Analisys of 5 source separation algorithms on simulated EEG signals

Ricardo Salido-Ruiz[2], Rebeca Romo-Vázquez[1], Radu Ranta[1], Lorenzo Leija[2]

[1] Centre de Recherche en Automatique de Nancy
(CRAN - UMR 7039), Nancy-University, CNRS
[2] CINVESTAV- IPN México D.F.
radu.ranta@ensem.inpl-nancy.fr)

Abstract. In this paper we evaluate the performance of 5 BSS algorithms (AMUSE, SOBI, SOBI-RO, SONS, JADE-TD) on simulated EEG signals. A first result evaluates the influence of the noise and signal characteristics (frequency, length, SNR) on the algorithms performance. A second objective is to introduce a new performance criterion, IEV which can be used to compare two matrices and is potentially useful on real signals. We validate this new index by comparing it with classic performance indices used in source separation.

1 Introduction

The electroencephalogram (EEG) is a medical examination based on brain's electric activity. The signal is recorded using electrodes placed on the scalp of the patient. One of the most common brain's diseases investigated through an EEG examination is the epilepsy. Epilepsy is a cerebral disease which is characterized by repeated crisis due to an excessive burst of synchronized neural activity. EEG signals are useful to detect this kind of anomalies as we mentioned before; however, these signals present several inconvenients:

- Recorded brain's signals are corrupted by artifacts (extra-cerebral signals) and noise, which are superimposed to the informative signals and make harder the interpretation for the physicians.
- Scalp EEG signals are by themselves a mixture between intra-cranial unknown sources and its mixing process is itself unknown.

A first step towards an easier interpretation for the physicians can be the development of a technique that allows the elimination of the artifacts and noise that the EEG signals present.

One current hypothesis is that these artifacts are independent from brain activity, either normal or pathologic. Under this hypothesis, a frequently used method is the blind source separation (BSS). The goal of BSS is to recover independent sources, given only sensor observations. This sensor observation is modeled as a linear mixture of independent source signals. The term blind indicates that both the source signals and the way the signals are mixed are unknown. Several algorithms for BSS were developed in the last 15 years [1, 2].

The main objective of this work is to test different source separation algorithms in order to examine their future use on real EEG signals. Besides classical performance criteria, in our opinion it is also necessary to asses the sensitivity of the algorithms to noise type and power, to the mixing model, as well as to the characteristics of the signal (frequency content, duration).

This communication is organized as follows. In the second section, we describe the source separation problem and its relation with EEG; in the third section we present the simulated EEG and noise test signals, and the evaluation criteria. The fourth section presents the obtained results and it is followed by a fifth section that concludes and presents the perspectives of this work.

2 Source Separation Problem

A method for solving the BSS problem it is to find a linear transformation of the measured sensor signals such that the resulting source signals are as statistically independent from each other as possible. The most widely used model considers N the number of unknown sources equal to the number of electrodes In this case the noisy mixture writes:

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{n}(k) \tag{1}$$

where $\mathbf{x}$ is the vector of the mixed signals (sensors), $\mathbf{A} \in \mathbb{R}^{N\times N}$is the unknown nonsingular mixing matrix, $\mathbf{s}$ is the vector of independent source signals, $\mathbf{n}$ is an additive vector noise, k being the time index after sampling.

The objective is to find a linear transformation $\mathbf{B}$ of the sensor signals $\mathbf{x}$ that makes the outputs as independent as possible:

$$\mathbf{y}(k) = \mathbf{B}\mathbf{x}(k) = \mathbf{B}\mathbf{A}\mathbf{s}(k) + \mathbf{B}\mathbf{n}(k) \tag{2}$$

where, $\mathbf{y}$ is the estimation of the sources and $\mathbf{B}$ is the separation matrix. The ideal separation is obtained when $\mathbf{B} = \mathbf{A}^{-1}$ and, consequently, $\mathbf{y}$ is a (noisy) estimate of $\mathbf{s}$.

As it has been pointed out by different authors [1, 2], obtaining the exact inverse of the $\mathbf{A}$ matrix is, in most of the cases, impossible. Therefore, source separation algorithms search a $\mathbf{B}$ matrix such as the product $\mathbf{BA}$ is a permuted diagonal and scaled matrix. Consequently, sources can be recovered up to their order (permutation) and their amplitude (scale).

