A NEW ESTIMATOR FOR IMAGE DENOISING USING A 2D DUAL-TREE M-BAND WAVELET DECOMPOSITION

Caroline Chaux[1], Laurent Duval[2], Amel Benazza-Benyahia[3], and Jean-Christophe Pesquet[1]

[1] IGM and UMR-CNRS 8049
Université de Marne-la-Vallée,
Champs-sur-Marne
77454 Marne-la-Vallée, France
e-mail: {chaux,pesquet}@univ-mlv.fr

[2] Institut Français du Pétrole
Technology, Computer Science
and Applied Mathematics Department
92500 Rueil Malmaison, France
e-mail : laurent.duval@ifp.fr

[3] URISA, Ecole Supérieure des
Communications de Tunis
Route de Raoued 3.5 Km
2083 Ariana, Tunisia
e-mail : ben.yahia@planet.tn

ABSTRACT

We propose a new estimator for image denoising using a 2D dual-tree M-band wavelet transform. Our work extends existing block-based wavelet thresholding methods by exploiting simultaneously coefficients in the two M-band wavelet trees. The contributions of this paper are two-fold. Firstly, we perform a statistical analysis of the noise in the considered redundant decomposition. Secondly, we propose an efficient method to remove the noise. Our approach relies on an extension of Stein's formula which allows us to take into account the specific correlations of the noise components. Simulation results are then presented to validate the proposed method.

1. INTRODUCTION

Wavelet shrinkage has become an efficient method for image denoising. It consists in projecting discrete data onto a basis (usually an orthonormal one) and applying a nonlinear operator to the transformed coefficients. A simple thresholding rule is often used as a nonlinear estimator. The denoised signal is then recovered by the inverse transform. Discrete wavelet transforms (DWT) possess good decorrelation properties and provide sparse representations for a variety of regular images. For the Visushrink and the SUREshrink methods, Donoho *et al.* have derived optimal scalar thresholds [1]. Many improvements on scalar thresholding have been investigated subsequently, such as block-thresholding which accounts for local dependence between neighboring wavelet coefficients [2]. From the transform choice viewpoint, the DWT is maximally decimated, which hampers its robustness to signal shifts. Undecimated wavelets or more general overcomplete expansions have thus been proposed to alleviate some of the wavelet decomposition shortcomings. However, frame decompositions introduce correlations between the signal/noise components which have to be taken into account in the design of the regression rule. The *dual-tree* discrete wavelet transform [3, 4] is one of the most promising frame decompositions due to its reasonable computational cost, limited redundancy and improved selectivity features for image applications. This decomposition is based on two classical DWT operating in parallel, employing an Hilbert pair of mother wavelets. In our recent work [5], we have proposed M-band extensions to the *dual-tree* transform (DTT) and investigated their properties. The objective of this paper is to build a reliable estimator in the M-band DTT domain. The novelty of our approach consists in both analyzing the statistical properties of the noise coefficients and proposing an appropriate block thresholding method for removing the noise. The derivation of this new estimator is based on an extension of Stein's principle. This paper is organized as follows. The main properties of M-band DTTs are briefly recalled in Section 2. In Section 3, explicit expressions for the statistics of the coefficients of a white noise are provided, thus generalizing the results given in [6] for the 1D case. In Section 4, the specific correlation structure of the noise coefficients in an M-band dual-tree wavelet decomposition is exploited in order to build a new adaptive block thresholding estimator. In Section 5, some simulation examples are presented to evaluate the benefits of the proposed denoising method and, finally some concluding remarks are given in Section 6. Throughout the paper, the following notations will be used: let M be an integer greater than or equal to 2, $\mathbb{N}_M = \{0, \ldots, M - 1\}$ and $\mathbb{N}_M^\star = \{1, \ldots, M - 1\}$. Besides, $\widehat{a}$ denotes the Fourier transform of a function a, $(\delta_m)_{m \in \mathbb{Z}}$ is the Kronecker sequence (equal to 1 if $m = 0$ and 0 otherwise) and $(f)_+ = f$ if $f > 0$ and 0 otherwise.

