ESANN'2006 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2006, d-side publi., ISBN 2-930307-06-4.

Semi-Blind Approaches for Source Separation and Independent component Analysis

Massoud Babaie-Zadeh[1] and Christian Jutten[2] *

1- Sharif University of Technology - Electrical Engineering Department
Azadi avenue, Tehran - Iran

2- INPG - LIS (UMR CNRS 5083)
46, avenue Félix Viallet, 38031 Grenoble Cedex - France

Abstract. This paper is a survey of semi-blind source separation approaches. Since Gaussian iid signals are not separable, simplest priors suggest to assume non Gaussian iid signals, or Gaussian non iid signals. Other priors can also been used, for instance discrete or bounded sources, positivity, etc. Although providing a generic framework for semi-blind source separation, Sparse Component Analysis and Bayesian ICA will just sketched in this paper, since two other survey papers develop in depth these approaches.

1 Introduction

Source separation consists in retrieving unknown signals, $\mathbf{s} = (s_1(t), \dots, s_n(t))^T$, which are observed through unknown mixture of them. Denoting the observations $\mathbf{x}(t) = (x_1(t), \dots, x_p(t))^T$, one can write :

$$\mathbf{x}(t) = \mathcal{F}(\mathbf{s}(t)), \tag{1.1}$$

where $\mathcal{F}(.)$ denotes the unknown mixture, a function from $\mathbb{R}^n$ to $\mathbb{R}^p$. If the number of observations p is greater or equal to the number of sources, n, the main idea for separating the sources is to estimate a transform $\mathcal{G}(.)$ which inverses the mixture $\mathcal{F}(.)$, and, without extra effort, provides estimates of the unknown sources.

Of course, without other assumptions, this problem cannot be solved. Basically, it is necessary to have priors about

- the nature of the mixtures : it is very important to chose a separating transform $\mathcal{G}(.)$ suited to the mixture transform $\mathcal{F}(.)$,
- the sources : sources properties - even weak - are necessary for driving the $\mathcal{G}(.)$ estimation.

Because of the very weak assumptions, the problem is referred as blind source separation (BSS), and the method based on the property of source independence

*This work has been partially funded by Sharif University of Technology, and by Center for International Research and Collaboration (ISMO) through the France-Iran integrated action program (Gundi-Shapour).

has been called independent component analysis (ICA) [1, 2].

In fact, one often has priors on signals. A natural idea is to add these priors in the model, for simplifying or improving the separation methods. This paper, with its two companions, the Gribonval's paper on sparse component analysis (SCA) [3] and the Mohamad-Djafari's paper on Bayesian ICA [4], constitute a review of semi-blind methods in which various priors are used. For this reason, Sparse and Bayesian ICA, although related to generic approaches for, will not be developed in this paper.

This paper is organized as follows. In Section 2, we recall usual assumptions in the so-called blind source separation. Section 3 is devoted to Gaussian non iid sources. In Section 4, we show that prior like discrete-valued or bounded sources leads to simple geometrical algorithms. In Section 5, we suggest other priors, like video cue for enhancing speech separation, or positivity. Section 6 is a short conclusion, which briefly presents the papers of the special session on semi-blind source separation (SBSS) of the conference ESANN 2006.

2 Blind source separation

Source separation methods have been developed intensively for linear mixtures, instantaneous as well as convolutive, and more recently by a few researchers for nonlinear mixtures. In the most general case, the only assumption done on the sources is that they are statistically independent. From Darmois's result [5], one deduces that this problem has no solution for mutually independent sources which are Gaussian, with (temporally) independent and identically distributed (iid) samples. Then, since the Gaussian iid model has no solution, one must add priors, which are threefold [6]:

- Non Gaussian iid,
- Gaussian but non temporally independent (first *i* of *iid*), *i.e.* temporally correlated,
- Gaussian, but non identically distributed (*id* of *iid*), *i.e.* non stationary sources.

