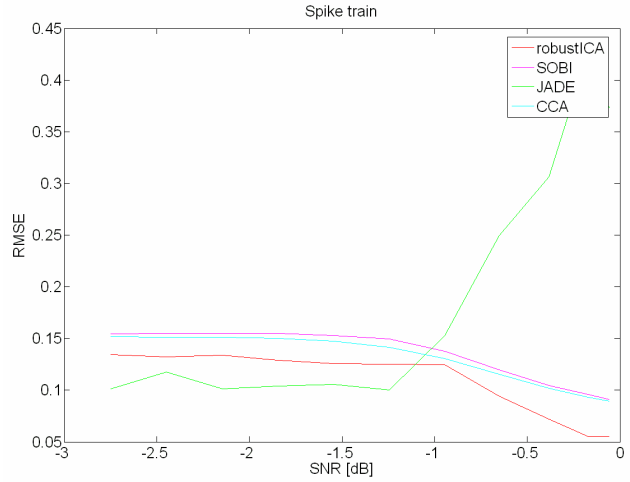


Figure 4. RMSE for spike train signal

4 Discussion

The obtained results do confirm the expected theoretical assumptions that the time structure based algorithms (SOBI, CCA) are efficient in extracting sine waves. Additionally, SOBI showed very good separation performance even for the low amplitudes of ECG artifacts. In separating the spike train signal, algorithms based on maximizing non–Gaussianity (RobustICA, JADE) slightly outperformed time structure based algorithms (SOBI, CCA).

Generally, when all parameters are taken into account, we may conclude that RobustICA should be used for spike train signal extraction, whereas SOBI should be applied for the separation of sine wave shaped signals. As we have estimated that ECG artifact can be composed of both of these signal, we may conclude that SOBI should be used in extracting real ECG artifacts, as the one with the best separation quality.

5 Acknowledgement

This work has been realized while the first author was visiting the BioMed research group at K.U. Leuven. Research supported by: Research Council KUL: GOA-AMBioRICS, GOA MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), IOF-KP06/11 FunCopt, FWO: G.0341.07 (Data fusion), research communities (ICCoS, ANMMM); IWT: TBM070713-Accelero, TBM070706-IOTA3, PhD Grants; Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, `Dynamical systems, control and optimization', 2007-2011); EU: ETUMOUR (FP6-2002-LIFESCIHEALTH 503094), Healthagents (IST–2004–27214), FAST (FP6-MC-RTN-035801), Neuromath (COST-BM0601).

References

[1] M. Klemm, J. Haueisen and G. Ivanova, “Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity”. *Med Biol Eng Comput*, vol 47, pp. 413–423, 2009.

[2] V. Zarzoso and P. Comon, “Comparative Speed Analysis of FastICA”. Proceedings ICA–2007, pp. 293–300, 2007.

[3] A. Belouchrani, K. Abed-Meraim, J.F. Cardoso and E. Moulines, “A blind source separation technique using second order statistics”. *IEEE Trans. Signal Processing*, 45:434–444, 1997.

[4] A. Hyväarinen, J. Karhunen and E. Oja, “Independent Component Analysis”. John Wiley & Sons, 2001.

[5] W. Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen and S. Van Huffel “Canonical Correlation Analysis Applied to Remove Muscle Artifacts From the Electroencephalogram”. *IEEE Transactions on Biomedical Engineering,* 2006.