TWO DENOISING SURE-LET METHODS FOR COMPLEX OVERSAMPLED SUBBAND DECOMPOSITIONS

Jérôme Gauthier[1], Laurent Duval[2], and Jean-Christophe Pesquet[1]

[1] Université Paris-Est
Institut Gaspard Monge and UMR CNRS 8049, 77454 Marne-La-Vallée Cedex 2, France
jerome.gauthier@univ-mlv.fr, jean-christophe.pesquet@univ-paris-est.fr
[2] Institut Français du Pétrole (IFP)
1 et 4, avenue de Bois-Préau, 92852 Rueil-Malmaison Cedex, France
laurent.duval@ifp.fr

ABSTRACT

Redundancy in wavelets and filter banks has the potential to greatly improve signal and image denoising. Having developed a framework for optimized oversampled complex lapped transforms, we propose their association with the statistically efficient Stein's principle in the context of mean square error estimation. Under Gaussian noise assumptions, expectations involving the (unknown) original data are expressed using the observation only. Two forms of Stein's Unbiased Risk Estimators, derived in the coefficient and the spatial domain respectively, are proposed, the latter being more computationally expensive. These estimators are then employed for denoising with linear combinations of elementary threshold functions. Their performances are compared to the oracle, and addressed with respect to the redundancy. They are finally tested against other denoising algorithms. They prove competitive, yielding especially good results for texture preservation.

1. INTRODUCTION

Estimation of unknown data from noisy observations is a central problem in statistics. Stein's principle [1] allowed for instance a remarkable improvement of the standard estimator of a multivariate normal mean by shrinking the standard estimator towards zero [2]. This principle demonstrated its usefulness in signal and image processing in a wavelet transformed domain, leading to Stein's Unbiased Risk Estimator (SURE) wavelet shrinkage (SureShrink) proposed by Donoho and Johnstone [3]. This approach does not require *a priori* knowledge of the underlying signal: assumptions only lie on the noise. It additionally benefits from the wavelet concentration property, which tends to better represent structured signal out of a noisy environment. Since then, Stein's principle has been exploited in more involved settings [4]. The relatively simple yet effective soft or hard thresholding from earlier works have been improved for instance in [5, 6]with a linear parameterization of thresholds using elementary functions. Recently, Blu and Luisier [7] associated redundant wavelet transforms and the SURE principle to such a linear expansion of thresholds, leading to a method minimizing the risk in the image domain instead of the transform domain. For applications where textures are especially important, it may be more effective to resort to transformations that better preserve high frequency directional details. An example of such a transform was used in a previous work [8], in which we proposed a practical way to compute optimized synthesis filter banks (FBs) associated with a complex analysis FB.

The contributions of this work are twofold. First, we extend the results by Blu and Luisier to the case of oversampled *complex* transforms for optimized complex FB denoising. Second, we compare the efficiency of the SURE-LET approach when applied in the transform or the image domain

In Section 2, we first recall Stein's principle and introduce some notations. In Sections 3 and 4, we devise two estimators working on transformed coefficients and on reconstructed samples. Section 5 describes the SURE-LET approach used to minimize those estimators. Finally, we explain their implementation for image denoising and compare in Section 7 the proposed methods to their oracles as well as various denoising algorithms.

2. NOTATIONS AND STEIN'S PRINCIPLE

2.1 Notations

Throughout this paper, we consider a discrete-time signal $\mathbf{x}$ of length L corrupted by a zero-mean Gaussian white noise $\mathbf{b}$ of variance σ^2 and independent of $\mathbf{x}$. The observed signal is denoted $\mathbf{y} = \mathbf{x} + \mathbf{b}$ (*i.e.* $y_n = x_n + b_n$, for all $n \in \{1,...,L\}$). We suppose that a linear complex transform is used to analyze this signal. Let L' be the length of the transformed signal. $\boldsymbol{D} \in \mathbb{C}^{L' \times L}$ represents the transform matrix. We also suppose that a linear synthesis operator, represented by a matrix $\boldsymbol{R} \in \mathbb{C}^{L \times L'}$, exists and achieves perfect reconstruction, *i.e.* $\boldsymbol{RD} = \boldsymbol{I}_L$. Let $\mathbf{Y} = \boldsymbol{D}\mathbf{y}$, $\mathbf{X} = \boldsymbol{D}\mathbf{x}$ and $\mathbf{B} = \boldsymbol{D}\mathbf{b}$ be the transformed vectors, we can then write: $\mathbf{Y} = (Y_1,...,Y_{L'})^\top = \mathbf{X} + \mathbf{B}$.