2. M-BAND DUAL-TREE WAVELET ANALYSIS

An M-band multiresolution analysis of $\mathrm{L}^2(\mathbb{R})$ is defined using one scaling function $\psi_0 \in \mathrm{L}^2(\mathbb{R})$ and $(M - 1)$ mother wavelets $\psi_m \in \mathrm{L}^2(\mathbb{R})$, $m \in \mathbb{N}_M^\star$. In the frequency domain, the so-called scaling equations are expressed as:

$$\forall m \in \mathbb{N}_M, \qquad \sqrt{M}\widehat{\psi}_m(M\omega) = H_m(\omega)\widehat{\psi}_0(\omega). \qquad (1)$$

The following para-unitarity conditions must hold: for all $(m, m') \in \mathbb{N}_M^2$,

$$\sum_{p=0}^{M-1} H_m(\omega + p\frac{2\pi}{M})H_{m'}^*(\omega + p\frac{2\pi}{M}) = M\delta_{m-m'} \qquad (2)$$

in order to generate an orthonormal M-band wavelet basis of $\mathrm{L}^2(\mathbb{R})$. The filter with frequency response H_0 is low-pass whereas the filters with frequency response H_m, $m \in \mathbb{N}_M^\star$, are band-pass or high-pass. In this case, cascading M-band para-unitary analysis and synthesis filter banks allows us to decompose and to perfectly reconstruct any 1D signal. A "dual" M-band multiresolution analysis is built by defining another M-band wavelet orthonormal basis associated with a scaling function ψ_0^{H} and mother wavelets ψ_m^{H}, $m \in \mathbb{N}_M^\star$. More precisely, the mother wavelets are the Hilbert transforms of the "primal" ones ψ_m, $m \in \mathbb{N}_M^\star$:

$$\forall m \in \mathbb{N}_M^\star, \qquad \widehat{\psi}_m^{\mathrm{H}}(\omega) = -\imath \, \mathrm{sign}(\omega)\widehat{\psi}_m(\omega) \qquad (3)$$

where $\mathrm{sign}(\cdot)$ is the signum function. Conditions for designing the involved frequency responses G_m, $m \in \mathbb{N}_M$, of the corresponding synthesis para-unitary Hilbert filter bank have been recently provided in [7]. For all $(m, m') \in \mathbb{N}_M^2$, we recall that the deterministic cross-correlation function of the primal and dual wavelets in $\mathrm{L}^2(\mathbb{R})$ is defined as:

$$\forall \tau \in \mathbb{R}, \qquad \gamma_{m,m'}(\tau) = \int_{-\infty}^{\infty} \psi_m(x)\psi_{m'}^{\mathrm{H}}(x-\tau)\,dx. \quad (4)$$

3. NOISE STATISTICAL PROPERTIES

We now aim at analyzing the statistics of the transform coefficients of a real-valued, zero-mean 2D white noise n with spectrum density σ^2. We denote by $(n_{.}(\mathbf{k}))_{\mathbf{k} \in \mathbb{Z}^2}$ the coefficients resulting from a $2D$ separable M-band wavelet decomposition [8] of the noise, in a given subband $(j, \mathbf{m}) \in \mathbb{Z} \times \mathbb{N}_M^2$. To simplify the notations, the indices $(j, \mathbf{m})$ have been dropped. The wavelet coefficients at the output of the dual tree are denoted by $(n_{.}^{\mathrm{H}}(\mathbf{k}))_{\mathbf{k} \in \mathbb{Z}^2}$. We obtain the following expressions of the covariance fields: for all $j \in \mathbb{Z}$, $\mathbf{m} = (m_1, m_2) \in \mathbb{N}_M^2$, $\mathbf{m}' = (m_1', m_2') \in \mathbb{N}_M^2$, $\mathbf{k} = (k_1, k_2) \in \mathbb{Z}^2$ and $\mathbf{k}' = (k_1', k_2') \in \mathbb{Z}^2$,

$$\left. \begin{array}{l} \mathsf{E}\{n_{j,\mathbf{m}}(\mathbf{k})n_{j,\mathbf{m}'}(\mathbf{k}')\} \\ \mathsf{E}\{n_{j,\mathbf{m}}^{\mathrm{H}}(\mathbf{k})n_{j,\mathbf{m}'}^{\mathrm{H}}(\mathbf{k}')\} \end{array} \right\} = \sigma^2 \delta_{m_1-m_1'}\delta_{m_2-m_2'}\delta_{k_1-k_1'}\delta_{k_2-k_2'}$$

$$\mathsf{E}\{n_{j,\mathbf{m}}(\mathbf{k})n_{j,\mathbf{m}'}^{\mathrm{H}}(\mathbf{k}')\} = \sigma^2 \gamma_{m_1,m_1'}(k_1'-k_1)\gamma_{m_2,m_2'}(k_2'-k_2).$$

As a consequence of Eq. (3), we note that, for $\mathbf{m} \neq \mathbf{0}$ and $\mathbf{k} = \mathbf{k}'$, the vector $[n_{.}(\mathbf{k})\quad n_{.}^{\mathrm{H}}(\mathbf{k})]^T$ has uncorrelated components with equal variance.