Many different algorithms are available; these can be summarized by the following fundamental approaches, depending on the cost functions minimized to find the separation matrix $\mathbf{B}$:

– The most popular approach exploits as cost function some measure of signals statistical independence, non-gaussianity or sparseness. When original sources are assumed to be statistically independent (regardless of their temporal structure) the higher-order statistics (HOS) are essential (implicitly or explicitly) to solve the BSS problem. In such case, the method does not allow more than one Gaussian source [1–3].

– If sources have temporal structures, then each source has non-vanishing temporal correlation, and less restrictive conditions than statistical independence can be used, namely, second-order statistics (SOS) are often sufficient to estimate the mixing matrix and sources. As they exploit temporal correlations, SOS methods do not allow the separation of sources with identical power spectra shapes or i.i.d. (independent and identically distributed).

Most of BSS methods (HOS and SOS) include a SOS only pre-processing step: the spatial decorrelation or whitening. The conventional whitening exploits the equal-time correlation matrix of the data $\mathbf{x}$, which is often considered a necessary criterion, but not sufficient for the independence. The whitening of $\mathbf{x}$ consists of the decorrelation and the normalization of its components. The idea is to find a matrix $\mathbf{W_b}$ known as whitening matrix, such as,

$$\bar{\mathbf{x}} = \mathbf{W_b}\mathbf{x} \tag{3}$$

with the covariance matrix of $\bar{\mathbf{x}}$ equal to the identity matrix: $\mathbf{R_{\bar{x}}} = \mathbf{I}$. One can show that the whitening matrix $\mathbf{W_b}$ can be written as:

$$\mathbf{W_b} = \mathbf{\Sigma}\mathbf{V^T} \tag{4}$$

where $\mathbf{\Sigma}$ is a diagonal matrix and $\mathbf{V}$ an orthogonal matrix, obtained from the eigen decomposition of $\mathbf{R_x}$, the covariance matrix of the data.

Independent estimates of the sources will be obtained from the whitening signals $\bar{\mathbf{x}}$ by a second transformation:

$$\mathbf{y} = \mathbf{J}\bar{\mathbf{x}} = \mathbf{J}\mathbf{W}_b\mathbf{x} \tag{5}$$

As the covariance matrix $\mathbf{R_y}$ has to be also equal to the identity matrix (estimates are independents, so uncorelated), $\mathbf{J}$ is necessarily an orthogonal matrix. The minimisation of the cost functions reminded earlier leads to this matrix.

Another whitening method is the robust whitening based on time-delayed correlation matrices. This method is used to minimize influence of the (white) noise in different algorithms (SOBI-RO, SONS, AMUSE).

Temporal, spatial and spatio-temporal decorrelations play important roles in EEG/MEG data analysis. Therefore, a lot of algorithms used in this domain are based only on second-order statistics (SOS) [5–8], although other authors prefer HOS algorithms [11–13].

The 5 algorithms compared in this work are:

1. JADE-TD (HOS - Joint Approximate Diagonalization of Eigen matrices with Time Delays), uses a combination of source separation algorithms of second order time structure (TDSEP) [14] and high order cumulant information (JADE) [15]. In principle, it is able to separate simultaneously time-correlated and non-Gaussian signals [10].
2. SOBI (SOS - Second Order Blind Identification), is an algorithm adapted for temporally correlated sources. It is based on the 'joint diagonalization' [5] of an arbitrary set of covariance matrices and relies only on second-order statistics of the received signals. It allows separation of Gaussian sources [9].

3. SOBI-RO (SOS - Robust SOBI with Robust Orthogonalization), combines robust whitening (in the presence of temporally uncorrelated additive noise) and time-delayed decorrelation, as SOBI. It improves the classical SOBI method by integrating robust whitening instead of simple whitening [6].
4. AMUSE (SOS - Algorithm for Multiple Unknown Source Extraction), is based on the EVD (eigenvalue decomposition) of a single time-delayed covariance matrix for prewhitened data. This algorithm also integrates a method for ordering automatically the estimated sources [7].
5. SONS (SOS - Second Order Nonstationary Source Separation) algorithm exploits the nonstationarity and temporal structure of the sources. This method needs only multiple time-delayed correlation matrices of the observed data at several different time-windowed data frames to estimate the mixing matrix. This algorithm is not sensitive to additive white noise [8].