Let $\mathscr{C}_{d,\boldsymbol{\Gamma}}$ be the class of all functions $\mathbf{T} : \mathbb{R}^d \to \mathbb{R}^d$, continuous, almost everywhere differentiable such that:

1. $\mathsf{E}\left(\|\mathbf{T}(\mathbf{w})\|^2\right) < \infty$,

2. $\mathsf{E}\left(\left\|\dfrac{\partial \mathbf{T}(\mathbf{w})}{\partial \mathbf{w}^\top}\right\|_F\right) < \infty$,

3. $\forall \mathbf{z} \in \mathbb{R}^d, \displaystyle\lim_{\|\mathbf{t}\| \to +\infty} \mathbf{T}(\mathbf{t}) \exp\left(-\dfrac{(\mathbf{t}-\mathbf{z})^\top \boldsymbol{\Gamma}^{-1}(\mathbf{t}-\mathbf{z})}{2}\right) = 0$,

where $\|.\|_F$ represents the Frobenius norm and $\mathsf{E}(\cdot)$ designates the mathematical expectation.

Finally, $\mathrm{Tr}(A)$ represents the trace of A, A^R (resp. A^I) the matrix of the real (resp. imaginary) parts of A and $\mathrm{diag}(A)$ the vector constructed with the diagonal elements of A.

2.2 Fundamental lemma

Lemma 1 *Let v be a vector of length d perturbed by a Gaussian noise vector n with zero-mean and covariance matrix $\Gamma^{(n)}$. The observed signal is $w = v + n$. If a function $T : \mathbb{R}^d \to \mathbb{R}^d$ belongs to $\mathscr{C}_{d,\Gamma^{(n)}}$, then:*

$$\mathsf{E}(T(w)v^\top) = \mathsf{E}(T(w)w^\top) - \mathsf{E}\left(\frac{\partial T(w)}{\partial w^\top}\right)\Gamma^{(n)}.$$

This lemma, expressed in [1], is known as Stein's principle.

Interestingly, this lemma allows us to express a cross-correlation matrix involving the original signal only through expectations involving the observed signal. This property is the cornerstone of the results given in Sections 3 and 4.

3. SURE ESTIMATION ON THE TRANSFORMED COEFFICIENTS

In this section, we study a first form of Stein's estimation by working on the transformed coefficients. More precisely, the goal is to build an estimator $\mathbf{F}(\mathbf{Y})$ of $\mathbf{X}$ such that the error on the coefficients: $\mathsf{E}\left(\|\mathbf{F}(\mathbf{Y}) - \mathbf{X}\|^2\right)$ is minimal. To this end, we first use Stein's principle to estimate this error and then in Section 5 we explain the minimization strategy.

3.1 Representation of transformed coefficients

We first define the L' vectors of $\mathbb{R}^2$ representing the complex coefficients after decomposition by: for all $j \in \{1,...,L'\}$,

$$\mathbf{Y}_j = \begin{pmatrix} (Y_j)^R \\ (Y_j)^I \end{pmatrix} = \begin{pmatrix} \mathbf{D}_j^R \mathbf{y} \\ \mathbf{D}_j^I \mathbf{y} \end{pmatrix}, \text{ where } D = \begin{pmatrix} \mathbf{D}_1 \\ \vdots \\ \mathbf{D}_{L'} \end{pmatrix}.$$

We suppose that a thresholding function Θ (generally non-linear) is used on the coefficients. This operator is applied *pointwise*, in other words, it can be written: $\Theta(\mathbf{w}) = (\vartheta_j(w_j))_{1 \le j \le L'}$, for all $\mathbf{w} = (w_j)_{1 \le j \le L'} \in \mathbb{C}^{L'}$. For all $j \in \{1,...,L'\}$, ϑ_j is a scalar function from $\mathbb{C}$ to $\mathbb{C}$. More complex operators can be considered [9], but in the scope of this paper we have considered scalar operators that lead to a simple minimization process as seen in Section 5.

