# A survey of Sparse Component Analysis for blind source separation: principles, perspectives, and new challenges

Rémi Gribonval, Sylvain Lesage

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

# A survey of Sparse Component Analysis for Blind Source Separation: principles, perspectives, and new challenges

Rémi Gribonval, Sylvain Lesage

IRISA
Campus de Beaulieu
35042 Rennes Cedex, France

**Abstract**. In this survey, we highlight the appealing features and challenges of Sparse Component Analysis (SCA) for blind source separation (BSS). SCA is a simple yet powerful framework to separate several sources from few sensors, even when the independence assumption is dropped. So far, SCA has been most successfully applied when the sources can be represented sparsely in a given basis, but many other potential uses of SCA remain unexplored. Among other challenging perspectives, we discuss how SCA could be used to exploit both the spatial diversity corresponding to the mixing process and the morphological diversity between sources to unmix even underdetermined convolutive mixtures. This raises several challenges, including the design of both provably good and numerically efficient algorithms for large-scale sparse approximation with overcomplete signal dictionaries.

## 1 Introduction

Because of its many applications in diverse areas of science, the blind source separation (BSS) problem has been extensively studied in the last twenty years. The most common and widely considered blind source separation model is the linear instantaneous one $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{e}(t)$ where $\mathbf{x}(t)$ is the column vector of observed signals $x_p(t)$, $1 \leq p \leq P$ at time $t$ (for images, $t$ would denote the pixel location), $\mathbf{s}(t)$ the column vector of unknown source signals $s_n(t)$, $1 \leq n \leq N$, $\mathbf{A}$ is an unknown $P \times N$ matrix and $\mathbf{e}(t)$ is additive noise.

For years, Independent Component Analysis (ICA) has been the favourite tool to tackle the BSS problem, since under weak assumptions of independence and non-Gaussianity, the model is identifiable up to gain and permutation indeterminacies. When the number of sources $N$ does not exceed the number of sensors $P$ and the mixing matrix $\mathbf{A}$ is well conditioned, the sources are easily recovered by simply (pseudo)inverting the mixing matrix, i.e., $\widehat{\mathbf{s}}(t) = \widehat{\mathbf{A}}^{-1}\mathbf{x}(t)$ or $\widehat{\mathbf{s}}(t) = \widehat{\mathbf{A}}^{\dagger}\mathbf{x}(t)$. However, recovering the sources for so-called degenerate or underdetermined mixtures with less sensors than mixtures seems a much harder problem in a truly blind setup relying only on the independence of the sources. It is then necessary to rely on some other priors on the sources.

Among other possible priors[1], sparsity of the sources in a transformed domain has turned out in the last few years both a successful tool [1, 2, 3, 4, 5, 6, 7] and

[1] e.g., discrete priors for artificial sources such as digital communications signals.

mathematically grounded [8, 9, 10, 11, 12] approach to address the degenerate BSS problem. In addition to making source separation possible even in the degenerate case, the sparse model also allows the estimation of the mixing matrix by clustering the scatter plot [2] of a sparse representation of the mixture. The related techniques are known as Sparse Component Analysis (SCA).

Despite its many successes, SCA faces a number of challenges. The goal of this paper is to discuss and propose a vision for its future. With this purpose in mind, we review in Section 2 the basic principles of SCA, which relies on the assumption that the sources can be represented sparsely in a given basis. Section 3 is devoted to the more general setup of SCA with overcomplete signal dictionaries. We discuss in Section 4 some new perspectives offered by extensions of this framework, such as the solution of underdetermined convolutive BSS problems. To conclude the paper, we discuss the theoretical and computational challenges raised by such perspectives.

## 2 Sparse Component Analysis

The basic principle of SCA usually consists of four steps. First, a *sparsifying linear transform* is applied to the mixture. Given a family of $K$ signals $\varphi_k(t)$ called *atoms*, the correlations $\langle x_p, \varphi_k \rangle$ are computed[2] and collected in a $P \times K$ matrix

$$C_{\mathbf{x}} := (\langle x_p, \varphi_k \rangle)_{pk} = \begin{bmatrix} c_{x_1} \\ \cdots \\ c_{x_P} \end{bmatrix} = [C_{\mathbf{x}}(1), \ldots, C_{\mathbf{x}}(K)].$$

