Image Deconvolution by Stein Block Thresholding

Christophe Chesneau, Jalal M. Fadili, Jean-Luc Starck

IMAGE DECONVOLUTION BY STEIN BLOCK THRESHOLDING

C. Chesneau[a], M. J. Fadili[b], and J. L. Starck[c]

[a] Laboratoire de Mathématiques Nicolas Oresme, Université de Caen, 14032 Caen, France.
[b] GREYC CNRS-ENSICAEN-Université de Caen, 14050 Caen France.
[c] DAPNIA/SEDI-SAP CEA-Saclay 14050 Caen France 91191 Gif-sur-Yvette France.

ABSTRACT

In this paper, we propose a fast image deconvolution algorithm that combines adaptive block thresholding and Vaguelet-Wavelet Decomposition. The approach consists in first denoising the observed image using a wavelet-domain Stein block thresholding, and then inverting the convolution operator in the Fourier domain. Our main theoretical result investigates the minimax rates over Besov smoothness spaces, and shows that our block estimator can achieve the optimal minimax rate, or is at least nearly-minimax in the least favorable situation. The resulting algorithm is simple to implement and fast. Its computational complexity is dominated by that of the FFT in the Fourier-domain inversion step. We report a simulation study to support our theoretical findings. The practical performance of our block vaguelet-wavelet deconvolution compares very favorably to existing competitors on a large set of test images.

Index Terms— Image deconvolution, Block thresholding, Wavelets, Minimax.

1. INTRODUCTION

In this paper, we consider the two-dimensional convolution model with Gaussian white noise $\sim \mathcal{N}(0, \sigma^2)$. We observe the stochastic process $Y(.)$ where

$$Y(d\mathbf{x}) = T(f)(\mathbf{x})d\mathbf{x} + \sigma W(d\mathbf{x}), \qquad (1)$$

$\mathbf{x} \in [0,1]^2$, $W(.)$ is a (non-observed) white Gaussian noise, $T(f)(\mathbf{x}) = (f \star g)(\mathbf{x})$ is the two-dimensional convolution operator on $[0,1]^2$, g is a known kernel (called also point spread function PSF), both f and g are one-periodic functions belonging to $\mathbb{L}_2([0,1]^2)$. In the sequel, the Fourier transform of a function f will be denoted $\mathcal{F}(f)(\mathbf{l}) = \int_{[0,1]^2} f(\mathbf{x}) e^{-i2\pi <\mathbf{l},\mathbf{x}>} d\mathbf{x}$. The observation model (1) illustrates the action of a linear time-invariant system on an input image f when the data are corrupted with additional noise. The deconvolution is to estimate f from Y which is a longstanding inverse problem in image processing.

There is an extensive statistical literature on wavelet-based deconvolution problems. For obvious space limitations, we only focus on *some* of them. In 1D, Donoho in [1] gave the first discussion of wavelet thresholding in linear inverse problems and introduced the Wavelet-Vaguelet Decomposition (WVD). The WaveD algorithm of [2] is an adaptation of WVD to the one dimensional deconvolution problem. Abramovich and Silverman in [3] proposed another procedure; the Vaguelet-Wavelet Decomposition (VWD). The original estimator based on VWD is defined with standard term-by-term thresholding rules. It has been improved by [4] using a Stein block thresholding rule. As for VWD, the original WaveD procedure based on term-by-term thresholding has been recently improved by [5] using again block thresholding.

In 2D, the WVD approach was refined in [6] who proposed a mirror wavelet basis adapted to capture the singularity of the spectrum of the inverse of h. The authors in [7] advocated a hybrid approach known as ForWarD. In [8], the authors proposed an adaptive wavelet estimator based on two-dimensional version of the WaveD algorithm of [2] which enjoys good numerical performance. Deconvolution methods based on variational or bayesian formulations with sparsity promoting regularization over wavelet coefficients have been recently proposed; see e.g. [9, 10, 11, 12]. These algorithms are based on iterative thresholding.

