Blind Image Separation using Sparse Representation

Wided Souidène, Abdeldjalil Aissa El Bey, Karim Abed-Meraim, Azzedine Beghdadi

BLIND IMAGE SEPARATION USING SPARSE REPRESENTATION

W. Souidène[1,2], A. Aïssa-El-Bey[2]

[1]Université Paris 13, L2TI
99, Av. J.B. Clément, 93430, Villetaneuse.
souidene, elbey@tsi.enst.fr

K. Abed-Meraim[2], A. Beghdadi[1]

[2]ENST-Paris, TSI department,
46, rue Barrault, 75634, Paris Cedex 13, France.
abed@tsi.enst.fr, beghdadi@l2ti.univ-paris13.fr

ABSTRACT

This paper focuses on the blind image separation using their sparse representation in an appropriate transform domain. A new separation method is proposed that proceeds in two steps: (i) an image pre-treatment step to transform the original sources into sparse images and to reduce the mixture matrix to an orthogonal transform (ii) and a separation step that exploits the transformed image sparsity via an ℓ_p-norm based contrast function. A simple and efficient natural gradient technique is used for the optimization of the contrast function. The resulting algorithm is shown to outperform existing techniques in terms of separation quality and computational cost.

Index Terms— Separation, image restoration, sparse matrices.

1. INTRODUCTION

Blind source separation (BSS) is an important research field in signal and image processing. In particular, separating linear mixtures of several 'independent' images has application in biomedical imaging [1–3], in cosmology and multispectral imaging [4,5], in polarimetric imaging [6], etc.

Recently, an important research activity has been observed for solving the BSS problem using a sparse representation of the source signals. Solution for the blind separation of image sources using sparsity include the wavelet-transform domain method in [6] and the method in [7] using projection onto sparse dictionaries.

In this work, we propose a new solution based on the transformed image sparsity. The new BSS algorithm is shown to be more efficient that other existing techniques in the literature and leads to improved separation quality with lower computational cost.

2. NOTATIONS AND DATA MODEL

We assume that N images $f_1, \cdots, f_N$ each of size (m_f, n_f) are merged and M linear mixtures of these original images are observed. The latter mixtures can be modeled by the following linear system:

$$\boldsymbol{g}(m,n) = \boldsymbol{A}\boldsymbol{f}(m,n) + \boldsymbol{w}(m,n) \qquad (1)$$

where, $\boldsymbol{f}(m,n) = [f_1(m,n), \cdots, f_N(m,n)]^T$ is a $N \times 1$ image source vector consisting of the stack of corresponding pixels of source images, $\boldsymbol{w}(m,n) = [w_1(m,n), \cdots, w_M(m,n)]^T$ is the $M \times 1$ gaussian complex noise vector which affects each image mixture pixel, $\boldsymbol{A}$ is the $M \times N$ full column rank mixing matrix (i.e., $M \geq N$), $\boldsymbol{g}(m,n) = [g_1(m,n), \cdots, g_M(m,n)]^T$ is an $M \times 1$ vector of mixture image pixels and the superscript T denotes the transpose operator. The purpose of blind image separation is to find a separating matrix, i.e. a $N \times M$ matrix $\boldsymbol{B}$ such that

$\widehat{\boldsymbol{f}}(m,n) = \boldsymbol{B}\boldsymbol{g}(m,n)$ is an estimate of original images. In practice, the separating matrix estimation is performed up to a permutation and a certain fixed scalar, i.e. $\boldsymbol{B}$ is a separating matrix iff:

$$\boldsymbol{B}\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda} \qquad (2)$$

where $\boldsymbol{P}$ is a permutation matrix and $\boldsymbol{\Lambda}$ a non-singular diagonal matrix that represent the inherent ambiguities of the BSS problem.