Initially, even if it was not clearly stated [7], the problem has been related to the non Gaussian iid model, and has been refered as blind source separation (BSS). The non Gaussian property appears clearly in the Comon's theorem [2] for linear mixtures.

Theorem 2.1 *Let* $\mathbf{x} = \mathbf{A}\mathbf{s}$ *be a p-dimension regular mixture of mutually independent random variables, with at most one Gaussian,* $\mathbf{y} = \mathbf{B}\mathbf{x}$ *has mutually independent components iff* $\mathbf{BA} = \mathbf{PD}$*, where* $\mathbf{P}$ *and* $\mathbf{D}$ *are permutation and diagonal matrices, respectively.*

This theorem is only based on the independence of random variables. The independence criterion involves (explicitly or implicitly) higher order (than 2) statistics, but does not take into account the order of samples. It means that the iid assumption is not required, it is just a default assumption: consequently, it works for iid as well as for not iid sources, but it cannot work for Gaussian sources.

The different (blind) ICA algorithms then use different ideas for achieving the required higher order independences. Some of different ideas used for ICA are:

- Non-linear decorrelation [8, 9, 10, 11].
- Methods based on Higher (than 2) Order Statistics (HOS)[12, 13]).
- Cancellation the mutual information (MI) of the outputs [2, 14, 15, 16, 17]. This approach may be shown to provide asymptotically a Maximum Likelihood (ML) estimation of the source signals [18].
- Algorithms based on non-Gaussianity [19, 20]. These algorithms may be shown to have a close correspondence to the algorithms based on MI minimization (refer to section 10.2 of [21]).

More complicated mixing systems have also been studied in the literature. For example, in (linear) convolutive mixtures, the mixing model is $\mathbf{x}(n) = \mathbf{B}_0\mathbf{x}(n)+\mathbf{B}_1\mathbf{x}(n-1)+\cdots+\mathbf{B}_p\mathbf{x}(n-p) = [\mathbf{B}(z)]\mathbf{x}(n)$, which has been shown [22] to be separable. Non-linear mixtures are not in general separable (refer to chapter 3 of [23]). A practically important case of non-linear mixtures is Post Non-Linear (PNL) mixtures [24, 23, 25], in which a linear mixture is followed by non-linear sensors. It has been shown that PNL mixtures are separable, too [24, 23].

However, if some weak prior information about the source signals is available, then the performance of the source separation algorithms may be significantly improved. Thus, these methods are not 'Blind' but 'Semi-Blind'. In the rest of this paper, some of most frequently used priors have been considered. It should be noted, however, that the 'sparsity prior' and 'Bayesian approaches' are mostly considered in the two companions of this paper [4, 3].

3 Separation of non iid sources

Suppose that we know a priori that the source samples are not iid, *i.e.* if sources are temporally correlated, or non stationary.

3.1 Separation of correlated sources

Several approaches had been proposed for separating correlated sources [26, 27, 28]. Pham and Garat [29] showed that time-correlated Gaussian sources can be separated provided than their spectra are different. In that case, the separation can be achieve by estimating a separation matrix $\mathbf{B}$ which minimizes the criterion

$$C(\mathbf{B}) = \sum_{l=1}^{L} w_l \text{off}(\mathbf{B}\hat{\mathbf{R}}(\tau_l)\mathbf{B}^T), \tag{3.1}$$

where w_l are weighting coefficients, off(.) is a measure of deviation from diagonality, which is positive and vanishes iff (.) is diagonal and which satisfies:

$$\text{off}(\mathbf{R}) = \text{D}(\mathbf{R} \mid \text{diag}\mathbf{R}), \tag{3.2}$$

where $\text{D}(\mathbf{R}_i \mid \mathbf{R}_j)$ denotes the Kullback-Leibler divergence of two zero mean multivariate normal densities, with variance-covariance matrices $\mathbf{R}_i$ and $\mathbf{R}_j$, and diag$\mathbf{R}$ is the diagonal matrice composed by diagonal entries of $\mathbf{R}$ and zeros elsewhere.