However, so far, these wavelet deconvolution algorithms were based on term-by-term thresholding which underperforms for many images. The drawback of individual cannot be circumvented by fine-tuning the regularization/threshold parameter. All these reasons motivated us to develop an adaptive estimator of f based on combining two-dimensional Stein block thresholding and VWD. The approach consists in first denoising the observed image using a wavelet-domain block thresholding, and then inverting the convolution operator in the Fourier domain. It can be viewed as a multi-dimensional version of the procedure developed by [4]. From a theoretical point of view, taking the minimax approach over the Besov balls $\mathbf{B}^s_{p,q}(M)$ (to be defined in Section 2) and under the $\mathbb{L}_2$ risk, we prove that our estimator achieves near optimal rates of convergence. These rates are for instance better than those attained by the two-dimensional WaveD of [8]. From a practical point of view, our algorithm is very simple to implement and runs very fast. Its performances compare very favorably to alternative deconvolution algorithms such as [9, 10, 7, 8]

over a large set of test images.

The paper is organized as follows. Section 2 briefly reviews wavelets and Besov balls. Section 3 describes the block thresholding-based deconvolution estimator. The minimax performances of this estimator are investigated in Section 4. Section 5 contains experimental results.

2. WAVELETS AND BESOV BALLS

We consider the tensor product wavelet basis on $\mathbb{L}_2([0,1]^2)$. Let us briefly recall the construction of such a basis (see, for instance, [13]).

Consider compactly supported scaling and wavelet functions, ϕ and ψ. Let us define the tensor-product wavelets Φ, Ψ^1, Ψ^2 and Ψ^3 as $\Phi(\mathbf{x}) = \phi(x)\phi(y)$, $\Psi^1(\mathbf{x}) = \psi(x)\phi(y)$, $\Psi^2(\mathbf{x}) = \phi(x)\psi(y)$ and $\Psi^3(\mathbf{x}) = \psi(x)\psi(y)$, $\forall \mathbf{x} = (x,y) \in [0,1]^2$. For any orientation $i \in \{1,2,3\}$, scale $j \geq 0$ and spatial location $\mathbf{k} = (k_1,k_2) \in D_j = \{0,\dots,2^j-1\}^2$, we define the translated and scaled versions $\Phi_{j,\mathbf{k}}(\mathbf{x}) = 2^j\Phi(2^jx - k_1, 2^jy - k_2)$ and $\Psi_{j,i,\mathbf{k}}(\mathbf{x}) = 2^j\Psi^i(2^jx - k_1, 2^jy - k_2)$. Any function $f \in \mathbb{L}_2([0,1]^2)$ can be expanded into a wavelet series

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in D_{j_0}} \alpha_{j_0,\mathbf{k}}\Phi_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{i=1}^{3}\sum_{j\geq j_0}\sum_{\mathbf{k}\in D_j} \beta_{j,i,\mathbf{k}}\Psi_{j,i,\mathbf{k}}(\mathbf{x}),$$

$\mathbf{x} \in [0,1]^2$, where $\alpha_{j,\mathbf{k}} = \int_{[0,1]^2} f(\mathbf{x})\Phi_{j,\mathbf{k}}(\mathbf{x})d\mathbf{x}$ and $\beta_{j,i,\mathbf{k}} = \int_{[0,1]^2} f(\mathbf{x})\Psi_{j,i,\mathbf{k}}(\mathbf{x})d\mathbf{x}$ are the wavelet coefficients of f.

We say that a function f in $\mathbb{L}_2([0,1]^2)$ belongs to the bi-dimensional (isotropic) Besov ball $\mathbf{B}^s_{p,q}(M)$ if, and only if, $\int_{[0,1]^2} f^2(\mathbf{x})d\mathbf{x} \leq M$ and there exists a constant M_*, depending on M, such that the wavelet coefficients of f satisfy

$$\left(\sum_{i=1}^{3}\sum_{j\geq 0}\left(2^{j(s+1-2/p)}\left(\sum_{\mathbf{k}\in D_j}|\beta_{j,i,\mathbf{k}}|^p\right)^{1/p}\right)^q\right)^{1/q} \leq M_*,$$

with a smoothness parameter $s > 0$, and the norm parameters: $0 < p \leq \infty$ and $0 < q \leq \infty$. Such Besov spaces contain both smooth images and those with sharp edges.