3. SEPARATION ALGORITHM

As shown in [6,7], exploiting the sparsity of some representations of the original images afford us to achieve the BSS problem. Indeed, the mixture destroys or 'reduces' the sparsity of the considered signals that is restored after source separation. Reversely, it is shown in [6,7] that restoring (maximizing) the sparsity leads to the desired source separation. Based on this, we propose in the sequel a two-step BSS solution consisting in a linear pre-treatment that transforms the original sources into sparse signals followed by a BSS algorithm that minimizes the ℓ_p norm of the transformed image mixtures using natural gradient technique.

3.1. Image pre-treatment

The algorithm proposed in this article is efficient for separating sparse sources. For some signals, one can assume that the spatial or temporal representation is naturally sparse, whereas for natural scenes, this assumptions falls down. We propose to make the image sparse by simply taking into account its Laplacian transform:

$$\boldsymbol{\mathcal{F}} = \nabla \boldsymbol{f} = \frac{\partial^2 \boldsymbol{f}}{\partial x^2} + \frac{\partial^2 \boldsymbol{f}}{\partial y^2} \ , \qquad (3)$$

or, in discrete form

$$\begin{aligned} \boldsymbol{\mathcal{F}}(m,n) \ = \ & \boldsymbol{f}(m+1,n) + \boldsymbol{f}(m-1,n) + \boldsymbol{f}(m,n+1) \\ & + \boldsymbol{f}(m,n-1) - 4\boldsymbol{f}(m,n) \ . \end{aligned}$$

Our motivation for choosing this transformation is two fold. First the Laplacian transform is a sparse representation of the image since it acts as an edge detector which provides a two-level image, the edges and the homogeneous background. Second, the Laplacian is a linear transformation. This latter property is 'critical' since the separation matrix estimated to separate the image mixtures is the same to separate the mixture of Laplacian images:

$$\boldsymbol{\mathcal{G}} = \frac{\partial^2 \boldsymbol{A}\boldsymbol{f}}{\partial x^2} + \frac{\partial^2 \boldsymbol{A}\boldsymbol{f}}{\partial y^2} = \boldsymbol{A}\boldsymbol{\mathcal{F}} \qquad (4)$$

where $\boldsymbol{\mathcal{G}}$ is the Laplacian transform of the mixtures. In the literature, some other linear transformations were proposed in order to

make the image sparse, including the projection into a sparse dictionary [7]. In Figure 1, the original cameraman image is displayed as well as its Laplacian transform and their respective histograms that clearly show the sparsity of the latter.
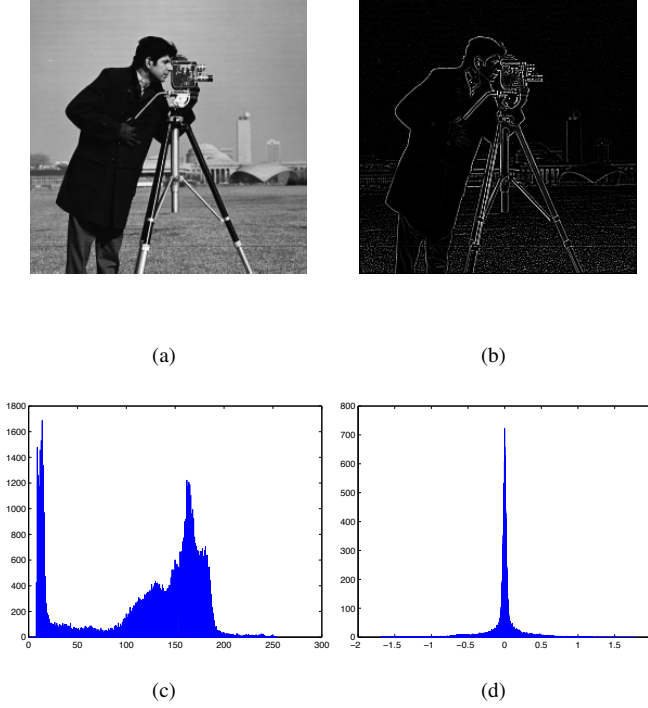

(a)   (b)

(c)   (d)