The above relations impact on the 2×2 linear combination of the primal and dual wavelet coefficients which is usually implemented at the last stage of a dual-tree decomposition. As explained in [7], the main advantage of such a postprocessing is to better capture the directional features in the analyzed image. More precisely, the postprocessing consists of the following unitary transform of the detail coefficients: for all $\mathbf{m} \in \mathbb{N}_M^{\star 2}$,

$$\forall \mathbf{k} \in \mathbb{Z}^2, \qquad w_{.}(\mathbf{k}) = \frac{1}{\sqrt{2}}(n_{.}(\mathbf{k}) + n_{.}^{\mathrm{H}}(\mathbf{k})) \quad (5)$$

$$w_{.}^{\mathrm{H}}(\mathbf{k}) = \frac{1}{\sqrt{2}}(n_{.}(\mathbf{k}) - n_{.}^{\mathrm{H}}(\mathbf{k})). \quad (6)$$

Thus, it is straightforward to compute the covariances of the transformed fields $(w_{.}(\mathbf{k}))_{\mathbf{k} \in \mathbb{Z}^2}$ and $(w_{.}^{\mathrm{H}}(\mathbf{k}))_{\mathbf{k} \in \mathbb{Z}^2}$ of noise coefficients: for all $\mathbf{m} \in \mathbb{N}_M^{\star 2}$ and $(\mathbf{k}, \mathbf{k}') \in \mathbb{Z}^2 \times \mathbb{Z}^2$,

$$\mathsf{E}\{w_{.}(\mathbf{k})w_{.}(\mathbf{k}')\} = \mathsf{E}\{n_{.}(\mathbf{k})n_{.}(\mathbf{k}')\} + \mathsf{E}\{n_{.}(\mathbf{k})n_{.}^{\mathrm{H}}(\mathbf{k}')\} \quad (7)$$

$$\mathsf{E}\{w_{.}^{\mathrm{H}}(\mathbf{k})w_{.}^{\mathrm{H}}(\mathbf{k}')\} = \mathsf{E}\{n_{.}(\mathbf{k})n_{.}(\mathbf{k}')\} - \mathsf{E}\{n_{.}(\mathbf{k})n_{.}^{\mathrm{H}}(\mathbf{k}')\} \quad (8)$$

$$\mathsf{E}\{w_{.}(\mathbf{k})w_{.}^{\mathrm{H}}(\mathbf{k}')\} = 0. \quad (9)$$

It is worth noticing that the post-transform not only improves the directional analysis of the image of interest but it plays an important role w.r.t. the noise statistics. Indeed, it allows to completely decorrelate the two noise coefficient fields obtained for any value of $(j, \mathbf{m})$ such that $\mathbf{m} \in \mathbb{N}_M^{\star 2}$.

4. DUAL-TREE BASED ESTIMATOR

4.1. Stein's formula

Noise statistics will be exploited to derive an adaptive estimator. To this purpose, we firstly state an extended form of Stein's principle [9] which will be useful in the next subsection. Hereafter, the considered random vectors are assumed to be real-valued.

Proposition 1 *Let $B \in \mathbb{N}$, $B > 1$, and*

$$\bar{\mathbf{r}} = \bar{\mathbf{s}} + \bar{\mathbf{n}} \quad (10)$$

where $\bar{\mathbf{n}}$ is a B-dimensional zero-mean Gaussian random vector and $\bar{\mathbf{s}}$ is a B-dimensional random vector which is independent of $\bar{\mathbf{n}}$. These vectors are decomposed as

$$\bar{\mathbf{r}} = \begin{bmatrix} r \\ \tilde{\mathbf{r}} \end{bmatrix}, \qquad \bar{\mathbf{s}} = \begin{bmatrix} s \\ \tilde{\mathbf{s}} \end{bmatrix}, \qquad \bar{\mathbf{n}} = \begin{bmatrix} n \\ \tilde{\mathbf{n}} \end{bmatrix} \quad (11)$$

where r, s and n are scalar random variables. Let $T : \mathbb{R}^B \to \mathbb{R}$ be a continuous, almost everywhere differentiable function satisfying some technical requirements [9]. Then,

$$\mathsf{E}[T(\bar{\mathbf{r}})s] = \mathsf{E}[T(\bar{\mathbf{r}})r] - \mathsf{E}[n^2]\mathsf{E}\Big[\frac{\partial T(\bar{\mathbf{r}})}{\partial r}\Big] - \mathsf{E}\Big[\frac{\partial T(\bar{\mathbf{r}})}{\partial \tilde{\mathbf{r}}^{\top}}\Big]\mathsf{E}[\tilde{\mathbf{n}}n].$$