Letting $\mathbf{\Phi}$ be the $K \times T$ *dictionary matrix*, whose rows are the atoms $\varphi_k$, the representation is obtained as[3] $C_{\mathbf{x}} = \mathbf{x}\mathbf{\Phi}^{\mathrm{H}}$ where $\cdot^{\mathrm{H}}$ denotes complex transposition. With obvious notations, it satisfies $C_{\mathbf{x}} = \mathbf{A}C_{\mathbf{s}} + C_{\mathbf{e}}$. Often used sparsifying transforms are orthogonal wavelet or wavelet packet transforms, and short time Fourier transforms (STFT). The transform should be chosen to sparsify the source representation, so that each source representation $c_{s_n}$ has few significant coefficients $c_{s_n}(k)$, which are located on a set $k \in I_n \subset \{1, \ldots, K\}$. If the sparsity is sufficient, then the sets $I_n$ can be almost disjoint for different sources, so that for most $k \in I_n$ the relation $C_{\mathbf{x}}(k) = \mathbf{A}C_{\mathbf{s}}(k) + C_{\mathbf{e}}(k) \approx c_{s_n}(k)\mathbf{a}_n$ holds, where $\mathbf{a}_n$ is the column of the mixing matrix which corresponds to the source $s_n$. In such a case, the scatter plot $\{C_{\mathbf{x}}(k)\}_{k=1}^{K}$ is made of points almost aligned with the columns of the mixing matrix, and one can cluster it to estimate $\widehat{\mathbf{A}}$.

The second step precisely consists in *estimating* $\mathbf{A}$ *from the scatter plot* of $C_{\mathbf{x}}$. Besides natural gradient ICA approaches [1], a common approach today is to rely on clustering techniques, with variants of weighted K-means. For such techniques to perform well, the key hypothesis is that *at most one source* contributes significantly to *each* point of the scatter plot. In practice, this hypothesis is at best valid for *most* points of the scatter plot, provided that the sources have

[2]The notation $x_p$ (resp. $\mathbf{x}$, $s_n$, $\mathbf{s}$) denotes the row vector $(x_p(t))_{1 \le t \le T} \in \mathbb{C}^T$ or $\mathbb{R}^T$.

[3]Note that the transform $\mathbf{\Phi}^{\mathrm{H}}$ operates on the right hand side, which is the time domain, as opposed to the mixing matrix $\mathbf{A}$ which acts on the left hand side, the spatial domain.

not only sparse but also almost disjoint representations [13]. Several authors [14, 15, 16] have proposed techniques to estimate $\mathbf{A}$ based on the far weaker hypothesis that for each source, there is *some* (possibly small) subset of the scatter plot where it is the only significantly active one, even if the transformed sources overlap in almost all the scatter plot. For the latter techniques to work, the sources representations are assumed to have some structure: in the case of audio sources, it is typically assumed that, in the time-frequency domain, the activity of each source displays some *local persistence* within the small time-frequency regions where they are "visible". Such models even make it possible to even estimate the number of sources in the mixture [16].

The third step is to *estimate the source representations* based on the sparsity assumption. In a noiseless context, Bofill and Zibulevsky [17] proposed to estimate the source representation $\widehat{C}_{\mathbf{s}}$

$$\widehat{C}_{\mathbf{s}}(k) := \arg \min_{C \mid C_{\mathbf{x}}(k)=\widehat{\mathbf{A}}C} \|C\|_q, \qquad q = 1. \tag{1}$$

which can be interpreted as a maximum likelihood estimate assuming the sources coefficients have a Laplacian distribution. More recently, some variants where the $\ell^1$ norm is replaced with an $\ell^q$ norm for $q < 1$ have been reported to improve performance for audio BSS [18]. Another approach, which is more robust when the estimation of the mixing matrix is prone to errors [19], is binary masking [7]. It sets the estimated source components to $\varepsilon_n \mathbf{a}_n^T C_{\mathbf{x}}$ where $\varepsilon_n := 1$ for the source whose direction maximizes the correlation $\left|\mathbf{a}_n^T C_{\mathbf{x}}\right|$ with the mixture component, and $\varepsilon_n := 0$ for the other sources.

The last step is the *reconstruction of the sources* by an inverse of the sparsifying transform. For unitary sparsifying transforms such as the discrete wavelet transform or the Fourier transform, this is simply obtained as $\widehat{\mathbf{s}} := \widehat{C}_{\mathbf{s}}\mathbf{\Phi}$.