Fig. 1. (a) Original image, (b) Laplacian transform, (c) Original image histogram, (d) Sparse Laplacian transform histogram

In the pre-treatment phase, we also propose an optional whitening step which aims to set the mixtures to the same energy level. Furthermore, this procedures reduces the number of parameters to be estimated. More precisely, the whitening step is applied to the Laplacian image mixtures before using our separation algorithm. The whitening is achieved by applying a $N \times M$ matrix $\boldsymbol{Q}$ to the Laplacian image mixtures in such a way $\mathrm{Cov}(\boldsymbol{Q}\boldsymbol{\mathcal{G}}) = \boldsymbol{I}$ in the noiseless case, where $\mathrm{Cov}(\cdot)$ stands for the covariance operator. As shown in [8], $\boldsymbol{Q}$ can be computed as the inverse square root of the noiseless covariance matrix of the Laplacian image mixtures (see [8] for more details). In the following, we apply our separation algorithm on the whitened data:

$$\boldsymbol{\mathcal{G}}_w(m, n) = \boldsymbol{Q}\boldsymbol{\mathcal{G}}(m, n).$$

3.2. Sparsity-based BSS algorithm

In this section, we propose an iterative algorithm for the separation of sparse signals, namely the ISBS for Iterative Sparse Blind Separation algorithm. It is well known that Laplacian image transform is characterized by its sparsity property in the spatial domain [1, 9]. This property can be measured by the ℓ_p norm where $0 \leq p < 2$. More specifically, one can define the following sparsity based contrast function,

$$G_p(\boldsymbol{\mathcal{F}}) = \sum_{i=1}^{N} \left[\mathcal{J}_p(\mathcal{F}_i)\right]^{\frac{1}{p}} \qquad (5)$$

where

$$\mathcal{J}_p(\mathcal{F}_i) = \frac{1}{m_f n_f} \sum_{n=1}^{m_f} \sum_{m=1}^{n_f} |\mathcal{F}_i(m, n)|^p . \qquad (6)$$

The algorithm finds a separating matrix $\boldsymbol{B}$ such as,

$$\boldsymbol{B} = \arg\min_{\boldsymbol{B}} \{\mathbb{G}_p(\boldsymbol{B})\} \qquad (7)$$

where

$$\mathbb{G}_p(\boldsymbol{B}) \triangleq G_p(\boldsymbol{\mathcal{H}}) \qquad (8)$$

and $\boldsymbol{\mathcal{H}}(m, n) \triangleq \boldsymbol{B}\boldsymbol{\mathcal{G}}_w(m, n)$ represents the estimated image sources Laplacian. The approach we choose to solve (7) is inspired from [10]. It is a block technique based on the processing of $m_f n_f$ observed image pixels and consists in searching the minimum of the sample version of (7). Solutions are obtained iteratively in the form:

$$\boldsymbol{B}^{(k+1)} = (\boldsymbol{I} + \boldsymbol{\epsilon}^{(k)})\boldsymbol{B}^{(k)} \qquad (9)$$
$$\boldsymbol{\mathcal{H}}^{(k+1)}(m, n) = (\boldsymbol{I} + \boldsymbol{\epsilon}^{(k)})\boldsymbol{\mathcal{H}}^{(k)}(m, n) . \qquad (10)$$

At iteration k, a matrix $\boldsymbol{\epsilon}^{(k)}$ is determined from a local linearization of $G_p(\boldsymbol{B}\boldsymbol{\mathcal{G}}_w)$. It is an approximate Newton technique with the benefit that $\boldsymbol{\epsilon}^{(k)}$ can be very simply computed (no Hessian inversion) under the additional assumption that $\boldsymbol{B}^{(k)}$ is close to a separating matrix. This procedure is illustrated in the following.
At the $(k+1)^{th}$ iteration, the proposed criterion (6) can be developed as follows:

$$\mathcal{J}_p(\mathcal{H}_i^{(k+1)}) =$$
$$= \frac{1}{m_f n_f} \sum_{m=1}^{m_f} \sum_{n=1}^{n_f} \left| \mathcal{H}_i^{(k)}(m, n) + \sum_{j=1}^{N} \epsilon_{ij}^{(k)} \mathcal{H}_j^{(k)}(m, n) \right|^p .$$