The criterion (3.1) involves a set of variance-covariance matrices with various delays τ_l : $\hat{\mathbf{R}}(\tau_l) = \hat{E}[\mathbf{y}(t-\tau_l)\mathbf{y}(t)^T]$, where $\hat{E}[.]$ is estimated using an empirical mean. Basically, minimizing this criterion is equivalent to estimate the separation matrix $\mathbf{B}$ which diagonalizes jointly the set of the variance-covariance matrices. The advantages of this approache are:

- it only requires second-order statistics,
- it can then separate Gaussian sources,
- there exist many very fast and efficient algorithms for jointly diagonalizing matrices [30, 31].

3.2 Separation of nonstationary sources

Source nonstationarity have been first used by Matsuoka *et al.* [32]. More recently, Pham et Cardoso developed a rigourous formalization, and proved that nonstationary Gaussian sources can be separated provided than the variance ratios $\sigma_i^2(t)/\sigma_j^2(t)$ are not constant. In that case, the separation can be achieve by estimating a separation matrix $\mathbf{B}$ which minimizes the criterion

$$C(\mathbf{B}) = \sum_{l=1}^{L} w_l \text{off}(\mathbf{B}\hat{\mathbf{R}}_l\mathbf{B}^T), \tag{3.3}$$

where we use the same notations than in the previous subsection. In Eq. (3.3), matrices $\hat{\mathbf{R}}_l$ are variance-covariance matrices estimated by empirical mean on successive sample blocks T_l. Among a few algorithms, the separation matrix $\mathbf{B}$ can be computed as the matrix which jointly diagonalizes the set of the variance-covariance matrices $\mathbf{R}_l$.

The method has the same advantages than the method exploiting the temporal correlation.

ESANN'2006 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2006, d-side publi., ISBN 2-930307-06-4.

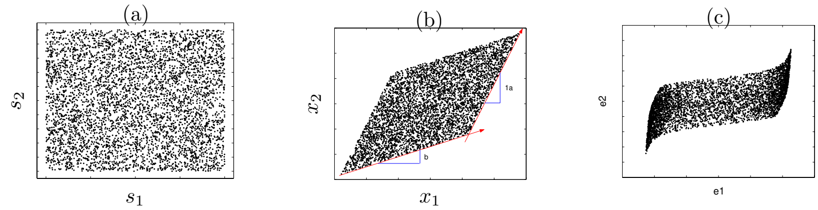

Fig. 1: Distribution of a) source samples, b) observation samples in linear mixtures, observation samples in Post Non-Linear (PNL) mixtures.

Moreover, it can be easily extended to linear convolutive mixtures, considered in the frequency domain. In fact, after Fourier transform, in each frequency band the signal tends to be close to a Gaussian signal, and consequently the method based on non Gaussian iid model are not efficient. Conversely, if the source is non stationary, one can extend the above algorithm in the frequency domain. This idea provides a very efficient method for speech signal [33].

4 Geometrical methods for source separation

4.1 Bounded sources

Suppose we know a priori that the sources are all bounded. This simple prior leads to simple geometrical interpretations and methods for source separation (firstly introduced in [34]).

Consider, for example, separating two sources from two mixtures. Because of the scale indeterminacy, the mixing matrix may be assumed to be of the form:

$$\mathbf{A} = \begin{pmatrix} 1 & a \\ b & 1 \end{pmatrix} \tag{4.1}$$

where a and b are constants to be estimated from the observed signals. Since the sources are bounded, the Probability Density Function (PDF) of each source has a bounded support, *i.e.* $p_i(s_i)$ (the PDF of the ith source) is non-zero only inside an interval $\alpha_i < s_i < \beta_i$. Then, the joint PDF $p_{\mathbf{s}}(\mathbf{s}) = p_1(s_1)p_2(s_2)$ is non-zero only in the rectangular region $\{(s_1, s_2) \mid \alpha_1 < s_1 < \beta_1, \alpha_2 < s_2 < \beta_2\}$. Consequently, if we have 'enough samples' $(s_1(n), s_2(n))$ from the sources, they form a rectangular region in the s-plane (see Fig. 1.a). This rectangle will be transformed, by the linear transformation $\mathbf{x} = \mathbf{A}\mathbf{s}$, into a parallelogram and the slopes of the borders of this parallelogram determine a and b (Fig. 1.b).