We also define the functions $\theta_j : \mathbb{R}^2 \to \mathbb{C}$, such that: $\vartheta_j(w_j) = \theta_j(w_j^R, w_j^I)$ for all $j \in \{1,...,L'\}$. For all $j \in \{1,...,L'\}$, we introduce:

$$\Theta_j(\mathbf{Y}_j) = \begin{pmatrix} \theta_j^R(\mathbf{Y}_j) \\ \theta_j^I(\mathbf{Y}_j) \end{pmatrix}.$$

With these notations, the quadratic error can be written:

$$\|\mathbf{F}(\mathbf{Y}) - \mathbf{X}\|^2 = \sum_{j=1}^{L'} \left\| \Theta_j(\mathbf{Y}_j) - \mathbf{X}_j \right\|_{\mathbb{R}^2}^2.$$

Finally, following the same notations, we introduce the L' noise vectors: $\mathbf{B}_j = \begin{pmatrix} \mathbf{D}_j^R \mathbf{b} \\ \mathbf{D}_j^I \mathbf{b} \end{pmatrix}$. Each vector $\mathbf{B}_j$ represents a

2×1 Gaussian random vector with zero mean and covariance matrix Γ_j, with $j \in \{1,...,L'\}$, defined as:

$$\Gamma_j = \sigma^2 \begin{pmatrix} \mathbf{D}_j^R \mathbf{D}_j^{R\top} & \mathbf{D}_j^R \mathbf{D}_j^{I\top} \\ \mathbf{D}_j^I \mathbf{D}_j^{R\top} & \mathbf{D}_j^I \mathbf{D}_j^{I\top} \end{pmatrix}.$$

3.2 Estimator expression

The following result uses Stein's principle to build an estimator of the error $\mathsf{E}\left(\|\mathbf{F}(\mathbf{Y}) - \mathbf{X}\|^2\right)$.

Proposition 1 *If for all $j \in \{1,...,L'\}$ the function $\Theta_j : \mathbb{R}^2 \to \mathbb{R}^2$ belongs to $\mathscr{C}_{2,\Gamma_j}$, then*

$$\varepsilon_c = \sum_{j=1}^{L'} \left(\left\| \Theta_j(\mathbf{Y}_j) - \mathbf{Y}_j \right\|_{\mathbb{R}^2}^2 + 2\mathrm{Tr}\left(\frac{\partial \Theta_j(\mathbf{Y}_j)}{\partial \mathbf{Y}_j^\top} \Gamma_j \right) - \mathrm{Tr}(\Gamma_j) \right),$$

is an unbiased estimator of $\mathsf{E}\left(\|\mathbf{F}(\mathbf{Y}) - \mathbf{X}\|^2\right)$.

In contrast with the case of a real transform, we have to resort here to a bivariate estimator to approximate the error.

4. SURE ESTIMATION ON THE RECONSTRUCTED SAMPLES

4.1 Estimator expression

We now aim at proposing a function $\mathbf{G} : \mathbb{R}^L \to \mathbb{R}^L$ of $\mathbf{y}$, where $\mathbf{G}(\mathbf{y}) = (g_n(\mathbf{y}))_{1 \le n \le L}$, to estimate the signal $\mathbf{x}$ by minimizing the mean square error: $\mathrm{MSE} = \mathsf{E}\left(\|\mathbf{G}(\mathbf{y}) - \mathbf{x}\|^2\right)$.

The difference with the study in the previous section is that we work now on the data after reconstruction. Once again, by applying Lemma 1, we can get the following statistical (proved in [7]).

Proposition 2 *If $G : \mathbb{R}^L \to \mathbb{R}^L$ belongs to $\mathscr{C}_{L,\sigma^2 I_L}$, then*

$$\varepsilon_s = \|G(y) - y\|^2 + 2\sigma^2 \mathrm{div}(G(y)) - L\sigma^2$$

is an unbiased estimator of the MSE.

4.2 Expression of divergence

In Proposition 2, we see that the construction of the estimator involves a divergence term. In the complex case, we reexpress this term in a more convenient way.

The function $\mathbf{G}$ can be written in this case as $\mathbf{G} = R\Theta(\mathbf{D})$ where Θ is a pointwise thresholding operator: $\Theta = (\vartheta_\ell(w_\ell))_{1 \le \ell \le L'}$ with $w_\ell \in \mathbb{C}$. We introduce the following pointwise thresholding function on $\mathbb{R}^{L'} \times \mathbb{R}^{L'}$:

$$\underline{\Theta}(\mathbf{w}^R, \mathbf{w}^I) = \left(\theta_\ell(w_\ell^R, w_\ell^I)\right)_{1 \le \ell \le L'} = \left(\vartheta_\ell(w_\ell^R + \iota w_\ell^I)\right)_{1 \le \ell \le L'}.$$

The components of $\mathbf{G}(\mathbf{y})$ can be written, for all $n \in \{1,\ldots,L\}$ and $\mathbf{y} \in \mathbb{R}^L$: $g_n(\mathbf{y}) = \sum_{\ell=1}^{L'} R_{n,\ell}\theta_\ell(Y_\ell^R, Y_\ell^I)$.