Under the assumption that $\boldsymbol{B}^{(k)}$ is close to a separating matrix, we have

$$|\epsilon_{ij}^{(k)}| \ll 1$$

and thus, a first order approximation of $\mathcal{J}_p(\mathcal{H}_i^{(k+1)})$ is given by:

$$\mathcal{J}_p(\mathcal{H}_i^{(k+1)}) \approx \frac{1}{m_f n_f} \sum_{m=1}^{m_f} \sum_{n=1}^{n_f} \left\{ |\mathcal{H}_i^{(k)}(m, n)|^p + \right.$$
$$\left. p\sum_{j=1}^{N} \epsilon_{ij}^{(k)} \left(|\mathcal{H}_i^{(k)}(m, n)|^{p-1} \mathrm{sgn}\left(\mathcal{H}_i^{(k)}(m, n)\right) \mathcal{H}_j^{(k)}(m, n) \right) \right\}$$
$$(11)$$

where $\mathrm{sgn}(\cdot)$ represents the sign value operator. Using equation (11), equation (5) can be rewritten in more compact form as:

$$\mathbb{G}_p\left(\boldsymbol{B}^{(k+1)}\right) = \mathbb{G}_p\left(\boldsymbol{B}^{(k)}\right) + Tr\left(\boldsymbol{\epsilon}^{(k)}\boldsymbol{\mathcal{R}}^{(k)T}\boldsymbol{D}^{(k)}\right) \qquad (12)$$

where $Tr(\cdot)$ is the matrix trace operator, the ij^{th} entry of matrix $\boldsymbol{\mathcal{R}}^{(k)}$ is given by:

$$\mathcal{R}_{ij}^{(k)} = \frac{1}{m_f n_f} \sum_{m=1}^{m_f} \sum_{n=1}^{n_f} |\mathcal{H}_i^{(k)}(m, n)|^{p-1} \mathrm{sgn}\left(\mathcal{H}_i^{(k)}(m, n)\right) \mathcal{H}_j^{(k)}(m, n) .$$

and

$$\boldsymbol{D}^{(k)} = \left[\mathrm{diag}\left(\mathcal{R}_{11}^{(k)}, \dots, \mathcal{R}_{NN}^{(k)}\right) \right]^{\frac{1}{p}-1} . \qquad (13)$$

Using a gradient technique, $\boldsymbol{\epsilon}^{(k)}$ can be written as:

$$\boldsymbol{\epsilon}^{(k)} = -\mu\boldsymbol{D}^{(k)}\boldsymbol{\mathcal{R}}^{(k)} , \qquad (14)$$

where $\mu > 0$ is the descent step. Replacing (14) into (12) leads to,

$$\mathbb{G}_p\left(\boldsymbol{B}^{(k+1)}\right) = \mathbb{G}_p\left(\boldsymbol{B}^{(k)}\right) - \mu\|\boldsymbol{D}^{(k)}\boldsymbol{\mathcal{R}}^{(k)}\|^2 , \qquad (15)$$

so μ controls the decrement of the criterion. Now, to avoid the algorithm's convergence to the trivial solution $\boldsymbol{B} = \boldsymbol{0}$, one normalizes the outputs of the separating matrix to unit-power, i.e. $\rho_{\mathcal{H}_i}^{(k+1)} \triangleq E\left(|\mathcal{H}_i^{(k+1)}(m,n)|^2\right) = 1, \ \forall\, i$. Using first order approximation, this normalization leads to:

$$\epsilon_{ii}^{(k)} = \frac{1 - \rho_{\mathcal{H}_i}^{(k)}}{2\rho_{\mathcal{H}_i}^{(k)}} . \qquad (16)$$

The final estimated separation matrix $\boldsymbol{B} = \boldsymbol{B}^{(\mathcal{K})}\boldsymbol{Q}$ is applied to the image mixtures $\boldsymbol{g}$ to obtain an estimation of the original images. $\mathcal{K}$ denotes here the number of iterations that can be either chosen a priori or given by a stopping criterion of the form $\|\boldsymbol{B}^{(k+1)} - \boldsymbol{B}^{(k)}\| < \delta$ where δ is a small threshold value.

