AN INTERSCALE MULTIVARIATE MAP ESTIMATION OF MULTISPECTRAL IMAGES

Amel Benazza-Benyahia [1], *Jean-Christophe Pesquet* [2]

[1] Dept. MASC, Ecole Supérieure des Communications de Tunis,
Route de Raoued 3.5 Km, 2083 Ariana, Tunisia
[2] IGM and UMR-CNRS 5141,
Université de Marne-la-Vallée,
5, Bd Descartes, Champs sur Marne, 77454 Marne la Vallée Cédex 2, France.
e-mail: ben.yahia@planet.tn,pesquet@univ-mlv.fr

ABSTRACT

In this paper, we develop a multivariate statistical approach for image denoising in the wavelet transformed domain. To this respect, the wavelet coefficients of all the image channels at the same spatial position, in a given orientation and at the same resolution level, are grouped into a vector and a maximum a posteriori estimate is derived from a multivariate Bernouilli-Gaussian prior. The parameters of this statistical model are computed recursively from coarse to fine resolutions in order to exploit the inter-scale redundancies between the wavelet coefficients. Simulation tests performed on remote sensing multispectral images indicate that the proposed procedure improves the conventional wavelet-based denoising methods.

1. INTRODUCTION

Most images are corrupted by noise due to the imperfections of the acquisition systems and the communication channels. Therefore, much attention has been paid in developing efficient denoising methods which can be classified into non Bayesian methods and Bayesian ones [1]. The first category of methods often assumes that the unknown signal is deterministic whereas in the second one, the unknown image is considered as a realization of a random field with a given prior probability distribution. Therefore, the challenging task consists in selecting a realistic prior distribution that allows a both efficient and tractable denoising approach. To this respect, statistical models in the Wavelet Transform (WT) domain have been proposed in [2, 3]. The rationale is that the statistics of many natural images, once transformed, are simplified thanks to the good energy concentration and decorrelation properties of the WT [4]. In earlier works, the wavelet coefficients have been considered as independent, identically distributed (iid) according to a Generalized Gaussian Distribution (GGD) and, the related Maximum A Posteriori (MAP) estimate have led to wavelet thresholding rules for a wide class of noise distributions [5, 6, 7]. Another alternative consists in considering the wavelet coefficients as independent variables distributed according to Gaussian mixtures [8, 9, 10], which allows to deduce appropriate denoising rules. However, in practice, the wavelet coefficients are not statistically independent since it has been noted that large-magnitude coefficients tend to occur at the same positions in subbands at adjacent scales [11]. An effort has been performed in order to design prior distribution that capture the inter-scale dependencies such as the hidden Markov models [11] or the Markov random field models in the WT domain [3, 12]. Recently, the bivariate shrinkage has appeared as an appealing tool due to its good performances [13].
In parallel to these works, much interest has been dedicated to noise removal in multichannel images acquired by sensors operating in different spectral ranges. Two main approaches are possible: process separately each component or process simultaneously the spectral component according to a multivariate approach [14]. In the latter case, an attempt has already been made in order to exploit the spectral redundancies. Indeed, in [15] we have developed an efficient MAP estimation of the wavelet coefficients based on a *multivariate* Bernouilli-Gaussian models. These models reflect the sparseness of the wavelet representation as well as the statistical dependencies existing between the different components. In this paper, we investigate the inter-scale redundancies in order to improve the performances of this new multivariate estimation procedure. The paper is organized as follows. In Section 2, the theoretical background is presented. In Section 3, we propose a very simple method that takes into account the inter-scale dependencies. In Section 5, experimental results are given and some concluding remarks are drawn.

2. THEORETICAL BACKGROUND

2.1 The observation model

The unknown multichannel image consists of $B \in \mathbb{N}^*$ spectral components $s^{(b)}$ of size $L \times L$ with $b \in \{1, \ldots, B\}$. These components are corrupted by an additive noise and, the observation vector $\mathbf{r}$ is expressed in the spatial domain as follows:

$$\forall \mathbf{m} \in \{1, \ldots, L\}^2, \quad \mathbf{r}(\mathbf{m}) = \mathbf{s}(\mathbf{m}) + \mathbf{n}(\mathbf{m}) \qquad (1)$$