4.2. Proposed adaptive estimator
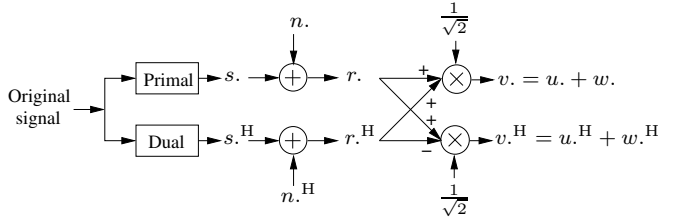
Fig. 1. Considered model.

As illustrated in Fig. 1, we consider the following additive noise model in the DTT domain:

$$r_{.}(\mathbf{k}) = s_{.}(\mathbf{k}) + n_{.}(\mathbf{k}) \quad (12)$$

where $r_{.}(\mathbf{k})$ is the wavelet coefficient of the observed noisy image at a given level j, in a given subband $\mathbf{m}$, at a spatial position $\mathbf{k}$, similar notations being used for the original image and the noise. Inspired by previous works on block thresholding [2, 10], we are interested in applying the following shrinkage function:

$$\eta(\|\bar{\mathbf{r}}_{.}(\mathbf{k})\|) = \left(\frac{\|\bar{\mathbf{r}}_{.}(\mathbf{k})\|^{\beta_{.}} - \lambda_{.}}{\|\bar{\mathbf{r}}_{.}(\mathbf{k})\|^{\beta_{.}}} \right)_{+} \quad (13)$$

where $\lambda_{.}$ and $\beta_{.}$ are positive parameters, $\bar{\mathbf{r}}_{.}(\mathbf{k})$ is a vector containing the wavelet coefficients to be estimated and some possible neighbors. These neighboring values can be taken from the primal tree in the same subband as in [2], or from the dual subband as well. It is important to point out that η includes well known shrinkage rules as particular cases, for specific values of λ and β. We have already applied such a kind of shrinkage to denoise multicomponent images in the conventional DWT domain [11].

As two sets of coefficients are generated by a dual-tree transform,

we aim at designing accurate estimators $\hat{s}_\cdot$ of $s_\cdot$ and $\hat{s}_\cdot^H$ of $s_\cdot^H$ having a common structure:

$$\hat{s}_\cdot(\mathbf{k}) = \eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|)\, r_\cdot(\mathbf{k}), \quad \hat{s}_\cdot^H(\mathbf{k}) = \eta^H(\|\bar{\mathbf{r}}_\cdot^H(\mathbf{k})\|)\, r_\cdot^H(\mathbf{k})$$

where, for the dual tree, $\bar{\mathbf{r}}_\cdot^H(\mathbf{k})$ plays a role symmetric to $\bar{\mathbf{r}}_\cdot(\mathbf{k})$. The shrinkage function η^H has the same form as the η one, making use of parameters $\lambda_\cdot^H$ and $\beta_\cdot^H$ instead of $\lambda_\cdot$ and $\beta_\cdot$. Here, the parameters $\lambda_\cdot$ and $\lambda_\cdot^H$ can be respectively considered as threshold values in the soft-thresholding of $\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{\beta_\cdot}$ and $\|\bar{\mathbf{r}}_\cdot^H(\mathbf{k})\|^{\beta_\cdot^H}$. Since the primal and the dual tree play analogous roles, we only develop the theoretical results for the primal tree. Indeed, expressions concerning the dual tree are easily obtained by replacing any variable g (scalar, vector or matrix) by its dual counterpart g^H. Our next objective is to find the threshold $\lambda_\cdot$ and the exponent $\beta_\cdot$ that minimize the quadratic risk $R(\lambda_\cdot, \beta_\cdot) = \mathsf{E}[|s_\cdot(\mathbf{k}) - \hat{s}_\cdot(\mathbf{k})|^2]$. The risk reads

$$R(\lambda_\cdot, \beta_\cdot) = \mathsf{E}[s_\cdot^2(\mathbf{k})] + \mathsf{E}\big[\big(\eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|)\, r_\cdot(\mathbf{k})\big)^2\big] \quad (14)$$
$$- 2\mathsf{E}[\eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|) r_\cdot(\mathbf{k})\, s_\cdot(\mathbf{k})].$$