4. PERFORMANCE EVALUATION

All simulations are carried on 256×256 parrot and cameraman images. The number of observed mixtures is $M = 2$. The algorithms are developed on MATLAB environment. Monte Carlo simulations are carried over 200 random realizations of the additive gaussian noise for variable Signal to Noise Ratios (SNR). In order to objectively evaluate the performance of the proposed algorithm, we consider two different criteria, the first one is the Interference to Signal Ratio (ISR) criterion [8] defined as:

$$ISR \triangleq \sum_{i=1}^{N} \sum_{j \neq i}^{N} \frac{E\left(|(\boldsymbol{BA})_{ij}|^2\right)\rho_j}{E\left(|(\boldsymbol{BA})_{ii}|^2\right)\rho_i} \qquad (17)$$

where $\rho_i = E(|f_i(m,n)|^2)$ is the ith source power. The second one is an objective image quality measure inspired from the Human Visual System (HVS) properties and developed in [11]. It is called $PSNR - WAV$ for Peak Signal to Noise Ratio based on Wavelet decomposition. The separation result of the proposed algorithm is depicted on Fig. 2 where we represent the two original images ($f_1, \ f_2$), the mixtures ($g_1, \ g_2$) and the recovered ones ($\widehat{f}_1, \ \widehat{f}_2$) by the proposed algorithm in the noiseless case. In Fig. 3, we compare the performance of the proposed algorithm to the Relative Newton algorithm developed by Zibulevsky et al. in [9] where the case of sparse sources is considered. We plot the residual interference between separated images (ISR) versus the SNR. It is clearly shown that our algorithm (ISBS) performs better in terms of ISR especially for low SNRs. We plot on Fig. 4 the objective distortion measure between the original and separated images versus the SNR. One can observe that, for each image, we reach the same conclusion for the $PSNR - WAV$ as for the ISR. In Fig. 5, we represent the evolution of the ISR as a function of the iteration number. A fast convergence rate is observed. Moreover, the complexity of the proposed algorithm is equal to $2N^2 m_f n_f + \mathcal{O}(N^2)$ flops per iteration whereas the complexity of the Relative Newton algorithm in [9] is $2N^4 + N^3 m_f n_f + N^6/6$.