3. THE DECONVOLUTION BLOCK ESTIMATOR

3.1. Smoothness of the kernel g

For the theoretical study, the following assumption on g will be essential. It is similar to the one employed in [1, 8, 7, 4]. We suppose that there exist four constants, $c > 0$, $C > 0$, $\delta_1 > 1/2$ and $\delta_2 > 1/2$, such that, for any $\mathbf{l} = (l_1,l_2) \in ((-\infty,-1]\cup[1,\infty))^2$, the Fourier transform of g satisfies

$$c|l_1|^{-\delta_1}|l_2|^{-\delta_2} \leq |\mathcal{F}(g)(\mathbf{l})| \leq C|l_1|^{-\delta_1}|l_2|^{-\delta_2}. \quad (2)$$

In words, this means that the Fourier transform of the blurring PSF decays in a polynomial fashion within its bandwidth. For example, it is easy to check that the square integrable one-periodic function g defined by $g(x,y) = h(x)h(y)$ where $h(x) = \sum_{m\in\mathbb{Z}} e^{-|x+m|}$, $x \in [0,1]$, satisfies (2). Indeed, for any $l \in \mathbb{Z}$, we have $\mathcal{F}(h)(l) = 2\left(1+4\pi^2l^2\right)^{-1}$. Hence, for any $\mathbf{l} = (l_1,l_2) \in ((-\infty,-1]\cup[1,\infty))^2$, $\mathcal{F}(g)(\mathbf{l}) = \mathcal{F}(h)(l_1)\mathcal{F}(h)(l_2)$ satisfies (2) with $c = 4(1+4\pi^2)^{-2}$, $C = (2\pi^2)^{-2}$ and $\delta_1 = \delta_2 = 2$. This assumption goes by the name of *ordinary smooth case*.

3.2. Vaguelet-Wavelet decomposition

Although the VWD is valid for more general operators T, we here restrict our description to the case of convolution where the VWD takes a simple form. Thus, under assumption (2), any function $f \in \mathbb{L}_2([0,1]^2)$ can be expanded into a vaguelet-wavelet series

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in D_{j_0}} \vartheta_{j_0,\mathbf{k}}\omega_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{i=1}^{3}\sum_{j\geq j_0}\sum_{\mathbf{k}\in D_j} \theta_{j,i,\mathbf{k}}w_{j,i,\mathbf{k}}(\mathbf{x}),$$

$\mathbf{x} \in [0,1]^2$, where $\vartheta_{j_0,\mathbf{k}} = \int_{[0,1]^2} T(f)(\mathbf{x})\Phi_{j_0,\mathbf{k}}(\mathbf{x})d\mathbf{x}$, $\theta_{j,i,\mathbf{k}} = \int_{[0,1]^2} T(f)(\mathbf{x})\Psi_{j,i,\mathbf{k}}(\mathbf{x})d\mathbf{x}$, $\omega_{j_0,\mathbf{k}}(\mathbf{x}) = T^{-1}(\Phi_{j_0,\mathbf{k}})(\mathbf{x})$ and $w_{j,i,\mathbf{k}}(\mathbf{x}) = T^{-1}(\Psi_{j,i,\mathbf{k}})(\mathbf{x})$. Note that, for any $h \in \mathbb{L}_2([0,1]^2)$, $T^{-1}(h)(\mathbf{x}) = \int_{[0,1]^2} \left(\mathcal{F}(h)(\mathbf{l})/\mathcal{F}(g)(\mathbf{l})\right) e^{i2\pi<\mathbf{l},\mathbf{x}>}d\mathbf{l}$.