The above idea may be even generalized for separating PNL mixtures [35]: in a PNL mixture, the parallelogram of Fig. 1.b is again transformed, by 'component-wise' nonlinearities (corresponding to sensor nonlinearities), into a nonlinear region (Fig. 1.c). It has been proved [35] that if this non-linear region is transformed again into a parallelogram by 'component-wise' nonlinearities, the sensor nonlinearities have been completely compensated. An iterative algorithm is then proposed in [35] for estimating the borders and inverting them.

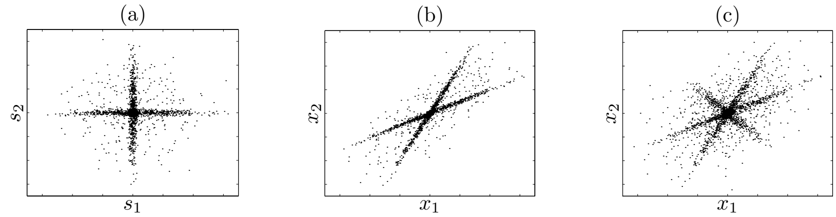

Fig. 2: Distribution of a) two sparse sources, (b) mixture of two sparse sources, (c) mixture of three sparse sources.

4.2 Sparse sources

Geometrical ideas are specially useful for separating sparse sources ('sparsity of sources' is itself another prior which lets us to significantly improve the results [3]). Note that for sparse sources, the probability that a sample $(s_1(n), s_2(n))$ is observed at the borders of the rectangular region of Fig. 1.a is very low, and hence we cannot rely on estimating the borders of the parallelogram of Fig. 1.b for source separation. However, for these sources, two 'axes' (parallel to the borders of the parallelogram) are easily visible, and their slopes again determine the mixing matrix (see Fig. 2.a and2.b, obtained from synthetic sparse signals). Moreover, for sparse sources, two new important advantages may be obtained:

1. Contrary to the traditional geometrical algorithm, it is easy to generalize the above geometric idea to higher dimensions (separating N sources from N mixtures) [36].

2. It enables us to estimate the mixing matrix (and even recovering the sources) in the underdetermined case, that is, where there is less sensors than sources [37]. Consider, for example, Fig. 2.c, for the case of 3 sources and 2 sensors. Three 'axes' are visible in this scatter plot, and they correspond to the 3 columns of the mixing matrix. This is because $\mathbf{x} = s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + s_3\mathbf{a}_3$, where $\mathbf{a}_i$'s are the columns of the mixing matrix, and consequently the axes of Fig. 2.c (which correspond to the instances where 2 of 3 sources are nearly zero) are in the directions of $\mathbf{a}_i$'s. This idea can be directly generalized to more number of sources and sensors.

A main restriction of the above idea for identifying the mixing matrix in underdetermined case, is that it is implicitly assumed that most of times there is just one 'active' (*i.e.* high-energy) source. The expected number of active sources at each instant is Np, where N is the number of sources, and p is the probability of a source being active (small by sparsity assumption). When Np is large (*e.g.* because of a very large N) the above idea fails. A solution to this problem has been proposed in [38].

Moreover, from the above geometric ideas, it is visually seen that for separating sparse sources, the independence of source signals is of minor importance. In fact, even this assumption may be dropped, leading to the name Sparce Component Analysis (SCA).