5. CONCLUSION

This article deals with a simple and efficient two-step algorithm of blind image separation. The proposed method consists in a sparsification of the natural observed mixtures followed by a blind separation of the original images. The sparsification is simply the Laplacian transform and has a low computational cost. The separation is performed using an iterative algorithm based on the minimizing of
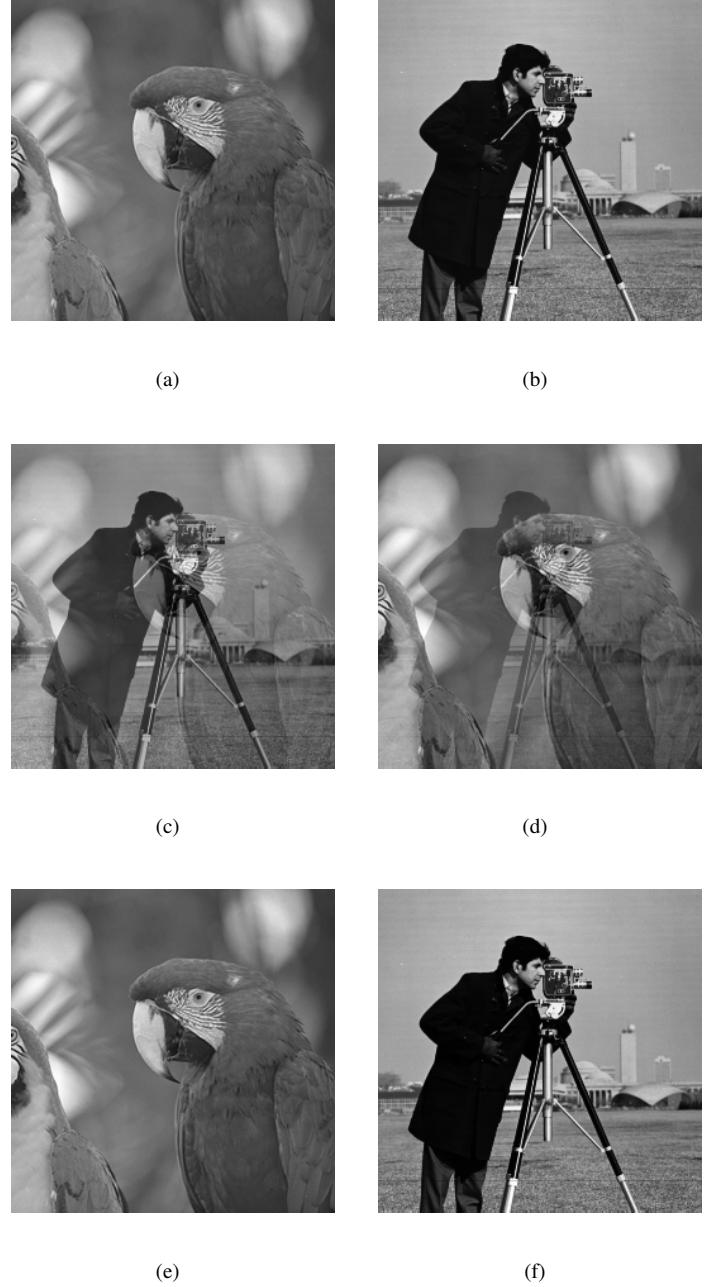

Fig. 2. (a)-(b) original images, (c)-(d) $M = 2$ observed mixtures, (e)-(f) restored images using ISBS algorithm.

the ℓ_p norm of the Laplacian image. The proposed technique outperforms existing solutions in terms of separation quality and computational cost.

6. REFERENCES

[1] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing*, Wiley & Sons, Ltd., UK, 2003.

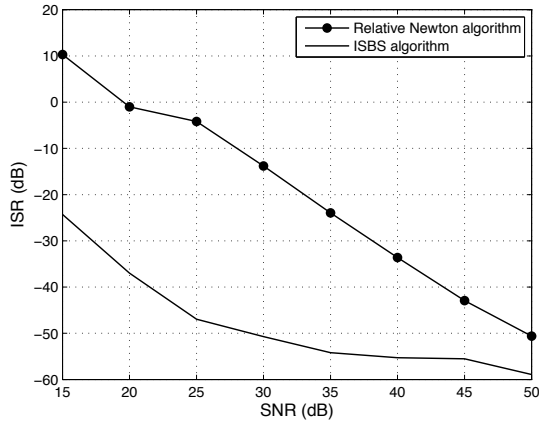[2] N. Zhang, J. LU, and T. Yahgi, "Nonlinear blind source separa-