3 Method

The goal of our evaluation is to assess the performance for the 5 BSS algorithms described before. They were tested on different simulated signals, mixing matrices and noise vectors, also they were compared using several evaluation criteria.

3.1 Simulated Signals

In order to test the BSS algorithms it was necessary to create different signals that simulate the EEG's source signals. We propose four signals having different characteristics close to the real EEG and a fifth signal simulating the eye blinking artifact. Four test sets were created using these five signals:

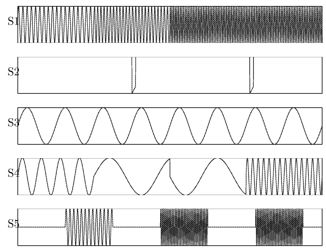
- The first one figure 1(a), having 2048 samples/signal and frequencies ranging from 0.5 Hz to 26 Hz.
- A second set contained only the first half of the previous one (1024 samples), thus having lower frequencies (range 0.5 Hz - 10 Hz).
- A 3^{rd} one contained the second half (1048 samples, mainly high frequencies).
- The last one was made duplicating the original set, having thus a 4096 samples signal and the same frequency range.

As we mentioned above the signals of each set were mixed using random mixing matrices (uniform distribution between -1 and 1). At the resulting mixture, two types of noises (Gaussian and Uniform) were added, with five different signals to noise ratios (0 dB, 5 dB, 10 dB, 15 dB and 20 dB).

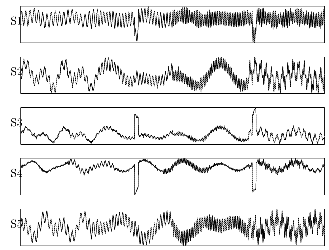
Figure 1(b) presents an exemple of the simulated EEG (the noisy mixture of the 5 sources from figure 1(a)).

3.2 Evaluation Criteria

Index of Separability (IS) To validate the separation, the first criterion we have chosen is the index of separability IS [4]. The index is computed from the



(a)



(b)

Fig. 1. Simulated EEG's: (a) original sources for a recording length of 8 seconds (sampling rate 256 Hz) resulting in signals of 2048 samples. Above (s1), a signal with 4 successive frequencies (9 Hz, 12 Hz, 14 Hz and 23 Hz), (s2) eyes artifact, (s3) a low frequency signal (1 Hz), (s4) a signal varying in frequency (2 Hz, 0.5 Hz, 7 Hz), (s5) three bursts of 10 Hz, 10 Hz and 26 Hz frequencies respectively; (b) noisy mixture.

$N \times N$ transfer matrix $\mathbf{G}$=$\mathbf{BA}$ between the original sources and the estimated ones. In order to obtain the IS it is necessary to take the absolute value of the elements of $\mathbf{G}$ and to normalize the rows $\mathbf{g}_i$ by dividing each element by the maximum absolute value of the row. The rows of the resulting matrix $\mathbf{G}'$ are:

$$\mathbf{g_i}' = \frac{|\mathbf{g_i}|}{\max|\mathbf{g_i}|} \tag{6}$$

The separability index is obtained from the new $\mathbf{G}'$ matrix:

$$IS = \frac{\sum_{j=1}^{N}\left(\sum_{j=1}^{N}\mathbf{G}'(i,j) - 1\right)}{N(N-1)} \tag{7}$$

Correlation (CC) The second criterion is the correlation ρ between the simulated sources $\mathbf{s}_i$, $i = 1...N$ and the estimated independent components $\mathbf{y}_j$. To avoid taking into account small values of correlation and estimated sources correlated with more than one original source, correlation values smaller than 0.5 were discarded:

$$\rho_{ij} = \frac{\texttt{cov}(\mathbf{s}_i\mathbf{y}_j)}{\sigma_{\mathbf{s}_i}\sigma_{\mathbf{y}_j}} \qquad r_{ij} = \begin{cases} \rho_{ij} & \text{if } \rho_{ij} \geq 0.5 \\ 0 & \text{if } \rho_{ij} < 0.5 \end{cases} \tag{8}$$