## 3 SCA with overcomplete signal dictionaries

SCA based on linear transforms has led to successful BSS techniques such as DUET [7], which works very well for the separation of anechoic sterophonic audio mixtures. Since the beginning of the 1990s, the use of redundant signal dictionaries for sparse signal representations has been advocated by many researchers [20, 21] as a mean to enforce higher sparsity than could be obtained with linear transforms by exploiting the variety of waveforms in the signal dictionary to get a better adaptation to the signal features. Such approaches consist in modeling the source signals as vectors in $\mathbb{R}^T$ or $\mathbb{C}^T$ with a sparse expansion

$$s_n \approx \sum_{k=1}^{K} c_n(k)\ \varphi_k = c_n.\mathbf{\Phi} \tag{2}$$

where $\mathbf{\Phi}$ is the redundant $K \times T$ dictionary matrix, used in synthesis mode, and $c_n \in \mathbb{R}^K$ or $\mathbb{C}^K$ is a sparse row vector of synthesis coefficients. Since $\mathbf{\Phi}$ is redundant, such an expansion is not unique: among the infinitely many possible expansions for a given signal, one can hope to get a highly sparse one. In the BSS context, the sparse source model (2) leads to a sparse mixture model

$$\mathbf{x} = \mathbf{A}C_{\mathbf{s}}\mathbf{\Phi} + \mathbf{e} = \sum_{n=1}^{N} \mathbf{a}_n \cdot c_n \cdot \mathbf{\Phi} + \mathbf{e} = \sum_{n=1}^{N}\sum_{k=1}^{K} c_n(k) \cdot \mathbf{a}_n \varphi_k \tag{3}$$

where $\mathbf{A}$ is the unknown $P \times N$ mixing matrix with columns $\mathbf{a}_n$, $C_{\mathbf{s}}$ an unknown (but sparse) $N \times K$ coefficient matrix with rows $c_n$, and $\mathbf{\Phi}$ is the dictionary in which all sources are assumed to have a sparse decomposition.

Following the SCA setup, but with adaptive sparse expansions in overcomplete dictionaries instead of linear transforms, a first step is therefore to compute an adaptive sparse representation $\mathbf{x} = C_{\mathbf{x}}\mathbf{\Phi} + \mathbf{e}$ of the mixture with $C_{\mathbf{x}}$ a $P \times K$ matrix of coefficients which can be clustered to estimate $\widehat{\mathbf{A}}$. This leads to solving a *simultaneous approximation* problem of the type

$$\widehat{C}_{\mathbf{x}} := \arg\min \frac{1}{2}\|\mathbf{x} - C_{\mathbf{x}}\mathbf{\Phi}\|_F^2 + \lambda \sum_{k=1}^{K} \|C_{\mathbf{x}}(k)\|_2^q \tag{4}$$

where $\lambda > 0$ determines a compromise between the quality of approximation, measured by the Fröbenius norm[4] $\|\cdot\|_F$, and the objective of selecting few active atoms with $\|C_{\mathbf{x}}(k)\|_2 > 0$. Note the use of $\|C_{\mathbf{x}}(k)\|_2$ and not $\|C_{\mathbf{x}}(k)\|_q$ in (4), which avoids giving a preference to any spatial orientation for the selected atoms.

The following steps consist in computing a sparse source representation $\widehat{C}_{\mathbf{s}}$ compatible with (3), where $\mathbf{A}$ is replaced by the estimated matrix $\widehat{\mathbf{A}}$, and to reconstruct the sources as $\widehat{\mathbf{s}} := \widehat{C}_{\mathbf{s}}\mathbf{\Phi}$. This can be done with an $\ell^q$ minimization approach similar to (1): in a noisy setting it becomes

$$\widehat{C_{\mathbf{s}}} = \arg\min_{C} \frac{1}{2}\|\mathbf{x} - \widehat{\mathbf{A}}C\mathbf{\Phi}\|_F^2 + \lambda \sum_{k=1}^{K} \|C_{\mathbf{s}}(k)\|_q^q. \tag{5}$$

This approach was developped for $q = 1$ [3] with an interpretation as a MAP estimation of the sources under a Laplacian model and Gaussian white noise $\mathbf{e}$.

## 4 Perspectives

Today, most SCA techniques for BSS rely on the sparsity of the sources in a common basis to separate them based on their *spatial diversity*. Its principle is essentially to decompose the mixture into elementary components, which are then grouped according to the similarity between their spatial patterns, before being finally recombined to reconstruct the sources. Such a principle is amenable to more general types of decompositions [22], provided that a proper grouping criterion makes it possible to distinguish which components of the decomposition correspond to which source. Below we discuss how this has been or could be used in various frameworks.