4.3 Discrete-valued sources

Another prior used in some papers [39, 40, 41, 42] is to assume that the sources are discrete (*e.g.* binary or n-valued), and the observations are continuous mixtures of them. Since the discrete sources are also bounded, the methods for separating bounded sources may be used for separating these mixtures, too. However, they can be modified to gain more advantages (*e.g.* simplicity, accuracy, or considering noisy mixtures). Moreover, for underdetermined mixtures of discrete sources, it is possible to identify and even recover the sources (much easier than the case of sparse sources). This can be seen by having in mind a geometrical interpretation like the previous section. Furthermore, for discrete sources, even the independence assumption may be dropped [40].

In [40], a geometrical approach (similar to what is presented in the previous section) is presented for separating discrete (n-valued) sources, in which the independence of the sources is not required. A Maximum Likelihood method for separating these mixtures (which works for underdetermined mixtures, too) has been proposed in [39], in which, it is assumed that the source distribution, too, is known a priori. The case of binary valued sources has been considered in [42] and a method based on creating virtual observations has been proposed. The same authors have proposed a solution based on a polynomial criterion [43] for PSK communication sources. In [41] the underdetermined BSS problem has been considered in a general case, and then a solution has been proposed for the case of discrete sources. An extension to the case of Post Non-Linear mixtures, where the source alphabet (except its size) is not known a priori, is considered in [44].

In the two previous subsections, sparsity is evidently a source property. More generally, and it will be explained in details in [3], a pre-processing step can transform (linearly for preserving the mixing model) the initial sources in new sparse (or sparser) sources in the transformed domain[1], so that methods exploiting sparsity can be used.

5 Others priors

Outside the general framework of Bayesian approaches, many other priors can be used for improving source separation methods. In this section, we show how visual information can enhanced speech separation. We also consider positivity constraints.

In the two next subsections, the speech (linear instantaneous or convolutive) mixtures, $\mathbf{x}(t)$ are completed by the video recording of the speaker face, $V(t')$ with a 20ms sampling period. Moreover, we consider extraction of one speech signal[2], the one associated to the visual cue.

[1] *e.g.* wavelet or time-frequency domains
[2] instead separation of all the sources as usual in source separation

5.1 Extraction based on audio-video spectrum estimation

The basic idea is to use the simple visual information, $V(t') = [h(t'), w(t')]^T$ associated to the height, $h(t')$, and width, $w(t')$, lip opening, for estimating a rough estimation of the speech spectrum of the speaker. Since lip motions are related to sounds but present ambiguities, from a set of audio-visual data, we first estimated (learning) a probabilistic audio-visual model. Then, by maximizing the audio-video likelihood by the EM algorithm, we can extract the audio source associated to the video. This method have been compared to Jade [30] and is much more efficient. It has mainly two advantages:

- it is very efficient for low SNR,
- it select the source of interest among all the sources.

The method can be extended for convolutive mixtures, in the frequency domain. In that case, a similar approach is done in each frequency band. Moreover, the video information is also very efficient for cancelling the permutation indeterminacies. [45, 46]

5.2 Extraction based on voice visual activity (VVA) detection

Another idea is to use the video signal for detecting the voice activity. As a simple idea, we claim that, on the frame t', there is voice activity if the lip motion is greater than a threshold, *i.e.* if:

$$\text{vva}(t') = \left|\frac{\partial h(t')}{\partial t'}\right| + \left|\frac{\partial w(t')}{\partial t'}\right|. \tag{5.1}$$

For avoiding noisy estimations, the actual VVA is decided after smoothing on the T previous frames:

$$\text{VVA}(t') = \sum_{k=0}^{T} a_k \text{vva}(t' - k), \tag{5.2}$$

where a_k are the coefficients of a truncated first-order IIR low-pass filter. This visual voice activity detector is very efficient for cancelling permutation indeterminacies in frequency domain source separation algorithms for convolutive mixtures [47].

5.3 Positive and non independent sources

In many problems, observations are positive mixtures of positive data. It is for instance the case of nuclear magnetic resonance spectroscopy of chemical compounds [48], or of hyperspectral images [49]. Moreover, in these cases, the spectra of the different species are basically non independent. Consequently, using ICA for recovering the spectra generally fails, or provides spectra with spurious peaks. Taking into account the positivity of the mixture matrix entries, improves the solution, but is generally not sufficient (due to the spectrum dependence, ICA can fail). Currently, in such cases, Bayesian methods, able to manage all the priors, are the most efficient [50].