where the noise $\mathbf{n}(\mathbf{m}) = (n^{(1)}(\mathbf{m}), \ldots, n^{(B)}(\mathbf{m}))^T$ is iid $\mathcal{N}(\mathbf{0}, \mathbf{R}^{(n)})$, independent of $\mathbf{s}(\mathbf{m}) = (s^{(1)}(\mathbf{m}), \ldots, s^{(B)}(\mathbf{m}))^T$. A separable J-stage WT is applied to each component $r^{(b)}$ of $\mathbf{r}$ and generates at each resolution j, 3 wavelets coefficients $r_j^{(b,o)}$ oriented horizontally ($o = 1$), vertically ($o = 2$) or diagonally ($o = 3$), all being of size $L_j \times L_j$ with $L_j = L/2^j$. Then, the coefficients $r_j^{(b,o)}$ of all the B channels at the same spatial position $\mathbf{k}$ in a subband of the same orientation o and, at a given resolution level j, are stacked into an observation vector $\mathbf{r}_j^{(o)}(\mathbf{k}) \triangleq (r_j^{(1,o)}(\mathbf{k}), \ldots, r_j^{(B,o)}(\mathbf{k}))^T$. As the WT is linear, $\mathbf{r}_j^{(o)}(\mathbf{k}) = \mathbf{s}_j^{(o)}(\mathbf{k}) + \mathbf{n}_j^{(o)}(\mathbf{k})$ where $\mathbf{s}_j^{(o)}(\mathbf{k})$ and $\mathbf{n}_j^{(o)}(\mathbf{k})$ are defined similarly to $\mathbf{r}_j^{(o)}(\mathbf{k})$. It is easy to check that $\mathbf{n}_j^{(o)}(\mathbf{k}) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_j^{(n,o)})$ where $\mathbf{\Gamma}_j^{(n,o)} = \mathbf{R}^{(n)}$. Unlike the conventional approach, it is worth noting that the multivariate approach allows to take into account the inter-channel correlation between the noise samples by selecting a non diagonal matrix $\mathbf{\Gamma}_j^{(n,o)}$. In the sequel, the noise covariance matrix $\mathbf{R}^{(n)}$ is assumed to be known.

2.2 Prior Model

In the Bayesian framework, the "clean" wavelet coefficients $\mathbf{s}_j^{(o)}$ are viewed as realizations of random processes. Following [8, 9, 10], we have extended this approach to the vector case by introducing *multivariate* Bernouilli-Gaussian (BG) priors able to reflect the sparness of the wavelet representation. Hence, the probability distribution $p_j^{(o)}$ of $\mathbf{s}_j^{(o)}$ is expressed as:

$$\forall \mathbf{u} \in \mathbb{R}^B, \quad p_j^{(o)}(\mathbf{u}) = (1-\epsilon_j^{(o)})\delta(\mathbf{u}) + \epsilon_j^{(o)} g_{\mathbf{0},\mathbf{\Gamma}_j^{(s,o)}}(\mathbf{u}), \quad (2)$$

where δ is the Dirac distribution, $g_{\mathbf{0},\mathbf{\Gamma}_j^{(s,o)}}$ denotes the probability density of a zero-mean multivariate Gaussian vector with covariance matrix $\mathbf{\Gamma}_j^{(s,o)}$ and, $\epsilon_j^{(o)} \in [0,1]$ is the mixture parameter. In order to avoid degenerate MAP estimates, it is used to couple the mixture model with hidden iid Bernouilli random variables $q_j^{(o)}(\mathbf{k})$ such that:

$$\begin{array}{rcl} p(\mathbf{s}_j^{(o)}(\mathbf{k})/q_j^{(o)}(\mathbf{k})=0) & = & \delta(\mathbf{s}_j^{(o)}(\mathbf{k})) \\ p(\mathbf{s}_j^{(o)}(\mathbf{k}) = /q_j^{(o)}(\mathbf{k})=1) & = & g_{\mathbf{0},\mathbf{\Gamma}_j^{(s,o)}}(\mathbf{s}_j^{(o)}(\mathbf{k})), \end{array} , \quad (3)$$

with $\mathsf{P}(q_j^{(o)}(\mathbf{k})=1) = \epsilon_j^{(o)}$. In this context, the denoising problem reduces to a classical problem of estimation theory. The hidden variables $q_j^{(o)}$ can be estimated and the MAP estimator $\hat{\mathbf{s}}_j^{(o)}$ of $\mathbf{s}_j^{(o)}$ is derived. Finally, the inverse WT is applied to the estimate $\hat{\mathbf{s}}_j^{(o)}$ in order to get the estimate of the spectral components, in the spatial domain.

3. INTERSCALE MAP ESTIMATION

3.1 Estimation of the Hidden Variables

In the case of binary random variables, the MAP estimator is the Bayesian estimator corresponding to a hit-or-miss cost. Therefore, as developed in [15], an estimate $\hat{q}_j^{(o)}(\mathbf{k}) = 1$ of $q_j^{(o)}(\mathbf{k})$ is obtained if

$$p(\mathbf{r}_j^{(o)}(\mathbf{k})/q_j^{(o)}(\mathbf{k})=0) < p(\mathbf{r}_j^{(o)}(\mathbf{k})/q_j^{(o)}(\mathbf{k})=1). \quad (4)$$