Moreover, by using the definition $Y_\ell = \sum_{m=1}^{L} D_{\ell,m}y_m$, we deduce that for all $n \in \{1,\ldots,L\}$:

$$\left(\frac{\partial Y_\ell^R}{\partial y_n}, \frac{\partial Y_\ell^I}{\partial y_n}\right) = \left(D_{\ell,n}^R, D_{\ell,n}^I\right).$$

With these equations, we can express the divergence as:

$$\text{div}(\mathbf{G}(\mathbf{y})) = \sum_{n=1}^{L} \frac{\partial g_n(\mathbf{y})}{\partial y_n} = \sum_{n=1}^{L} \sum_{\ell=1}^{L'} R_{n,\ell} \frac{\partial \theta_\ell(\mathbf{D}_\ell^R, \mathbf{D}_\ell^I)(\mathbf{y})}{\partial y_n}$$

$$= \sum_{n=1}^{L} \sum_{\ell=1}^{L'} R_{n,\ell} \left(\nabla \theta_\ell(Y_\ell^R, Y_\ell^I)\right)^\top \left(\frac{\partial Y_\ell^R}{\partial y_n}, \frac{\partial Y_\ell^I}{\partial y_n}\right)^\top$$

$$= \sum_{\ell=1}^{L'} \left(\frac{\partial \theta_\ell(Y_\ell^R, Y_\ell^I)}{\partial x_1} \left(\boldsymbol{D}^R \boldsymbol{R}\right)_{\ell,\ell} + \frac{\partial \theta_\ell(Y_\ell^R, Y_\ell^I)}{\partial x_2} \left(\boldsymbol{D}^I \boldsymbol{R}\right)_{\ell,\ell} \right)$$

$$= \text{diag}(\boldsymbol{D}^R \boldsymbol{R})^\top \underline{\Theta}_{x_1}(\mathbf{Y}^R, \mathbf{Y}^I) + \text{diag}(\boldsymbol{D}^I \boldsymbol{R})^\top \underline{\Theta}_{x_2}(\mathbf{Y}^R, \mathbf{Y}^I),$$

where $\underline{\Theta}_{x_j}(\mathbf{Y}^R, \mathbf{Y}^I) = \left(\frac{\partial \theta_\ell(Y_\ell^R, Y_\ell^I)}{\partial x_j} \right)_{1 \leq \ell \leq L'}$, for all $j \in \{1, 2\}$.

This result is an extension to the complex case of the one derived in [7].

5. DENOISING WITH SURE-LET METHOD

5.1 Principle

We now consider that the thresholding functions $\mathbf{F}$ and $\mathbf{G}$ are linear combinations of K elementary thresholding functions $\mathbf{F}^{(k)}$ and $\mathbf{G}^{(k)}$ weighted by a vector of parameters $\boldsymbol{a} \in \mathbb{R}^K$ [5, 10]. This combination has been called a *Linear Expansion of Thresholds* in [7]. The main interest of this approach is that the design of the estimators minimizing ε_c and ε_s can be achieved in a straightforward way.

The method consists in finding a vector $\boldsymbol{a}$ minimizing $\varepsilon_c = J_c(\boldsymbol{a})$ or $\varepsilon_s = J_s(\boldsymbol{a})$. These functions are convex quadratic in the parameter vector $\boldsymbol{a}$, thus the solutions of the minimization problems are obtained by solving a linear system. We now provide the expressions of this system for each method.

5.2 Minimization of ε_c

Considering the thresholding function $\mathbf{F}(\mathbf{Y}) = \sum_{k=1}^{K} a_k \mathbf{F}^{(k)}(\mathbf{Y})$, we can write:

$$\Theta_j(\mathbf{Y}_j) = \sum_{k=1}^{K} a_k \Theta_j^{(k)}(\mathbf{Y}_j),$$

where for all $j \in \{1, ..., L'\}$ and $k \in \{1, ..., K\}$, the elementary pointwise thresholding functions verify: $\Theta_j^{(k)} \in \mathscr{C}_{2,\Gamma_j}$. By imposing $\frac{\partial J_c(\mathbf{a})}{\partial a_k} = 0$, we have to solve the following system: $\boldsymbol{M}\mathbf{a} = \mathbf{c}$ where for all $(k, \ell) \in \{1, \dots, K\}^2$:

$$M_{k,\ell} = \sum_{j=1}^{L'} \left(\Theta_j^{(\ell)}(\mathbf{Y}_j) \right)^\top \Theta_j^{(k)}(\mathbf{Y}_j),$$

$$c_k = \mathbf{Y}_j^\top \Theta_j^{(k)}(\mathbf{Y}_j) - \sum_{j=1}^{L'} \text{Tr} \left(\frac{\partial \Theta_j^{(k)}(\mathbf{Y}_j)}{\partial \mathbf{Y}_j^\top} \Gamma_j \right).$$