6 Conclusion and introduction to the special session

Now, it must be clear that *blind* source separation does not exist. First, although this point has not been adressed in this paper, it is important to have priors on the mixture models, and to consider a suitable separation model. Second, prior on sources are essential. From a statistical point of view, since the problem has no solution for Gaussian iid signals, 3 types of priors are possible : sources are non Gaussian iid, sources are Gaussian temporally correlated, sources are Gaussian nonstationary. Remember that, in the 2 former cases, Gaussian means that second order statistics is sufficient, and that it is then possible to consider Gaussian sources, but the methods works for non Gaussian sources too.

Additionaly, other priors can provide original, simple and efficient algorithms. For instance, bounded sources or discrete sources leads to geometrical algorithms. It is also possible to exploit other informations like positivity of sources and/or mixing entries, or to add a visual cue to enhance speech processing.

Two very interesting approaches, which provide a general framework, are the Bayesian ICA which is able to take into account any priors, and the Sparse Component Analysis, which both exploits the data sparsity and looks for sparse representations. These two approaches are explained in details in the survey papers of A. Mohammad-Djafari [4] and Gribonval and Lesage [3].

The rest of the special session on Semi-Blind Source Separation is a *melting pot* of works illustrating the various priors that can be used: correlated sources [51], bounded sources [52, 53], sparse sources [38, 54], Bayesian source separation [55], prior on the source power spectral density [56], non-independent sources which leads to independent subspaces instead to independent component [57, 58].

References

[1] C. Jutten and J. Hérault. Independent components analysis versus principal components analysis. In *Proc. European Signal Processing Conf. EUSIPCO 88*, pages 643–646, Grenoble, France, September 1988.

[2] P Comon. Independent component analysis, a new concept ? *Signal Processing*, 36(3):287–314, 1994.

[3] R. Gribonval and S. Lesage. A survey of sparse components analysis for blind source separation: principles, perspectives and new challenges. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[4] A. Mohamed-Djafari. Bayesian source separation: beyong PCA and ICA. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[5] G. Darmois. Analyse générale des liaisons stochastiques. *Rev. Inst. Intern. Stat.*, 21:2–8, 1953.

[6] J.-F. Cardoso. The three easy routes to independent component analysis: contrasts and geometry. In *Proceeding of ICA 2001*, San Diego, USA, 2001.

[7] C. Jutten and J. Hérault. Blind separation of sources, Part I: an adaptive algorithm based on a neuromimetic architecture. *Signal Processing*, 24(1):1–10, 1991.

[8] J. Hérault and C. Jutten. Space or time adaptive signal processing by neural networks models. In *Intern. Conf. on Neural Networks for Computing*, pages 206–211, Snowbird (Utah, USA), 1986.

[9] A. Cichocki, R. Unbehauen, and E. Rummert. Robust learning algorithm for blind separation of signals. *Electronics Letters*, 30(17):1386–1387, 1994.

[10] J.-F. Cardoso and B. Laheld. An information-maximization approach to blind separation and blind deconvolution. *IEEE Trans. on SP*, 44:3017–3030, 1996.

[11] S. I. Amari, A. Cichocki, and Yang H. H. A new learning algorithm for blind source separation. *Advances in neural information processing systems*, 8:757–763, 1996.

[12] J.-L. Lacoume and P. Ruiz. Sources identification: a solution based on cumulants. In *IEEE ASSP Workshop*, Mineapolis, USA, August 1988.

[13] P. Comon. Separation of sources using higher-order cumulants. In *SPIE Vol. 1152 Advanced Algorithms and Architectures for Signal Processing IV*, San Diego (CA), USA, August 8-10 1989.

[14] D. T. Pham. Mutual information approach to blind separation of stationary sources. In *Proceedings of ICA'99*, pages 215–220, Aussois, France, January 1999.