Unfortunately, as the wavelet coefficients $s_\cdot(\mathbf{k})$ are unknown, it may appear impossible to calculate explicitly the last term in the expression of $R(\lambda_\cdot, \beta_\cdot)$. However, for an additive Gaussian noise, Prop. 1 applied to our estimator yields

$$\mathsf{E}[\eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|) r_\cdot(\mathbf{k})\, s_\cdot(\mathbf{k})] = \mathsf{E}[\eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|)\, r_\cdot^2(\mathbf{k})]$$

$$-\sigma^2 \mathsf{E}\Big[\frac{\partial\big(\eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|) r_\cdot(\mathbf{k})\big)}{\partial r_\cdot(\mathbf{k})}\Big] - \mathsf{E}\Big[\frac{\partial\big(\eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|) r_\cdot(\mathbf{k})\big)}{\partial \tilde{\mathbf{r}}_\cdot^\top(\mathbf{k})}\Big]\mathbf{\Gamma}^{(\tilde{n}_\cdot, n_\cdot)}$$

where the vector $\bar{\mathbf{r}}_\cdot(\mathbf{k})$ is decomposed as $[r_\cdot(\mathbf{k})\ \tilde{\mathbf{r}}_\cdot^\top(\mathbf{k})]^\top$ and, using similar notations for the noise components, $\mathbf{\Gamma}^{(\tilde{n}_\cdot, n_\cdot)} = \mathsf{E}[\tilde{\mathbf{n}}_\cdot(\mathbf{k}) n_\cdot(\mathbf{k})]$. From (13), we deduce after some calculations that:

$$\frac{\partial(\eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|) r_\cdot(\mathbf{k}))}{\partial r_\cdot(\mathbf{k})} = \eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|) + \lambda_\cdot\, \rho_\cdot(\mathbf{k}) r_\cdot(\mathbf{k}) \quad (15)$$

and

$$\frac{\partial(\eta(\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|) r_\cdot(\mathbf{k}))}{\partial \tilde{\mathbf{r}}_\cdot(\mathbf{k})} = \lambda_\cdot\, \rho_\cdot(\mathbf{k})\, \tilde{\mathbf{r}}_\cdot(\mathbf{k}) \quad (16)$$

where

$$\rho_\cdot(\mathbf{k}) = \mathbf{1}\{\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{\beta_\cdot} > \lambda_\cdot\}\frac{\beta_\cdot}{\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{\beta_\cdot+2}}\, r_\cdot(\mathbf{k}). \quad (17)$$

4.3. Closed form expression of the risk

From the above calculations, the risk $R(\lambda_\cdot, \beta_\cdot)$ can be expressed as the expected value of

$$R(\lambda_\cdot, \beta_\cdot, \mathbf{k}) = a_\cdot(\mathbf{k})\lambda_\cdot^2 + b_\cdot(\mathbf{k})\lambda_\cdot + c_\cdot(\mathbf{k}) \quad (18)$$

where

$$a_\cdot(\mathbf{k}) = \mathbf{1}\{\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{\beta_\cdot} > \lambda_\cdot\}\frac{r_\cdot^2(\mathbf{k})}{\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{2\beta_\cdot}}$$

$$b_\cdot(\mathbf{k}) = 2\sigma^2\Big(\beta_\cdot\, r_\cdot(\mathbf{k})\frac{r_\cdot(\mathbf{k}) + \kappa_\cdot(\mathbf{k})}{\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{\beta_\cdot+2}} - \|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{-\beta_\cdot}\Big)$$

$$\mathbf{1}\{\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{\beta_\cdot} > \lambda_\cdot\}$$

$$c_\cdot(\mathbf{k}) = r_\cdot^2(\mathbf{k}) - \sigma^2 + \mathbf{1}\{\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|^{\beta_\cdot} > \lambda_\cdot\}\big(2\sigma^2 - r_\cdot^2(\mathbf{k})\big)$$

$$\kappa_\cdot(\mathbf{k}) = \sigma^{-2}\tilde{\mathbf{r}}_\cdot^\top(\mathbf{k})\mathbf{\Gamma}^{(\tilde{n}_\cdot, n_\cdot)}.$$