To prevent bad conditioning of $\boldsymbol{M}$ issues, and thus possible numerical problems, this system is solved by calculating the Moore-Penrose pseudo-inverse of $\boldsymbol{M}$.

5.3 Minimization of ε_s

The thresholding function considered in this case is:

$$\mathbf{G}(\mathbf{y}) = \sum_{k=1}^{K} a_k \mathbf{G}^{(k)}(\mathbf{y}) = \sum_{k=1}^{K} a_k \boldsymbol{R} \Theta^{(k)}(\boldsymbol{D}\mathbf{y}),$$

where $\Theta^{(k)} : \mathbb{C}^{L'} \to \mathbb{C}^{L'}$. In a similar way to Section 4.2 we define an equivalent thresholding function: $\underline{\Theta}^{(k)}(\mathbf{w}^R, \mathbf{w}^I) = \Theta^{(k)}(\mathbf{w})$.

Once again, by setting $\frac{\partial J_s(\mathbf{a})}{\partial a_k} = 0$, the minimization is performed by solving the linear system: $\widetilde{\boldsymbol{M}}\mathbf{a} = \widetilde{\mathbf{c}}$ where for all $(k, \ell) \in \{1, \dots, K\}^2$:

$$\widetilde{M}_{k,\ell} = \mathbf{G}^{(\ell)}(\mathbf{y})^\top \mathbf{G}^{(k)}(\mathbf{y}) \text{ and } \widetilde{c}_k = \mathbf{G}^{(k)}(\mathbf{y})^\top \mathbf{y} - \sigma^2 \text{div}(\mathbf{G}^{(k)}(\mathbf{y})).$$

As in the previous section, this system is solved with a pseudo-inversion.

6. IMPLEMENTATION DETAILS

6.1 Extension to two dimensions

In this section, we apply the SURE-LET methods on images. Let $\mathscr{X}$ be an $L \times L$ image, corrupted by a zero-mean white Gaussian noise $\mathscr{B}$ with variance σ^2. The noisy observed image is: $\mathscr{Y} = \mathscr{X} + \mathscr{B}$. To apply the results of the previous sections, we consider the vector obtained by using the column stacking operation: $\mathbf{y}_{i_1 + (i_2-1)L} = \mathscr{Y}_{i_1, i_2}$ for all $(i_1, i_2) \in \{1, \dots, L\}^2$.

Assuming the 2D transform separable and using the notations introduced for the 1D case, we can write the decomposition as follows:

$$\boldsymbol{D}\mathscr{Y}\boldsymbol{D}^\top = \left(\sum_{i_1=1}^{L} \sum_{i_2=1}^{L} D_{n_1, i_1} \mathscr{Y}_{i_1, i_2} D_{n_2, i_2} \right)_{1 \leq n_1, n_2 \leq L'},$$

which can be expressed in the equivalent form:

$$\mathbf{Y}_n = \sum_{i=1}^{L^2} \mathscr{D}_{n,i} \mathbf{y}_i,$$

with $\mathscr{D}_{n_1 + (n_2-1)L', i_1 + (i_2-1)L} = D_{n_1, i_1} D_{n_2, i_2}$, for all $(n_1, n_2) \in \{1, \dots, L'\}^2$ and $(i_1, i_2) \in \{1, \dots, L\}^2$. Similarly, the synthesis matrix reads: $\mathscr{R}_{i_1 + (i_2-1)L, n_1 + (n_2-1)L'} = R_{i_1, n_1} R_{i_2, n_2}$, for all $(n_1, n_2) \in \{1, \dots, L'\}^2$ and $(i_1, i_2) \in \{1, \dots, L\}^2$.

Since $\mathscr{R}\mathscr{D} = \boldsymbol{I}_{L^2}$ (by construction), we directly apply the SURE-LET methods described in the previous sections.