[15] G. A. Darbellay and P. Tichavský. Independent component analysis through direct estimation of the mutual information. In *ICA2000*, pages 69–74, Helsinki, Finland, June 2000.

[16] L. B. Almeida. MISEP – linear and nonlinear ica based on mutual information. *Journal of Machine Learning Research*, 4:1297–1318, December 2003.

[17] M. Babaie-Zadeh and C. Jutten. A general approach for mutual information minimization and its application to blind source separation. *Signal Processing*, 85(5):975–995, May 2005.

[18] J.-F. Cardoso. Blind signal separation: statistical principles. *Proceedings IEEE*, 9:2009–2025, 1998.

[19] A. Hyvärinen and E. Oja. A fast fixed point algorithm for independent component analysis. *Neural computation*, 9:1483–1492, 1997.

[20] N. Delfosse and Ph. Loubaton. Adaptive blind separation of independent sources: A deflation approach. *Signal Processing*, 45:59–83, 1995.

[21] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons, 2001.

[22] D Yellin and E. Weinstein. Criteria for multichannel signal separation. *IEEE Trans. Signal Processing*, pages 2158–2168, August 1994.

[23] M. Babaie-Zadeh. *On blind source separation in convolutive and nonlinear mixtures*. PhD thesis, INP Grenoble, 2002.

[24] A. Taleb and C. Jutten. Source separation in post nonlinear mixtures. *IEEE Tr. on SP*, 47(10):2807–2820, 1999.

[25] S. Achard and C. Jutten. Identifiability of post nonlinear mixtures. *IEEE Signal Processing Letters*, 12(5):423–426, May 2005.

[26] L. Tong, V. Soon, R. Liu, and Y. Huang. Amuse: a new blind identification algorithm. In *Proc. ISCAS*, New Orleans, USA, 1990.

[27] A. Belouchrani and J.-F. Cardoso. Maximum likelihood source separation by the expectation-maximization technique. In *NOLTA 95*, Las Vegas, USA, 1995.

[28] L. Molgedey and H.G. Schuster. Separation of a mixture of independent signals using time delayed correlations. *Physical Review Letters*, 72(23):3634–3637, June 1994.

[29] D. T. Pham and Ph. Garat. Blind separation of mixture of independent sources through a quasimaximum likelihood approach. *IEEE Trans. on SP*, 45:1712–1725, 1997.

[30] J.-F. Cardoso and A. Souloumiac. An efficient technique for blind separation of complex sources. In *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics*, pages 275–279, South Lac Tahoe, USA (CA), June 1993.

[31] D.T. Pham. Blind separation of instantaneous mixtures via an independent component analysis. *SIAM J. on Matrix Anal. and Appl.*, 22(4):1136–1152, 2001.

[32] K. Matsuoka, M. Ohya, and M. Kawamoto. A neural net for blind separation of nonstationary signals. *Neural Networks*, 8(3):411–419, 1995.

[33] D.T. Pham, C. Servière, and H. Boumaraf. Blind separation of convolutive audio mixtures using nonstationarity. In *Proc. Int. Conf. ICA 32003*, Nara, Japan, April 2003.

[34] C. Puntonet, A. Mansour, and C. Jutten. A geometrical algorithm for blind separation of sources. In *Actes du XV Colloque GRETSI 95*, Juan-Les-Pins, France, Septembre 1995.

[35] M. Babaie-Zadeh, C. Jutten, and K. Nayebi. A geometric approach for separating Post Non-Linear mixtures. In *EUSIPCO*, volume II, pages 11–14, Toulouse, France, September 2002.

[36] M. Babaie-Zadeh, C. Jutten, and A. Mansour. Sparse ica via cluster-wise pca. *Neurocomputing*, 2006. (to be appeared).

[37] P. Bofill and M. Zibulevsky. Source separation using sparse representations. *Signal Processing*, 81:2353–2362, 2001. (to be appeared).