4.4. Computation of the parameters $\lambda_\cdot$ and $\beta_\cdot$

Under mild conditions, $R(\lambda_\cdot, \beta_\cdot)$ is estimated by an empirical average $\hat{R}(\lambda_\cdot, \beta_\cdot)$ computed over the K_j observations in a given subband at resolution j. The optimal values of $\lambda_\cdot$ and $\beta_\cdot$ are found according to a similar procedure to the one used to derive the SUREshrink estimator. To this purpose, the observed variables $\|\bar{\mathbf{r}}_\cdot(\mathbf{k})\|$ are sorted in descending order, so that $\|\bar{\mathbf{r}}_\cdot(\mathbf{k}_1)\| \geq \|\bar{\mathbf{r}}_\cdot(\mathbf{k}_2)\| \geq \ldots \geq \|\bar{\mathbf{r}}_\cdot(\mathbf{k}_{K_j})\|$.
For $i_0 \in \{2, \ldots, K_j\}$, if $\|\bar{\mathbf{r}}_\cdot(\mathbf{k}_{i_0-1})\|^{\beta_\cdot} > \lambda_\cdot \geq \|\bar{\mathbf{r}}_\cdot(\mathbf{k}_{i_0})\|^{\beta_\cdot}$, the risk estimate can be expressed as

$$\hat{R}(\lambda_\cdot, \beta_\cdot) = \frac{1}{K_j}\Big(\sum_{i=1}^{i_0-1} R(\lambda_\cdot, \beta_\cdot, \mathbf{k}_i) + \sum_{i=i_0}^{K_j} R(\lambda_\cdot, \beta_\cdot, \mathbf{k}_i)\Big),$$

or equivalently as

$$K_j\, \hat{R}(\lambda_\cdot, \beta_\cdot) = \lambda_\cdot^2 \sum_{i=1}^{i_0-1} \frac{r_\cdot^2(\mathbf{k}_i)}{\|\bar{\mathbf{r}}_\cdot(\mathbf{k}_i)\|^{2\beta_\cdot}}$$

$$+ 2\sigma^2\lambda_\cdot \sum_{i=1}^{i_0-1}\Big(\beta_\cdot\, r_\cdot(\mathbf{k}_i)\frac{r_\cdot(\mathbf{k}_i) + \kappa_\cdot(\mathbf{k}_i)}{\|\bar{\mathbf{r}}_\cdot(\mathbf{k}_i)\|^{\beta_\cdot+2}} - \|\bar{\mathbf{r}}_\cdot(\mathbf{k}_i)\|^{-\beta_\cdot}\Big)$$

$$+ \sum_{i=i_0}^{K} r_\cdot^2(\mathbf{k}_i) - (K - 2i_0 + 2)\sigma^2.$$

For a given value of $\beta_\cdot$, an optimal value $\lambda_\cdot^*$ of $\lambda_\cdot$ is obtained by minimizing the so-defined piecewise second-order binomial function. Then, a search on the optimal value of $\beta_\cdot$ is carried out to minimize $\hat{R}(\lambda_\cdot^*, \beta_\cdot)$.

4.5. Different cases

The above expressions are valid for the primal tree but quite similar results are obtained for the dual one. In the expressions of the risk, the correlation of the noise, which is related to the redundancy of the dual-tree wavelet decomposition, is summarized by $\kappa_\cdot(\mathbf{k})$ and its dual counterpart $\kappa_\cdot^H(\mathbf{k}) = \sigma^{-2}(\tilde{\mathbf{r}}_\cdot^H(\mathbf{k}))^\top\mathbf{\Gamma}^{(\tilde{n}_\cdot, n_\cdot)}$. Let us write $\tilde{\mathbf{r}}_\cdot(\mathbf{k}) = [(r_\cdot(\mathbf{k}+\ell))_{\ell\in\mathcal{D}_\cdot}, (r_\cdot^H(\mathbf{k}+\ell))_{\ell\in\mathcal{V}_\cdot}]$ and choose a symmetric neighborhood form in the dual tree: $\tilde{\mathbf{r}}_\cdot^H(\mathbf{k}) = [(r_\cdot(\mathbf{k}+\ell))_{\ell\in\mathcal{V}_\cdot}, (r_\cdot^H(\mathbf{k}+\ell))_{\ell\in\mathcal{D}_\cdot}]$. From the results in Section 3, the noise correlation terms take the following expressions, for $\mathbf{m} = (m_1, m_2)$ and $\mathbf{k} = (k_1, k_2)$,