3.3. Gaussian sequence model

The first step to estimate f consists in estimating the unknown wavelet coefficients of $T(f)$: $(\vartheta_{j_0,\mathbf{k}})_{\mathbf{k}}$ and $(\theta_{j,i,\mathbf{k}})_{j,i,\mathbf{k}}$ from the observation Y in (1). It follows from (1) that

$$y_{j,i,\mathbf{k}} = \theta_{j,i,\mathbf{k}} + \sigma e_{j,i,\mathbf{k}}, \quad (3)$$

where $y_{j,i,\mathbf{k}} = \int_{[0,1]^2} \Psi_{j,i,\mathbf{k}}(\mathbf{x})Y(d\mathbf{x})$ and $e_{j,i,\mathbf{k}}$ is the noise: $e_{j,i,\mathbf{k}} = \int_{[0,1]^2} \Psi_{j,i,\mathbf{k}}(\mathbf{x})W(d\mathbf{x})$. Thanks to the orthonormality of the wavelet basis, the random variables $(e_{j,i,\mathbf{k}})_{j,i,\mathbf{k}}$ are i.i.d. $\sim \mathcal{N}(0,1)$. Therefore, $y_{j,i,\mathbf{k}}$ is a natural estimator for $\theta_{j,i,\mathbf{k}}$.

3.4. Two-dimensional block thresholding estimator

Let the observed image be defined on a $n \times n$ discrete grid of equally-spaced pixels $\left\{Y(i/n, j/n);\ (i,j) \in \{1,\dots,n\}^2\right\}$. Let $L = \lfloor(2\log(n))^{1/2}\rfloor$ be the block length, $j_0 = \lfloor\log_2 L\rfloor$ is the coarsest decomposition scale, and $J_* = \lfloor(1/(\delta_1 + \delta_2))\log_2(n)\rfloor$. Consider the sequence model (3). For any $\mathbf{k} \in D_{j_0}$, we set $\widehat{\vartheta}_{j_0,\mathbf{k}} = \int_{[0,1]^2} \Phi_{j_0,\mathbf{k}}(\mathbf{x})dY(\mathbf{x})$. For any $j \in \{j_0,\dots,J_*\}$, let $\mathcal{A}_j = \left\{1,\dots,\lfloor 2^jL^{-1}\rfloor\right\}^2$ be the set indexing the blocks at scale j, and for each block index $\mathbf{K} = (K_1,K_2) \in \mathcal{A}_j$, $U_{j,\mathbf{K}} = \{\mathbf{k} \in D_j;\ (K_1-1)L \leq k_1 \leq K_1L-1,\ (K_2-1)L \leq k_2 \leq K_2L-1\}$ is the set indexing the positions of coefficients within the $\mathbf{K}$th block $U_{j,\mathbf{K}}$.

For any $\mathbf{k} \in U_{j,\mathbf{K}}$, $\mathbf{K} \in \mathcal{A}_j$ and $i \in \{1,2,3\}$, we estimate the wavelet coefficients $\theta_{j,i,\mathbf{k}}$ of $T(f)$ from $y_{j,i,\mathbf{k}}$ in (3) as

- $\widehat{\theta}_{j,i,\mathbf{k}} = y_{j,i,\mathbf{k}}$ if $j \in \{0, \ldots, j_0 - 1\}$;
- $\widehat{\theta}_{j,i,\mathbf{k}} = y_{j,i,\mathbf{k}} \left(1 - \frac{\lambda_* \sigma^2}{\frac{1}{L^2} \sum_{\mathbf{k} \in U_{j,\mathbf{K}}} y_{j,i,\mathbf{k}}^2}\right)_+$ if $j \in \{j_0, \ldots, J_*\}$;
- $\widehat{\theta}_{j,i,\mathbf{k}} = 0$ if $j > J_*$.

where $(a)_+ = \max(a, 0)$, and λ_* is the root of $x - \log x = 3$, i.e. $\lambda_* = 4.50524\ldots$. To estimate f, we reconstruct it from these block-thresholded coefficients as

$$\widehat{f}(\mathbf{x}) = \sum_{\mathbf{k} \in D_{j_0}} \widehat{\vartheta}_{j_0,\mathbf{k}} \omega_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{i=1}^{3} \sum_{j=j_0}^{J_*} \sum_{\mathbf{K} \in \mathcal{A}_j} \sum_{\mathbf{k} \in U_{j,\mathbf{K}}} \widehat{\theta}_{j,i,\mathbf{k}} w_{j,i,\mathbf{k}}(\mathbf{x}), \quad (4)$$

$\mathbf{x} \in [0,1]^2$. It is easy to see that the latter vaguelet reconstruction formula is also equivalent to first applying an inverse wavelet transform to $(\widehat{\vartheta}_{j_0,\mathbf{k}}, \widehat{\theta}_{j,i,\mathbf{k}})$ to get the estimate $\widehat{T(f)}$, and then inverting the convolution operator in the Fourier domain.