[38] Z. He and A. Cichocki. K-subspace clustering and its application in sparse component analysis. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[39] A. Belouchrani and J.-F. Cardoso. Maximum likelihood source separation for discrete sources. In *EUSIPCO*, pages 768–771, Edinburgh, Scotland, September 1994.

[40] C. Puntonet, A. Prieto, C. Jutten, M. Rodriguez-Alvarez, and J. Ortega. Separation of sources: A geometry-based procedure for reconstruction of n-valued signals. *Signal Processing*, 46(3):267–284, 1995.

[41] A. Taleb and C. Jutten. On underdetermined source separation. In *ICASSP'99*, Arizona, March 1999.

[42] P. Comon and O. Grellier. Non-linear inversion of under-determined mixtures. In *Proceedings of International Workshop on Independent Component Analysis and Signal Separation (ICA'99)*, pages 461–465, Aussois, France, January 1999.

[43] O. Grellier and P. Comon. Blind separation of discrete sources. *IEEE Signal Processing Letters*, 5(8):212–214, August 1998.

[44] B. Lachover and A. Yeredor. Separation of polynomial post non-linear mixtures of discrete sources. In *Proceedings of The 2005 IEEE Workshop on Statistical Signal Processing (SSP2005)*, Bordeaux, France, July 2005.

[45] W. Wang, D. Cosker, Y. Hicks, S. Sanei, and J. Chambers. Video assisted speech source separation. In *Proceedings of the International Conference of Acoustic, Speech and Signal Processing (ICASSP2005)*, Philadelphia (USA), March 2005.

[46] B. Rivet, L. Girin, C. Jutten, and J.-L. Schwartz. Solving the indeterminations of blind source separation of convolutive speech mixtures. In *Proceedings of the International Conference of Acoustic, Speech and Signal Processing (ICASSP2005)*, Philadelphia (USA), March 2005.

[47] B. Rivet, C. Servière, L. Girin, D.T. Pham, and C. Jutten. Using a visual voice activity detector to cance the permutation indeterminacy in blind source separation of speech mixtures. In *Proceedings of EUSIPCO 06*, Florence, Italy, September 2006. submitted.

[48] S. Moussaoui, D. Brie, and C. Carteret. Non-negative source separation using the maximum likelihood approach. In *Proceedings of the International Workshop of Statistical Signal Processing (SSP2005)*, Bordeaux, France, July 2005.

[49] M. Lennon, G. Mercier, M. Mouchot, and L. Hubert-Moy. Spectral unmixing of hyperspectral images with the independent component analysis and wavelet packets. In *Int. Geoscience Remote Sensing Symposium*, Sydney, Australia, July 2001.

[50] S. Moussaoui, D. Brie, C. Carteret, and A. Mohammad-Djafari. Application of bayesian non-negative source separation to mixture analysis in spectroscopy. In *Proceedings of the 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt'2004)*, Garching, Germany, July 2004.

[51] Y. Deville, D. Bissessur, M. Puigt, Sh. Hosseini, and H. Carfantan. A time-scale correlation-based blind separation method applicable to correlated sources. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[52] J.A. Lee, F. Vrins, and M. Verleysen. Non-orthogonal support width ica. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[53] D.T. Pham and F. Vrins. Discriminacy of the minimum range approach to blind separation of bounded sources. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[54] M. Pedzisz and A. Mansour. A simple idea to separate convolutive mixtures in an undetermined scenario. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[55] N. Bali, A. Mohammad-Djafari, and A. Mohammadpour. Hierarchical markovian models for jointclassification, segmentation and data reduction of hyperspectral images. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[56] J. Igual, R. Llinares, and A. Camacho. Source separation with priors on the power spectrum of the sources. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[57] A. Ilin. Independent dynamic subspace analysis. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.

[58] A. Hyvärinen and U. Köster. Fastisa: A fast fixed-point algorithm for independent subspace analysis. In *Proc. ESANN 2006*, Bruges, Belgium, April 2006.