6.2 Thresholding operators

The following two scalar complex-valued functions are used to build the thresholding operators:

$$\tau_1(x, y) = x + \iota y,$$

$$\tau_2(x, y) = (x + \iota y) \left(1 - e^{-\frac{(x^2 + y^2)}{(\alpha \sigma)^2}} \right),$$

with $\alpha \in \mathbb{R}$. Function τ_2 can be seen as a smoothed version of a hard thresholding. The constant α is set to $3/2$ in this

work (in an empirical manner) but it should be adapted to the chosen transform. The resulting scalar threshold function is the linear combination of the two previous ones:

$$\vartheta_k(x+\iota y) = \theta_k(x,y) = a_{1,k}\tau_1(x,y) + a_{2,k}\tau_2(x,y), \ \forall(x,y) \in \mathbb{R}^2.$$

To properly define the functions $\mathbf{F}^{(k)}$ and $\mathbf{G}^{(k)}$, we must define the subbands on which they are applied. The transforms used in this paper, as described in [8], are complex oversampled filter banks based on a generalized Fourier transform. To ensure that the reconstructed image is real-valued, we similarly process subbands in a symmetric way. Therefore, each $\mathbf{F}^{(k)}$ and $\mathbf{G}^{(k)}$ simultaneously processes two symmetric subbands.

7. SIMULATIONS

We denote by FB-SURE-LET-C and FB-SURE-LET-S the denoising methods based on oversampled filter banks and SURE-LET in the **c**oefficient (as in Section 5.2) and **s**ample (as in Section 5.3) domain, respectively. The analysis and the associated optimized synthesis FBs are as in [8, 11].

7.1 FB-SURE-LET-C and FB-SURE-LET-S vs. oracle

The analysis FB is parametrized here by: $k = 3$ (overlapping factor), $k' = 3$ (redundancy) and $N = 4$ (downsampling factor). Two versions of the Lena image of size 256×256 and 512×512 were considered both corrupted by a zero-mean Gaussian white noise with standard deviation $\sigma = 20$. We first compare the performance of both algorithms with their oracle. It is obtained by minimizing either $\phi_c(a_1,...,a_K) = \left\| \sum_{k=1}^{K} a_k \mathbf{F}^{(k)}(\mathbf{Y}) - \mathbf{X} \right\|$ or $\phi_s(a'_1,...,a'_K) = \left\| \sum_{k=1}^{K} a'_k \mathbf{G}^{(k)}(\mathbf{y}) - \mathbf{x} \right\|$, for the FB-SURE-LET-C or FB-SURE-LET-S, respectively. For each case, we compute the median PSNR from 20 realizations. The results are reported in Table 1. Performances of FB-SURE-LET methods stand very close to those obtained with the oracle, by -0.4 dB and -0.2 dB in the coefficient and the sample domain, thus showing the effectiveness of Stein's principle. Comparing results with increasing image sizes indicates that with a larger image the difference between oracles and FB-SURE-LET methods decreases, which may be explained by the fact that the estimation becomes more consistent with more samples.

7.2 FB-SURE-LET vs. filter bank redundancy

In this section, we compare the proposed methods for different FBs of varying redundancy. The analysis FBs share the same overlapping and downsampling factors: $k = 3$ and $N = 8$. The redundancy k' varies as indicated in Table 2. Its increase allows to consistently improve the denoising performance for the Lena image (of size 256×256) corrupted by a white Gaussian noise with $\sigma = 30$, sometimes drastically in the FB-SURE-LET-C case. Again, the sample domain denoising method always better performs than the coefficient-based one. With the chosen FBs we observe that the gap in SNR between the two methods decreases as the redundancy increases. This result may appear surprising as the gap could be expected to increase as we depart from the orthogonal case. This behavior can be explained by the better quality of the synthesis FBs we use for higher redundancy factors. FB-SURE-LET-C is meanwhile much more cost effective,

since the optimization requires only one reconstruction, while FB-SURE-LET-S implies the computation of K reconstructions, K being the number of parameters to be estimated, *i.e.* the length of vector $\boldsymbol{a}$. To illustrate this difference, we have chosen an analysis FB parametrized by: $k = k' = 3$ and $N = 4$ and its associated optimized synthesis FB. On a 256×256 image, FB-SURE-LET-S is computed in 63 seconds against 3 seconds for FB-SURE-LET-C[1].

	Lena 256×256	*Lena* 512×512
Noisy image	22.1	22.1
FB-SURE-LET-S	30.0	31.3
Oracle-S	30.2	31.3
FB-SURE-LET-C	29.2	30.5
Oracle-C	29.6	30.7

Table 1: Denoising results (in PSNR) using FB-SURE-LET-S and FB-SURE-LET-C as well as their respective oracles for the Lena image with a $\sigma = 20$ noise.