$$\kappa_\cdot(\mathbf{k}) = \sum_{(\ell_1, \ell_2)\in\mathcal{V}_\cdot} r_\cdot^H(k_1+\ell_1, k_2+\ell_2)\gamma_{m_1, m_1'}(\ell_1)\, \gamma_{m_2, m_2'}(\ell_2)$$

$$\kappa_\cdot^H(\mathbf{k}) = \sum_{(\ell_1, \ell_2)\in\mathcal{V}_\cdot} r_\cdot(k_1+\ell_1, k_2+\ell_2)\gamma_{m_1, m_1'}(\ell_1)\, \gamma_{m_2, m_2'}(\ell_2).$$

Due to the symmetries of the correlation functions of the wavelets, it can be futher noticed that the indices such that $\ell_1 = 0$ (resp. $\ell_2 = 0$) can be omitted in the above summations when $m_1 \neq 0$ (resp. $m_2 \neq 0$). Besides, after the postprocessing of a subband $(j, \mathbf{m})$ with $\mathbf{m} \in \mathbb{N}_M^{\star 2}$, the additive noise model (12) becomes $v_\cdot(\mathbf{k}) = u_\cdot(\mathbf{k}) + w_\cdot(\mathbf{k})$, where $v_\cdot(\mathbf{k})$ (resp. $u_\cdot(\mathbf{k})$) is a post-transformed coefficient of the noisy (resp. clean) signal. Section 3 indicates that the second-order properties of the noise are modified so that, after transformation, the following expressions of the noise

correlation terms hold:

$$\zeta_\cdot(\mathbf{k}) = \sum_{\substack{(\ell_1,\ell_2)\in\mathcal{D}_\cdot \\ \ell_1\neq 0\,,\,\ell_2\neq 0}} v_\cdot(k_1+\ell_1,k_2+\ell_2)\gamma_{m_1,m_1'}(\ell_1)\,\gamma_{m_2,m_2'}(\ell_2)$$

$$\zeta_\cdot^{\mathrm{H}}(\mathbf{k}) = -\sum_{\substack{(\ell_1,\ell_2)\in\mathcal{D}_\cdot \\ \ell_1\neq 0\,,\,\ell_2\neq 0}} v_\cdot^{\mathrm{H}}(k_1+\ell_1,k_2+\ell_2)\gamma_{m_1,m_1'}(\ell_1)\,\gamma_{m_2,m_2'}(\ell_2)$$

with $\zeta_\cdot(\mathbf{k}) = \sigma^{-2}\tilde{\mathbf{v}}_\cdot^\top(\mathbf{k})\boldsymbol{\Gamma}^{(\tilde{\mathbf{w}},w)}$, $\zeta_\cdot^{\mathrm{H}}(\mathbf{k}) = \sigma^{-2}\left(\tilde{\mathbf{v}}_\cdot^{\mathrm{H}}(\mathbf{k})\right)^\top\boldsymbol{\Gamma}^{(\tilde{\mathbf{w}}^{\mathrm{H}},w^{\mathrm{H}})}$ and where straightforward extensions of our notations have been used. This shows that, in spite of the whiteness of the noise, spatial correlations must be taken into account in this case.