The retained estimated source $\mathbf{y}_j$ for the source $\mathbf{s}_i$ is the one for which the correlation is maximal $r_i = \max r_{ij}$. Finally, the CC criterion is defined as:

$$CC = \frac{1}{N}\sum_i r_i \tag{9}$$

Eigen Values Vector's Norm-1 Distance(IEV) Besides the previous evaluation criteria, we propose a new one base on the eigen values of the mixing and separation matrices. The basic idea is that, if the separation is successful, the mixing matrix and the inverse of the separation matrix must be similar (after reordering and normalization). The goal of the method is to evaluate this similarity. As seen previously (section 2), the BSS algorithms cannot find source estimates in the same order, with the same amplitudes and with the same signs as the original signals. Thus the inverse of the separation matrix, which is supposed to be equal to the mixing matrix has different values and the order of its rows may be changed. Therefore, before comparing two steps must be taken:

1. Normalization: first, the elements in $\mathbf{B}$ and in $\mathbf{A^{-1}}$ are taken in absolute values (to eliminate sign ambiguity). Next, each line is divided by his maximum value to normalize it to a maximum value of 1.
2. Permutation: the lines in matrix $\mathbf{A^{-1}}$ are permuted, using the same method as for the correlation index.

In this way, we obtain $\mathbf{A_{inp}}$, the inverted, normalized and permuted version of $\mathbf{A}$ and $\mathbf{B_n}$, the normalized version of $\mathbf{B}$. To compare the previously obtained matrices, we chose to compute their eigen values. For each matrix, we construct a vector containing their eigen values and we compute the norm 1 distance between those two vectors:

$$IEV = \left| \sum_{i=1}^{N} \left(\lambda_{A_{inp}} - \lambda_{B_n} \right) \right| \quad (10)$$

where IEV is the newly obtained index, $\lambda_{A_{inp}}$ and λ_{B_n} are the eigen values of the described matrices.

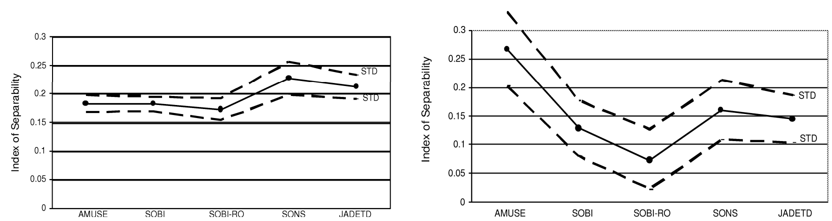
So, if we have two equal matrices (mixing matrix = inverse of separation matrix) our index will be zero, indicating that the separation was performed at 100%. Generally speaking, a small value of IEV indicates good separation.

4 Results

The first aspect evaluated in this work is the behavior of the BSS algorithms using different noise vectors and mixing matrices. The second objective of our work is to analyse the tested algorithms according to their performances for all combination of signal and noise, as described previously. Finally, a third goal is to evaluate the accuracy of our new performance index IEV by comparing it with the two others.

In order to evaluate the influence of the random noise on the separation index, we created 10 simulated EEG (noisy mixtures of the original sources 8s) by using one mixing matrix and 10 noise vectors (Gaussian noise 15 dB). The results (mean value and standard deviation STD) are presented in figure 2(a).

As we can see in the figure 2(a), the standard deviation of the IS index is small and affects the methods almost in the same way. We conclude from this simulation that the influence of the noise on the separation index is rather small (for a given SNR and probability law). Other simulations (not presented here), show that the IS values are much more influenced by the probability law of the noise and especially by the noise power (SNR).



(a) (b)

Fig. 2. Mean and STD for (a) one mixing matrix and 10 different noise vectors (b) 1000 different mixing matrices and noise vectors.

Concerning the influence of the mixing matrix on the index of separation (IS), we created 1000 simulated EEG (noisy mixtures of the original sources 8s) using the 1000 different mixing matrices and one noise vector for each one (Gaussian noise 15 dB). The results (mean value and standard deviation of the IS) are presented in figure 2(b). As we can see the relative standard deviation of the IS is larger than the one showed in figure 2(a). Our interpretation of this result is that, in order obtain a robust evaluation of the algorithms performances, it is necessary to test them by using an important number of mixing matrices, but it is not necessary to simulate an important number of noise vectors. A second conclusion is that SOBI-RO algorithm shows the best performances, at least for this signal and noise combination.