Redund. k'	5/4	3/2	7/4	2
FB-SURE-LET-C	16.7	24.3	25.5	25.9
FB-SURE-LET-S	26.0	26.7	26.9	27.0
Difference	9.3	2.4	1.4	1.1
Redund. k'	9/4	5/2	11/4	3
FB-SURE-LET-C	26.0	25.7	25.8	27.1
FB-SURE-LET-S	27.2	26.8	26.9	27.4
Difference	1.2	1.1	1.1	0.3

Table 2: PSNR in dB after reconstruction with FB-SURE-LET-S and FB-SURE-LET-C of Lena image with $\sigma = 30$, FB parameters: $N = 8$, $k = 3$ and a varying redundancy.

7.3 FB-SURE-LET vs. other denoising methods

We compare the proposed methods on several standard test images of size 256×256 with four other denoising algorithms, namely *Curvelets*[2] (redundancy: ~ 7.3), *SureShrink CS* [3] with cycle-spinning (redundancy: $3j_{\max} + 1 = 13$ where $j_{\max} = 4$ is the decomposition level), bivariate shrinkage [12] (*BiShrink*)[3] and undecimated wavelets SURE-LET: *UWT* SURE-LET (same redundancy: 13) [7]. We chose $N = 4$ and $k' = 3$ (thus the redundancy is: 9).

Table 3 reports the median PSNR on 20 realizations. FB-SURE-LET-S favorably compares with *UWT* SURE-LET and often outperforms the other algorithms. As earlier, the performance of FB-SURE-LET-C is somewhat weaker. Still, it generally yields results comparable with the other methods, except for the higher noise levels.

For $\sigma = 30$, Figure 1 provides a cropped version of Barbara image. The oriented textures from the scarf are much better preserved by FB-SURE-LET-S, due to the nature of the complex filter banks employed here, well-adapted to fine directional structure preservation.

---

[1] Using Matlab 7 on a computer with an Intel Core 2 T7400 2.16GHz CPU and 2Gb of RAM.
[2] Toolbox at `http://www.curvelet.org`
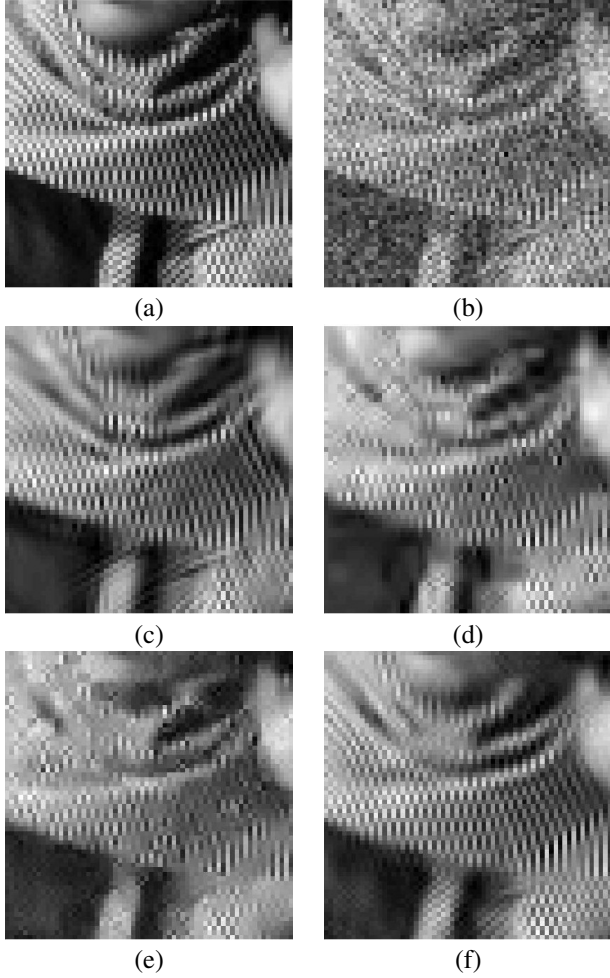[3] Toolbox at `http://taco.poly.edu/WaveletSoftware`

Figure 1: Detail on Barbara image: (a) original image, (b) noisy image ($\sigma = 30$), denoised image using (c) *Curvelets*, (d) *BiShrink*, (e) *UWT* SURE-LET and (f) FB-SURE-LET-S.