[4] the square root of the sum of all squared entries in a matrix.

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

*Combining spatial and morphological diversity.* If the sources happen to be sparse in *different* (known) dictionaries, it is also possible to rely on sparsity to separate the sources, even in the single channel separation case where no spatial diversity is available. Such an approach has been called *morphological component analysis* [23] since it relies in the different *typical shapes* of the sources to separate them, and it was successfully exploited for the separation of piecewise smooth content and textures in images [24], or the decomposition of musical sounds into a tonal and a transient layer for coding [25]. A technique to exploit *both* spatial and morphological diversity of the sources, called *multichannel morphological component analysis* [26], was recently proposed: with a specific algorithm, it decomposes a mixture as

$$\mathbf{x} = \sum_{n=1}^{N} \mathbf{a}_n \cdot c_n \cdot \mathbf{\Phi}^{(n)} + \mathbf{e} = \sum_{n=1}^{N} \sum_{k=1}^{K} c_n(k) \cdot \mathbf{a}_n \varphi_k^{(n)} \tag{6}$$

where $\mathbf{a}_n$ is the spatial direction of the source $s_n$ and $c_n \cdot \mathbf{\Phi}^{(n)}$ is a sparse approximation thereof in a specific dictionary $\mathbf{\Phi}^{(n)}$ (e.g., in audio a wavelet basis for drums, and a Gabor dictionary for a violin and a piano). The estimated sources are reconstructed as $\widehat{s}_n := \sum_k c_n(k)\varphi_k^{(n)}$.

*Convolutive degenerate mixtures.* So far, SCA has been mostly used to deal with linear instantaneous mixtures, with notable extensions to anechoic (time-delayed) mixtures [7, 27]. More generally, one can face the problem of separating (possibly underdetermined) convolutive mixtures $\mathbf{x} = \mathbf{A} \star \mathbf{s} + \mathbf{e}$. Examining whether and how sparse source models can at all help identify the mixing conditions today remains a serious challenge. Extending SCA to separate the mixed sources, when the mixing conditions are known, seems an easier perspective which is in the spirit of $\ell^1$-based deconvolution [28]. Formally, given an estimate of the mixing matrix of filters, one can decompose a mixture as

$$\mathbf{x} = \sum_{n=1}^{N} \mathbf{a}_n \star (c_n \cdot \mathbf{\Phi}^{(n)}) + \mathbf{e} = \sum_{n=1}^{N} \sum_{k=1}^{K} c_n(k) \cdot \mathbf{a}_n \star \varphi_k^{(n)} + \mathbf{e}, \tag{7}$$

and reconstruct the sources as $\widehat{s}_n := \sum_k c_n(k)\varphi_k^{(n)}$. The success of the approach certainly depends on the global diversity of the multichannel waveforms $\mathbf{a}_n \star \varphi_k^{(n)}$, and the practical validity of this approach is yet to be demonstrated.

## 5 Challenges and open problems

One of the main bottlenecks which prevents the wide exploration of potential applications of SCA with overcomplete signal dictionaries is the computational complexity of the optimization algorithms it involves. It is therefore of the utmost importance to carefully design sparse approximation algorithms that would ideally combine two crucial properties : good "quality" of the resulting decomposition, and numerically efficient design (in terms of moderate usage of computation resources such as flops and memory for large-scale problems).

*Provably good algorithms.* In the last decade, several sparse approximation algorithms have been proposed, either for simultaneous approximation Eq. (4) or for the demixing step Eq. (5). Among them, $\ell^1$ optimisation techniques and greedy algorithms have been quite extensively studied in a huge body of recent work (see *e.g.* [9, 10, 11, 12] and the references therein) initiated by Donoho and Huo [8]. These techniques have good decomposition performance –in some precise mathematical sense– provided that the sparse source model is satisfied with good accuracy. Nonconvex $\ell^q$ optimisation principles[5] have also been analyzed for $0 \leq q \leq 1$ [12]. The performance of convex $\ell^q$ optimization for $1 < q \leq 2$, which –quite surprinsingly– remains unexplored, certainly deserves its own analysis, and simultaneous approximation algorithms also raise their fair share of questions.