3.5. Deconvolution algorithm

The deconvolution algorithm can be summarized as follows:

Parameters: The observed blurred and noisy image Y, the PSF g.
Initialization:

- Block size $L = \lfloor (2\log(n))^{1/2} \rfloor$, coarsest decomposition scale $j_0 = \lfloor \log_2 L \rfloor$, threshold $\lambda_* = 4.50524$.

Block Stein thresholding:

- Wavelet transform: compute $\widehat{\vartheta}_{j_0,\mathbf{k}} = \int_{[0,1]^2} \Phi_{j_0,\mathbf{k}}(\mathbf{x}) dY(\mathbf{x})$ and $y_{j,i,\mathbf{k}} = \int_{[0,1]^2} \Psi_{j,i,\mathbf{k}}(\mathbf{x}) dY(\mathbf{x}), \forall j \geq j_0$.
- Keep $\widehat{\vartheta}_{j_0,\mathbf{k}}$ intact, and threshold $y_{j,i,\mathbf{k}}$ by blocks $U_{j,\mathbf{K}}$ to get $\widehat{\theta}_{j,i,\mathbf{k}} = y_{j,i,\mathbf{k}} \left(1 - \frac{\lambda_* \sigma^2}{\frac{1}{L^2} \sum_{\mathbf{k} \in U_{j,\mathbf{K}}} y_{j,i,\mathbf{k}}^2}\right)_+$.
- Inverse wavelet transform of $(\widehat{\vartheta}_{j_0,\mathbf{k}}, \widehat{\theta}_{j,i,\mathbf{k}})$ to get $\widehat{T(f)}$.

Output: Get deconvolved image $\widehat{f}$ from $\widehat{T(f)}$ by inverting the convolution operator in the Fourier domain.

It is worth noting that in practice, the last step of this algorithm can be modified to handle ill-conditioned convolution kernels (beyond the ordinary smooth case (2) required for theoretical reasons) by regularizing the inverse of the kernel in the Fourier domain, e.g. using Wiener deconvolution.

4. OPTIMALITY RESULT

Theorem 1 below investigates the minimax rates of convergence attained by $\widehat{f}$ over $\mathbf{B}^s_{p,q}(M)$ under the $\mathbb{L}_2$ risk.

Theorem 1 *Consider the model* (1). *Let* $\widehat{f}$ *be the estimator defined by* (4). *Then there exists a constant* $C > 0$ *such that*

$$\sup_{f \in \mathbf{B}^s_{p,q}(M)} \mathbb{E}\left(\int_{[0,1]^2} \left(\widehat{f}(\mathbf{x}) - f(\mathbf{x})\right)^2 d\mathbf{x}\right) \leq C v_\sigma,$$

where

$$v_\sigma = \begin{cases} \sigma^{2s/(s+\delta_1+\delta_2+1)}, & \text{for } 2 \leq p, \\ (\sigma \log(n))^{2s/(s+\delta_1+\delta_2+1)}, & \text{for } p < 2,\ sp > c, \end{cases} \quad (5)$$

$c = 2 \vee (2-p)(\delta_1 + \delta_2 + 1)$.

Using lower bound techniques, one can prove that v_σ is optimal except in the cases $p < 2$ where there is an extra logarithmic term. It can also be shown that v_σ is better than the one achieved by the conventional term-by-term thresholding estimators (WaveD [8], etc). The main difference is for the case $p \geq 2$ where there is no extra logarithmic term.