σ	10	20	30	40	50
Noisy	28.1	22.1	18.6	16.1	14.2
Lena					
Curvelets	33.0	29.3	27.3	25.9	24.8
SureShrink CS	32.6	28.9	26.4	25.5	24.0
BiShrink	32.8	29.0	26.9	25.5	24.6
UWT SURE-LET	33.0	29.4	27.4	26.1	25.2
FB-SURE-LET-C	33.0	29.3	27.1	25.5	24.3
FB-SURE-LET-S	**33.8**	**30.0**	**28.0**	**26.4**	**25.3**
Barbara					
Curvelets	31.2	27.2	25.1	23.6	22.6
SureShrink CS	30.7	26.5	24.1	21.9	21.4
BiShrink	30.1	26.6	24.4	23.1	22.2
UWT SURE-LET	30.8	26.4	24.3	23.2	22.4
FB-SURE-LET-C	31.6	27.6	25.4	23.7	22.3
FB-SURE-LET-S	**32.2**	**28.3**	**26.2**	**24.7**	**23.6**
Boat					
Curvelets	31.7	28.1	26.1	24.9	24.0
SureShrink CS	32.0	27.9	25.8	24.4	23.3
BiShrink	31.8	27.9	25.9	24.7	23.8
UWT SURE-LET	32.4	28.6	26.6	**25.3**	**24.5**
FB-SURE-LET-C	32.0	28.1	26.0	24.6	23.4
FB-SURE-LET-S	**32.6**	**28.7**	**26.7**	**25.3**	24.3

Table 3: PSNR obtained with FB-SURE-LET methods and other recent denoising methods on images Lena, Barbara and Boat for varying σ.

8. CONCLUSION

Two forms of Stein's Unbiased Risk Estimators (SURE) criteria have been derived for denoising with oversampled complex filter banks. Denoising results are close to what the oracles predict thus showing the robustness of the FB-SURE-LET approach. Results are slightly better than those obtained with other recent denoising methods, especially on textured areas of images. In a future work, we plan to investigate other thresholding functions, and study more precisely how to choose an adapted scaling parameter α.

REFERENCES

[1] C. Stein, "Estimation of the mean of a multivariate normal distribution," *Annals of Statistics*, vol. 9, no. 6, pp. 1135–1151, 1981.

[2] W. James and C. Stein, "Estimation with quadratic loss," *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, vol. 1, pp. 311–319, 1961.

[3] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. American Statist. Ass.*, vol. 90, pp. 1200–1224, Dec. 1995.

[4] A. Benazza-Benyahia and J.-C. Pesquet, "Building robust wavelet estimators for multicomponent images using Stein's principle," *IEEE Trans. on Image Proc.*, vol. 14, pp. 1814–1830, Nov. 2005.

[5] J.-C. Pesquet and D. Leporini, "A new wavelet estimator for image denoising," in *IEE Sixth Int. Conf. Im. Proc. Appl.*, vol. 1, pp. 249–253, Jul. 14-17 1997.

[6] M. Raphan and E. P. Simoncelli, "Optimal denoising in redundant bases," in *Proc. Int. Conf. on Image Processing*, vol. III, (San Antonio, TX, USA), pp. 113–116, Sep. 16-19 2007.

[7] T. Blu and F. Luisier, "The SURE-LET approach to image denoising," *IEEE Trans. on Image Proc.*, vol. 16, pp. 2778–2786, Nov. 2007.

[8] J. Gauthier, L. Duval, and J. C. Pesquet, "Oversampled inverse complex lapped transform optimization," in *Proc. Int. Conf. on Acoust., Speech and Sig. Proc.*, vol. 1, pp. 549–552, April 2007.

[9] C. Chaux, L. Duval, A. Benazza-Benyahia, and J.-C. Pesquet, "A nonlinear Stein based estimator for multichannel image denoising," *IEEE Trans. on Signal Proc.*, 2008. To appear.

[10] M. Raphan and E. P. Simoncelli, "Learning to be Bayesian without supervision," in *Proc. Neural Information Processing Systems*, (Vancouver, BC, Canada), Dec. 4-7 2006. Published as Advances in Neural Information Processing Systems, eds. B. Schlkopf, J. Platt and T. Hofmann, vol. 19, May 2007.

[11] J. Gauthier, L. Duval, and J.-C. Pesquet, "Inversion and optimization of oversampled complex filter banks: application to directional filtering of images," 2008. Submitted.

[12] L. Şendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. on Signal Proc.*, vol. 50, pp. 2744–2756, Nov. 2002.