*Numerically efficient algorithms.* From a computational point of view, the solution of (4) and (5) for $q = 1$ typically rely on convex programming techniques which unfortunately remain computationally very demanding, except for very small problem sizes. For $0 < q \leq 1$, iterative reweighted least squares (IRLS) techniques such as FOCUSS and M-FOCUSS [29, 30] converge in few iterations to a local minimum of the criterion, yet they need a good initialization and involve the costly inversion of large matrices (or the numerical solution of large linear systems). Hopefully, $\ell^q, q \leq 1$ optimization are not the only approaches to sparse signal decompositions. Iterative algorithms such as Matching Pursuit [20] indeed provide a good alternative with a much lower computational complexity to tackle both the simultaneous approximation problem (4) – with Multichannel Matching Pursuit [5, 22, 31]– and the source separation step itself (5) – with a variant known as Demixing Pursuit [19]. A fast $N \log N$ open implementation of Multichannel as well as Demixing Matching Pursuits (called MPTK) has recently been proposed [32] and is freely available [6]. With overcomplete dictionaries such as multiscale Gabor dictionaries, which are relevant for audio signal processing, MPTK can provide close to real time decompositions even in very large dimension. More recent algorithms such as iterative thresholding [33, 34, 35] are very promising in terms of numerical efficiency, though not yet defined in a simultaneous approximation setting.

*Learning the mixing matrix ... and the dictionary.* Last, but not least in a series of open questions about SCA for BSS comes the issue of analyzing the identifiability of the mixing matrix, and the robustness of the separation algorithms when $\mathbf{A}$ is unprecisely known. Related problems involve the analysis of the role of the choice of the dictionary $\mathbf{\Phi}$ in the separation performance, in close connection with learning and sparse coding issues, where the dictionary itself might be learned from data [1].

[5]For $0 < q \leq 1$, the $\ell^q$ quasi-norm $\|c\|_q := (\sum_k |c(k)|^q)^{1/q}$ satisfies the quasi-triangle inequality $\|c+d\|_q^q \leq \|c\|_q^q + \|d\|_q^q$ instead of the triangle inequality. For $q = 0$ the $\ell^0$ quasi-norm $\|c\|_0$ simply counts the number of nonzero entries in a vector.

[6]http://mptk.gforge.inria.fr

Overall, SCA has proved a very succesful tool for degenerate source separation, and a key to its most visible successes (such as the DUET algorithm [7]) is the combination of simplicity and efficiency of the model and corresponding algorithms. Whether or not SCA can fulfill its new promises will depend on on whether we overcome the computational challenges it raises, and explore the new lands it uncovers.

## References


[1] M.S. Lewicki and T.J. Sejnowski. Learning overcomplete representations. *Neural Computation*, 12(2):337–365, 2000.

[2] M. Van Hulle. Clustering approach to square and non-square blind source separation. In *IEEE Workshop on Neural Networks for Signal Processing (NNSP99)*, pages 315–323, August 1999.

[3] M. Zibulevsky and B.A. Pearlmutter. Blind source separation by sparse decomposition in a signal dictionary. *Neural Computation*, 13(4):863–882, 2001.

[4] L. Vielva, D. Erdogmus, and J.C. Principe. Underdetermined blind source separation using a probabilistic source sparsity model. In *Proc. Int. Conf. on ICA and BSS (ICA2001)*, pages 675–679, San Diego, California, December 2001.

[5] R. Gribonval. Sparse decomposition of stereo signals with matching pursuit and application to blind separation of more than two sources from a stereo mixture. In *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP'02)*, Orlando, Florida, May 2002.

[6] F.J. Theis, A. Jung, C.G. Puntonet, and Lang E.W. Linear geometric ICA: Fundamentals and algorithms. *Neural Computation*, 15(2):419–439, February 2003.

[7] O. Yilmaz and S. Rickard. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52(7):1830–1847, July 2004.

[8] D.L. Donoho and Xiaoming Huo. Uncertainty principles and ideal atomic decompositions. *IEEE Trans. Inform. Theory*, 47(7):2845–2862, November 2001.

[9] J. Tropp. Just relax: Convex programming methods for subset selection and sparse approximation. Technical Report ICES Report 04-04, UT-Austin, February 2004.

[10] R. Gribonval and M. Nielsen. Beyond sparsity : recovering structured representations by $\ell^1$-minimization and greedy algorithms. – Application to the analysis of sparse underdetermined ICA –. *Advances in Computational Mathematics*, 2006. accepted.

[11] D. Donoho, M. Elad, and V. Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Trans. Inf. Th.*, 52(1):6–18, January 2006.

[12] R. Gribonval, R. M. Figueras i Ventura, and P. Vandergheynst. A simple test to check the optimality of sparse signal approximations. *EURASIP Signal Processing, special issue on Sparse Approximations in Signal and Image Processing*, 86(3):496–510, March 2006.