5. EXPERIMENTAL RESULTS

The proposed block VWD deconvolution method has been compared to three deconvolution methods from the literature: ForWarD [7], wavelet-domain iterative thresholding (IT) with 100 iterations [9, 10], and WaveD [8]. For fair comparison, all methods used the Symmlet 6 wavelet except WaveD that uses Meyer wavelets, and the regularization parameter of iterative thresholding was tweaked manually to reach its best performance. For reliable comparison, we applied the deconvolution algorithms to six standard grayscale images of size 512×512 (Barbara, Lenna, Boat) and 256×256 (Cameraman, House, Peppers). The blurred images were corrupted by a zero-mean white Gaussian noise such that the blurred signal-to-noise ratio ($\mathrm{BSNR} = 10 \log_{10}(\|f \star g\|_\infty / \sigma^2)$) ranged from 10 to 40 dB. At each combination of test image and noise level, ten noisy versions were generated and each deconvolution algorithm was applied to each noisy realization. The output SNR improvement (ISNR) was averaged over the ten replications. The results are shown in Table 1 where the PSF was $g(i,j) = e^{-\left(|i|^{0.5} + |j|^{0.5}\right)}$ (similar results were obtained with other PSFs not shown due to space limitation). Each table corresponds to the ISNR as a function of BSNR for each image. These results clearly show that our approach compares very favorably to ForWarD and iterative thresholding. It is even able to outperform them particularly at low BSNR, while having substantially less computational cost as reported in Table 2. These quantitative results are confirmed by visual inspection of Fig. 1 which displays the result on Barbara for BSNR=30dB. Again, owing to block thresholding, our VWD deconvolution is able to recover many details (e.g. textured areas) much better that the other methods.

Barbara 512×512							Lenna 512×512						Boat 512×512					
BSNR (dB)	10	15	20	25	30	40	10	15	20	25	30	40	10	15	20	25	30	40
Ours	2.93	2.56	2.90	3.66	5.68	11.11	5.08	5.13	6.15	7.81	8.63	12.82	3.69	3.87	5.08	6.71	7.66	11.98
ForWarD [7]	0.00	2.00	2.91	3.58	4.92	10.24	0.77	4.25	6.16	7.46	8.76	12.18	1.00	3.70	5.17	6.38	7.75	11.42
IT [9, 10]	2.87	2.34	2.84	3.26	4.33	7.69	4.82	4.67	6.32	7.62	9.00	12.66	3.30	3.31	5.26	6.52	8.32	11.89
WaveD [8]	2.33	2.32	2.50	2.91	3.41	6.82	3.99	4.64	5.59	6.99	8.47	10.51	2.89	3.16	4.06	5.36	6.67	8.77
Cameraman 256×256							House 256×256						Peppers 256×256					
Ours	3.47	3.53	4.79	5.89	7.84	11.40	5.28	5.76	7.54	9.13	10.20	14.55	4.73	5.79	7.83	9.04	11.04	14.41
ForWarD [7]	0.05	2.15	4.32	6.29	8.16	12.40	1.27	4.63	7.76	9.75	11.06	13.84	1.97	5.04	7.66	9.67	11.51	15.46
IT [9, 10]	3.27	3.18	4.77	5.86	7.86	11.75	5.21	5.67	7.73	9.25	10.74	14.58	3.69	4.28	7.07	8.81	11.01	15.01
WaveD [8]	3.08	2.93	3.31	4.40	5.60	7.47	4.50	4.91	5.41	6.31	8.07	9.81	3.33	3.54	4.05	5.07	6.03	10.69

Table 1. Comparison of average ISNR in dB over ten realizations for various images.

6. CONCLUSION

In this paper, a wavelet-based deconvolution algorithm was presented. It combines the benefits of block-thresholding with vaguelet-wavelet decomposition. Its theoretical and practical performances were established. Although we focused on convolution, the approach can handle other operators $T(f)$. A possible perspective of the present work that we are currently investigating is the theoretical properties of the procedure when other transforms than wavelets (curvelets for instance) are used.