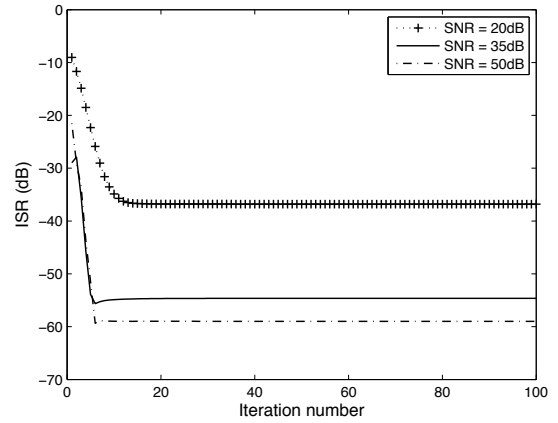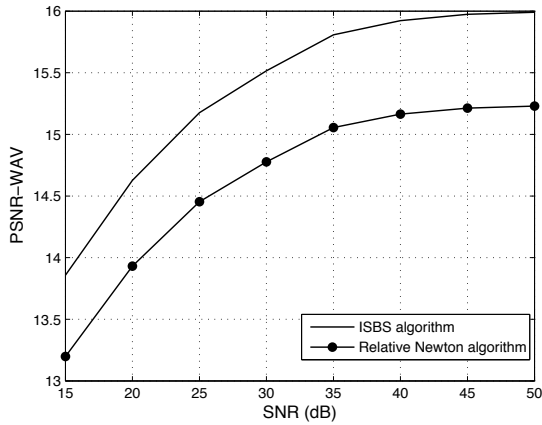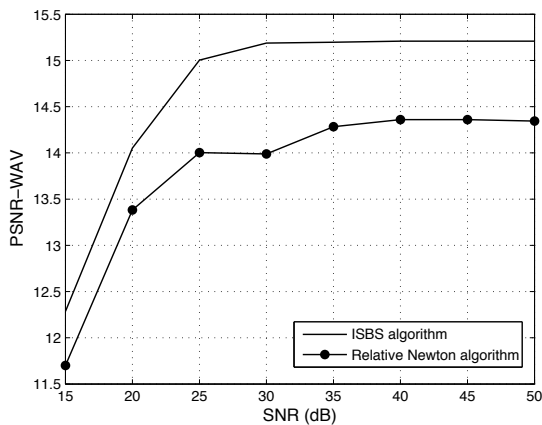**Fig. 3**. Interference to Signal Ratio (ISR) versus SNR

Fig. 5. Evolution of the Interference to Signal Ratio (ISR) as a function of the iteration number.

(a) $PSNR - WAV$ for first image source

(b) $PSNR - WAV$ for second image source

Fig. 4. Objective image quality measure ($PSNR - WAV$)

tion method for X-ray image separation," *IEICE Transactions on Fundamentals of Electronics, Communications and Com-*

puter Sciences, vol. E89-A, no. 4, pp. 924–931, 2006.

[3] J. Kolba and I. Jouny, "Blind source separation in tumor detection in mammograms," in *Proc. of the IEEE 32nd Annual Northeast Bioengineering Conference*, April 2006, pp. 65–66.

[4] M. S. Naceur, M. A. Loghmari, and M. R. Boussema, "The contribution of the sources separation method in the decomposition of mixed pixels," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 11, pp. 2642–2653, 2004.

[5] A. Bijaoui and D. Nuzillard, "Blind source separation of multispectral astronomical images," in *Proc. MPA/ESO/MPE Joint Astronomy Conf., Mining the Sky*, Germany, August 2000.

[6] M. M. Bronstein, A. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi, "Separation of reflections via sparse ICA," in *Proc. International Conference on Image Processing ICIP*, September 2003, vol. 1, pp. 313–316.

[7] M. Zibulevsky and B. A. Pearlmutter, "Blind source separation by sparse decomposition in signal dictionary," *Neural Computation*, vol. 13, no. 4, pp. 863–882, 2001.

[8] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Transaction on Signal Processing*, vol. 45, no. 2, pp. 434–444, February 1997.

[9] M. Zibulevsky, "Sparse source separation with relative Newton method," in *Proc. ICA*, Apr. 2003, pp. 897–902.

[10] D. T. Pham and P. Garat, "Blind separation of mixture of independent sources through a quasi-maximum likelihood approach," *IEEE Transactions on Signal Processing*, vol. 45, no. 7, pp. 1712–1725, July 1997.

[11] A. Beghdadi and B. Pesquet-Popescu, "A new image distortion measure based on wavelet decomposition," in *Proc. International Symposium on Signal Processing and Its Applications ISSPA*, Paris, France, July 2003, vol. 1, pp. 485–488.