[13] R. Balan and J. Rosca. Statistical properties of STFT ratios for two channel systems and applications to blind source separation. In *Proc. Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2000)*, pages 429–434, Helsinki, Finland, June 2000.

[14] F. Abrard and Y. Deville. A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources. *Signal Processing*, 85(7):1389–1403, July 2005.

[15] C. Févotte and C. Doncarli. Two contributions to blind source separation using time-frequency distributions. *IEEE Signal Processing Letters*, 11(3):386–389, March 2004.

[16] S. Arberet, R. Gribonval, and F. Bimbot. A robust method to count and locate audio sources in a stereophonic linear instantaneous mixture. In *Proc. of the Int'l. Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2006)*, Charleston, South Carolina, USA, March 2006.

[17] P. Bofill and M. Zibulevsky. Underdetermined blind source separation using sparse representations. *Signal Processing*, 81:2353–2362, 2001.

[18] R. Saab, Ö. Yilmaz, M.J. McKeown, and R. Abugharbieh. Underdetermined sparse blind source separation with delays. In *Proc. First Workshop on Signal Processing with Sparse/Structured Representations (SPARS'05)*, pages 67–70, Rennes, France, November 2005. see `http://spars05.irisa.fr`.

[19] S. Lesage, S. Krstulovic, and R. Gribonval. Under-determined source separation: comparison of two approaches based on sparse decompositions. In *Proc. of the Int'l. Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2006)*, Charleston, South Carolina, USA, March 2006.

[20] S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12):3397–3415, December 1993.

[21] S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, January 1999.

[22] R. Gribonval. Piecewise linear source separation. In M.A. Unser, A. Aldroubi, and A.F. Laine, editors, *Proc. SPIE '03*, volume 5207 Wavelets: Applications in Signal and Image Processing X, pages 297–310, San Diego, CA, August 2003.

[23] J.-L. Starck, Y. Moudden, J. Bobin, M. Elad, and D.L. Donoho. Morphological component analysis. In *Proceedings of the SPIE conference wavelets*, volume 5914, July 2005.

[24] J.-L. Starck, M. Elad, and D.L. Donoho. Image decomposition : Separation of textures from piecewise smooth content. In M.A. Unser, A. Aldroubi, and A.F. Laine, editors, *Wavelet: Applications in Signal and Image Processing X, Proc, SPIE '03*, volume 5207, San Diego, CA, aug 2003.

[25] Daudet L. and B. Torrésani. Hybrid representations for audiophonic signal encoding. *Signal Processing, special issue on Coding Beyond Standards*, 82(11):1595–1617, 2002.

[26] J. Bobin, Y. Moudden, J-L. Starck, and M. Elad. Multichannel morphological component analysis. In *Proc. First Workshop on Signal Processing with Sparse/Structured Representations (SPARS'05)*, Rennes, France, November 2005. see `http://spars05.irisa.fr`.

[27] M. Puigt and Y. Deville. Time-frequency ratio-based blind separation methods for attenuated and time-delayed sources. *Mechanical Systems and Signal Processing*, 19:1348–1379, 2005.

[28] F. Santosa and W.W. Symes. Linear inversion of band-limited reflection sismograms. *SIAM J. Sci. Statistic. Comput.*, 7:1307–1330, 1986.

[29] I. F. Gorodnitsky and B. D. Rao. Sparse signal reconstruction from limited data using focuss : a reweighted norm minimization algorithm. *IEEE Trans. Signal Proc.*, 45(3):600–616, mar 1997.

[30] S.F. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Trans. Signal Proc.*, 53(7):2477–2488, July 2005.

[31] D. Leviatan and V.N. Temlyakov. Simultaneous approximation by greedy algorithms. Technical Report 0302, IMI, Dept of Mathematics, University of South Carolina, Columbia, SC 29208, 2003.

[32] S. Krstulovic and R. Gribonval. MPTK: Matching Pursuit made tractable. In *Int. Conf. Acoust. Speech Signal Process. (ICASSP'06)*, Toulouse, France, 2006. to appear.

[33] M.A. Figueiredo and R.D. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Trans. Image Process.*, 12(8):906–916, August 2003.

[34] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, August 2004.

[35] M. Elad. Shrinkage for redundant representations. In *Proc. SPARS'05*, IRISA, Rennes, France, November 2005. http://spars05.irisa.fr.