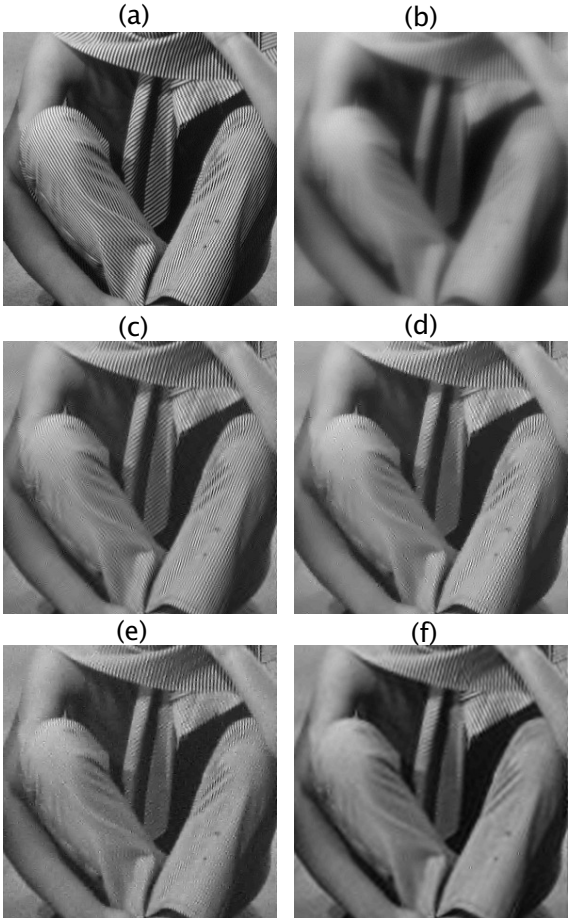

Fig. 1. Deconvolution of Barbara 512×512. (a) original, (b) blurred and noisy BSNR=30dB, (c) our method ISNR=5.66dB, (d) ForWarD [7] ISNR=4.9dB, (e) iterative thresholding [9, 10] ISNR=4.33dB, (f) WaveD [8] ISNR=3.14dB.

Algorithm	Ours	ForWarD [7]	IT [9, 10]	WaveD [8]
512×512	0.62	5.2	119	4.08
256×256	0.15	1.05	29	0.76

Table 2. Average execution times (seconds) over then replications for 512×512 and 256×256 images. The algorithms were run under Matlab with an 2.53GHz Intel Core Duo CPU, 4Gb RAM.

Acknowledgment. This work is supported by ANR grant NatImages, ANR-08-EMER-009.

7. REFERENCES

[1] D.L. Donoho, "Nonlinear solution of inverse problems by wavelet-vaguelette decomposition," *Applied and Computational Harmonic Analysis*, vol. 2, pp. 101–126, 1995.

[2] I.M. Johnstone, G. Kerkyacharian, D. Picard, and M. Raimondo, "Wavelet deconvolution in a periodic setting," *Journal of the Royal Statistical Society. Series B. Methodological*, vol. 66, pp. 547–573, 2004.

[3] F. Abramovich and B. W. Silverman, "Wavelet decomposition approaches to statistical inverse problems," *Biometrika*, vol. 85, pp. 115–129, 1998.

[4] T. Cai, "On adaptive wavelet estimation of a derivative and other related linear inverse problems," *J. Statistical Planning and Inference*, vol. 108, pp. 329–349, 2002.

[5] C. Chesneau, "Wavelet estimation via block thresholding : A minimax study under the L_p risk," *Statistica Sinica*, vol. 18, no. 3, pp. 1007–1024, 2008.

[6] J. Kalifa, S. Mallat, and B. Rougé, "Image deconvolution in mirror wavelet bases," in *IEEE ICIP*, 1998, vol. 1, pp. 565–569.

[7] R. Neelamani, H. Choi, and R. Baraniuk, "Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems," *IEEE Transactions on signal processing*, vol. 52, pp. 418–433, 2004.

[8] D. Donoho and M. Raimondo, "A fast wavelet algorithm for image deblurring," *The Australian & New Zealand Industrial and Applied Mathematics Journal*, vol. 46, pp. C29–C46, 2005.

[9] M. Figueiredo and R. Nowak, "An EM algorithm for wavelet-based image restoration," *ITIP*, vol. 12, no. 8, pp. 906–916, 2003.

[10] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure Appl. Math.*, vol. 57, pp. 1413–1541, 2004.

[11] M. J. Fadili and J.-L. Starck, "Sparse representation-based image deconvolution by iterative thresholding," in *ADA IV*, France, 2006, Elsevier.

[12] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs, "A variational formulation for frame-based inverse problems," *Inv. Prob.*, vol. 23, pp. 1495–1518, 2007.

[13] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 2nd edition, 1998.