5. SIMULATION RESULTS

Test images s (Barbara and Boat) of size 512×512 coded at 8 bpp are corrupted by a zero-mean additive white Gaussian noise which is independent of s. We are interested in evaluating the impact of choosing different decompositions and neighborhoods on the performances of our estimator. More precisely, we consider M-band Meyer wavelets with $M = 2$ or $M = 4$ and two possible neighborhoods called PN1 and PN2. PN1 is a purely "inter-tree" neighborhood since it does not include any spatial information. With the notations used in Section 4.5, PN1 corresponds to $\mathcal{D}_\cdot = \emptyset$ and $\mathcal{V}_\cdot = \emptyset$ (resp. $\mathcal{V}_\cdot = \{(0,0)\}$) for the subbands where the post-transform is (resp. is not) applied. The neighborhood PN2 combines both spatial and "inter-tree" information. We choose $\mathcal{D}_\cdot = \{-2,\ldots,2\}^2 \setminus \{(0,0)\}$ and $\mathcal{V}_\cdot = \emptyset$ (resp. $\mathcal{V}_\cdot = \{(0,0)\}$) for the subbands where the post-transform is (resp. is not) applied. The denoising performance is evaluated in terms of Signal to Noise Ratio (SNR). The achieved results are compared with those provided by the Neighblock [2] and SUREshrink [12] estimators. The wavelet decomposition is performed over 4 resolution levels when $M = 2$ and 2 levels when $M = 4$, in order to generate the same size for the approximation subband at the coarsest resolution. The resulting SNRs are listed in Table 1. It can be noted that our estimator achieves the best results when $M = 2$ or $M = 4$. On the one hand, the comparison of the two first columns shows that the addition of the information brought by the dual tree (resp. primal tree) is useful in the estimation of the coefficients in the primal tree (resp. dual tree). On the other hand, the comparison of the two last columns demonstrates the effectiveness of our adaptive threshold versus a fixed one as in the Neighblock method. A visual inspection of the denoised images leads to the same conclusions. Fig. 2 shows cropped denoised versions of Barbara: the left image is obtained by the Neighblock, the right one is provided by our estimator. Neighblock introduces more artefacts especially in the black right upper corner and on Barbara's left cheek. Besides, granular noise in the uniform areas is more important in the Neighblock estimated image.

6. CONCLUSION

In this paper, we have proposed a new DTT denoising method which was derived from a generalized Stein's formula. This method is applicable to images corrupted by a white Gaussian noise. The redundancy of the DTT introduces noise correlations which have been exploited in our approach. Experiments show that the proposed method leads to improved results compared with state-of-the-art methods. Several extensions of this work can be envisaged.

In particular, the case of neighborhoods exploiting interscale dependencies is under investigation.

7. REFERENCES

[1] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, Sep. 1994.

[2] T. T. Cai and B. W. Silverman, "Incorporating information on neighboring coefficients into wavelet estimation," *Sankhya*, vol. 63, pp. 127–148, 2001.

[3] N.G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *J. of Appl. and Comp. Harm. Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.

[4] I. W. Selesnick, "Hilbert transform pairs of wavelet bases," *Signal Processing Letters*, vol. 8, no. 6, pp. 170–173, Jun. 2001.

[5] C. Chaux, L. Duval, and J.-C. Pesquet, "Image analysis using a dual-tree M-band wavelet transform," *to appear in IEEE Trans. on Image Proc.*, 2006.

[6] C. Chaux, L. Duval, and J.-C. Pesquet, "Étude du bruit dans une analyse M-bandes en arbre dual," in *Proc. GRETSI*, Louvain, Belgique, Sep. 2005, pp. 229–232.

[7] C. Chaux, L. Duval, and J.-C. Pesquet, "2D dual-tree M-band wavelet decomposition," in *Proc. Int. Conf. on Acoust., Speech and Sig. Proc.*, Philadelphia, USA, March 18-23, 2005.

[8] P. Steffen, P. N. Heller, R. A. Gopinath, and C. S. Burrus, "Theory of regular M-band wavelet bases," *IEEE Trans. on Signal Proc.*, vol. 41, no. 12, pp. 3497–3511, Dec. 1993.

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

[10] L. Şendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," *Signal Processing Letters*, vol. 9, no. 12, pp. 438–441, Dec. 2002.

[11] C. Chaux, A. Benazza-Benyahia, and J.-C. Pesquet, "A block-thresholding method for multispectral image denoising," in *Proc. SPIE*, San Diego, CA, USA, Aug. 2005.

[12] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, vol. 90, pp. 1200–1224, Dec. 1995.

		SURE	PN1	NB	PN2
Barb.	DTT $M=2$	12.94	13.53	14.00	14.39
	DTT $M=4$	13.43	14.02	14.33	**14.72**
Boat	DTT $M=2$	13.49	13.65	13.78	**14.09**
	DTT $M=4$	13.34	13.54	13.69	13.91

Table 1. Denoising results for Barbara image (initial SNR of 6.17 dB) and for Boat image (initial SNR of 6.03 dB). The considered estimators are SUREshrink (SURE), the proposed estimator using PN1 neighborhood, Neighblock (NB) and the proposed estimator using PN2 neighborhood.

Fig. 2. Cropped versions of Barbara denoised using 4-bands wavelets: Neighblock estimator (left), proposed one using PN2 neighborhood (right).