Therefore, for the following simulations, we only generated one noise vector for each particular situation (a given signal, a given mixing matrix, a given SNR and a given probability of the noise). We used the 4 signal sets previously presented, 1000 random mixing matrices, 5 signal to noise ratios (SNR = 0, 5, 10, 15, 20 dB) and 2 noise probability distributions, which leads us to 40000 simulations of noisy mixtures (10000 for each signal). We added also a no noise simulation, again using 1000 mixing matrices for each of the 4 signals (4000 simulations of no noise mixtures). The averaged results are presented in the next tables.

In table 1, we can see the behavior of the BSS algorithm, being evaluated for four evaluation criteria. The results obtained for each evaluation criterion are coherent with the Index of Separability (IS), and show that the best BSS algorithms are founded between SOBI and SOBI-RO.

Table 1. Results obtained for 4s-1 for the 5 algorithms.

	AMUSE	SOBI	SOBIRO	SONS	JADETD	AMUSE	SOBI	SOBIRO	SONS	JADETD
	without noise					with noise				
CC	0.97	0.99	0.99	0.88	0.95	0.56	0.68	0.65	0.57	0.64
IS	0.07	0.03	0.03	0.16	0.12	0.31	0.17	0.13	0.2	0.22
IEV	0.56	0.43	0.44	0.57	0.76	0.74	0.71	0.65	0.71	1.21

Table 2. Results obtained for 4s-2 for the 5 algorithms.

	AMUSE	SOBI	SOBIRO	SONS	JADETD	AMUSE	SOBI	SOBIRO	SONS	JADETD
	without noise					with noise				
CC	0.99	0.99	0.99	0.94	0.98	0.58	0.69	0.65	0.59	0.65
IS	0.04	0.02	0.03	0.12	0.05	0.3	0.16	0.11	0.19	0.2
IEV	0.63	0.4	0.41	0.49	0.71	0.74	0.67	0.61	0.66	1.16

In table 2, the results of the BSS algorithms seem to be improved. We can see it if we compare the values obtained in table 1 with the ones obtained in table 2. Both sets of signals have a 4s duration, but there is a big difference between their frequency content: 4s-1 has low frequency components, 4s-2 who presents mainly high frequencies.

Table 3. Results obtained for 8s for the 5 algorithms.

	AMUSE	SOBI	SOBIRO	SONS	JADETD	AMUSE	SOBI	SOBIRO	SONS	JADETD
	without noise					with noise				
CC	1	1	1	0.93	1	0.57	0.68	0.65	0.59	0.67
IS	0.04	0.02	0.03	0.12	0.04	0.29	0.16	0.11	0.18	0.18
IEV	0.64	0.41	0.42	0.52	0.67	0.71	0.7	0.61	0.67	1.162

In table 3 we can observe the results obtained for a larger length, and high and low frequency content signals. The results are even better than those presented for small length signals. We can see it by comparing the values obtained for the evaluation criteria. The hypothesis here is that a large amount of data (and also higher frequencies) facilitate the separation procedure for the BSS algorithms. Again we see that SOBI and SOBI-RO are the BSS algorithms with the best performance.

Another important observation (not shown in these tables which only present average values) is that between the behavior of these algorithms present small variations when the values are compared for uniform and Gaussian noises in each SNR level.

Table 4. Results obtained for 16s for the 5 algorithms.

	AMUSE	SOBI	SOBIRO	SONS	JADETD	AMUSE	SOBI	SOBIRO	SONS	JADETD
	without noise					with noise				
CC	1.00	1.00	1.00	0.93	1.00	0.57	0.68	0.60	0.58	0.62
IS	0.04	0.02	0.03	0.12	0.04	0.29	0.16	0.10	0.18	0.17
IEV	0.64	0.40	0.41	0.54	0.70	0.72	0.68	0.57	0.69	1.16

The results shows different aspects of the behavior of the BSS algorithms. A first point is that for the sets of five simulated signals with short length (4 seconds) the evaluation criteria show worst performances for all the noise conditions and mixing matrices. The frequency content of the simulated signals played an important role in the performance of the BSS algorithms. Higher frequency signals are better separated. The set of simulated signals with the shortest duration and lower frequency content was the one with the worst performance for all performance criteria. The behavior of the BSS algorithms with respect to the type of the added noise was not so important. Very similar values were obtained for uniform and Gaussian noises. On the contrary, the noise level affected the separation performance for all signals. Almost all the algorithms struggle to obtain a good source separation when the signal to noise ratio is lower than 10 dB. Globally, all the results show that SOBI-RO and SOBI are the BSS algorithms which obtained the best scores in all the tested evaluation criteria the source separation performance. The IEV criterion indicate the same behaviour as the IS and CC indices.

5 Conclusion and Perspectives

As we mentioned above, the main objective of this work is to test different source separation algorithms in order to examine their future application on real EEG signals. Here we have presented a methodology that allows us to compare the BSS algorithms taking in account the nature of the EEG simulated signals. Namely, we simulated signals with different length, frequency, type and noise levels. Also, the important simulation number played an important role in our evaluation, because of the big and reliable data base which better support our conclusions.

Different interesting aspects of the behavior of the algorithms were presented here: the noise does not seem to play an important role in the performance of the separation except for its power. On the contrary the influence of the mixing matrix is much more important.

Some of the BSS algorithms presented considerable changes associated to the signal length, and to its frequency content. High frecuency long signals are better separated than low frecuency short ones. However, all the evaluation criteria show that SOBI-RO and SOBI algorithms are the best for the source separation on our simulated signals.

The new introduced criterion IEV allows us to compare directly two matrices and it aims to give a measure fo a kind of 'distance' between them: if its value es close to zero, the compared matrices are similar. Used on the (normalized) mixing marice and on the inverse of the normalized separation matrix, it gives similar indications as the other performance indices (IS and CC), which proves that IEV is a reliable performance criterion. Its main interest is its possible aplication on real signals.

References

1. Te-Won Lee: 'Independent component Analysis Theory and Applications' Kluwer Academic Publisher. Boston, 1998.
2. A. Cichocki, Shun-ichi Amari 'Adaptative blind Signal and Image Processing Learning Algorithms and Applications' John Wiley and Sons, ltd, 2002.
3. S. Choi, A. Cichocki, H. Park, S. Lee, 'Blind Source Separation and Independent Component Analysis : A Review' Neural Information Processing - Letters and Reviews, Vol. 6, no. 1, January 2005.
4. A. Cichocki, S. Amari, K. Siwek, T. Tanaka et al., ICALAB Toolboxes, http://www.bsp.brain.riken.jp/ICALAB.
5. A. Belouchrani, K. Abed-Meraim, J.F. Cardoso, and E. Moulines, 'Second-order blind separation of temporally correlated source', Proc. Int. Conf. on Digital Sig. Proc., (Cyprus), pp. 346-351, 1993.
6. A. Belouchrani, A. Cichocki, 'Robust whitening procedure in blind source separation contex' Electronics Letters, vol. 36, No. 24,pp. 2050-2053, 2000.
7. L. Tong, V. Soon, Y. F. Huang, and R. Liu, 'Indeterminacy and identifiability of blind identification', IEEE Trans. CAS, vole. 38, pp. 499-509, March 1991.
8. S. Choi and A. Cichocki, 'Blind separation of nonstationary sources in noisy mixtures', Electronics Letters, Vol. 36, pp. 848-849, April 2000.
9. A. Belouchrani, K. Abed-Meraim, J.F. Cardoso and E. Moulines, 'A blind source separation technique using second order statistics', IEEE Trans. on Signal Processing, vol. 45, No. 2, pp. 434-444, February 1997.
10. K.-R. Mller, P. Philips, and A. Ziehe, 'JADEtd: Combining Higher-Order statistics and temporal information for blind source separation with noise', Proc. Int. Workshop on Independent Component Analysis and Blind Separation of Signals (ICA '99), Aussois, 1999.
11. A. J. Bell and T. J. Sejnowski 'An information maximisation approach to blind separation and blind deconvolution', Neural Computation, 7, 6, 1129-1159, 1995.
12. M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, Kindermann, and T. J. Sejnowski, 'Analysis of fMRI by Blind Separation into Independent Spatial Components', Human Brain Mapping, 6(3):160-88 (1998).
13. A. Hyvrinen and E. Oja, 'A Fast Fixed-Point Algorithm for Independent Component Analysis', Neural Computation, 9(7):1483-1492, 1997.
14. A. Ziehe, K. - R. Mller, ' 'TDSEP - an efficient algorithm for blind separation using time structure', International Conference on Artificial Neural Networks, , Sweden, 2 - 4 September 1998.
15. Jean-Franois Cardoso, Antoine Souloumiac, 'Jacobi Angles For Simultaneous Diagonalization', SIAM J. Matrix Anal. Appl. 17 (1), 161 - 164, 1996.