Such a denoising scheme could be improved by exploiting inter-scale redundancies. More precisely, we first recall that each father wavelet coefficient at position $\mathbf{k}_\mathrm{f} = (k_{(1,\mathrm{f})}, k_{(2,\mathrm{f})})$ has four children placed in the next finest subband at positions $\mathbf{k}$ that are $(2k_{(1,\mathrm{f})},\ 2k_{(2,\mathrm{f})})$, $(2k_{(1,\mathrm{f})},\ 2k_{(2,\mathrm{f})}+1)$, $(2k_{(1,\mathrm{f})}+1,\ 2k_{(2,\mathrm{f})})$ or $(2k_{(1,\mathrm{f})}+1,\ 2k_{(2,\mathrm{f})}+1)$. It has been noted that very often, when a father wavelet coefficient has a significant magnitude, its children tend also to be signficant and carry valuable information. In other words, it could be expected that the hidden variables $q_j^{(o)}$ present some redundancies from a scale to the next one. Therefore, our contribution consists in exploiting such dependencies by using the value of the hidden variable $q_{j+1}^{(o)}(\mathbf{k}_\mathrm{f})$ of the father in the estimation of the hidden variable of its four children $q_j^{(o)}(\mathbf{k})$. To this respect, Equation (4) is replaced by:

$$\begin{array}{l} \mathsf{P}(q_j^{(o)}(\mathbf{k})=0/q_{j+1}^{(o)}(\mathbf{k}_\mathrm{f}), \mathbf{r}_j^{(o)}(\mathbf{k})) \\ \qquad < \mathsf{P}(q_j^{(o)}(\mathbf{k})=1/q_{j+1}^{(o)}(\mathbf{k}_\mathrm{f}), \mathbf{r}_j^{(o)}(\mathbf{k})). \end{array} \quad (5)$$

In the sequel, for the sake of clarity, the spatial indices are dropped. The Bayes formula states that:

$$\begin{array}{rcl} \mathsf{P}(q_j^{(o)}/q_{j+1}^{(o)}, \mathbf{r}_j^{(o)}) & \propto & \mathsf{P}(q_j^{(o)}/q_{j+1}^{(o)})\mathsf{P}(q_{j+1}^{(o)})p(\mathbf{r}_j^{(o)}/q_j^{(o)}) \\ & \propto & \mathsf{P}(q_j^{(o)}/q_{j+1}^{(o)})p(\mathbf{r}_j^{(o)}/q_j^{(o)}) \end{array} . \quad (6)$$

Therefore, an estimate $\hat{q}_j^{(o)} = 1$ of $q_j^{(o)}$ should satisfy:

$$\begin{array}{l} \mathsf{P}(q_j^{(o)}=0/q_{j+1}^{(o)})p(\mathbf{r}_j^{(o)}/q_j^{(o)}=0) \\ \qquad < \mathsf{P}(q_j^{(o)}=1/q_{j+1}^{(o)})p(\mathbf{r}_j^{(o)}/q_j^{(o)}=1). \end{array} \quad (7)$$

After some manipulations, we obtain the following MAP estimator:

$$\hat{q}_j^{(o)} = \left\{ \begin{array}{ll} 1 & \text{if } (\mathbf{r}_j^{(o)})^T \mathbf{M}_j^{(o)} \mathbf{r}_j^{(o)} > \chi_{j,j+1}^{(o)}, \\ 0 & \text{else} \end{array} \right. \quad (8)$$

where $\mathbf{M}_j^{(o)}$ is the definite positive matrix:

$$\mathbf{M}_j^{(o)} = (\mathbf{\Gamma}_j^{(n,o)})^{-1} - (\mathbf{\Gamma}_j^{(s,o)} + \mathbf{\Gamma}_j^{(n,o)})^{-1}, \quad (9)$$

and the positive threshold $\chi_{j,j+1}^{(o)}$ is determined by:

$$\chi_{j,j+1}^{(o)} = 2\ln\left(\frac{\mathsf{P}(q_j^{(o)}=0/q_{j+1}^{(o)}) \mid \mathbf{\Gamma}_j^{(s,o)} + \mathbf{\Gamma}_j^{(n,o)} \mid^{1/2}}{\mathsf{P}(q_j^{(o)}=1/q_{j+1}^{(o)}) \mid \mathbf{\Gamma}_j^{(n,o)} \mid^{1/2}} \right). \quad (10)$$

This threshold can take the values $\chi_{j,j+1,0}^{(o)}$ if $q_{j+1}^{(o)} = 0$ and, $\chi_{j,j+1,1}^{(o)}$ if $q_{j+1}^{(o)} = 1$ where

$$\begin{array}{rcl} \chi_{j,j+1,0}^{(o)} & = & \ln\left(\frac{\pi_{j,00} |\mathbf{\Gamma}_j^{(s,o)} + \mathbf{\Gamma}_j^{(n,o)}|^{1/2}}{(1-\pi_{j,00})|\mathbf{\Gamma}_j^{(n,o)}|^{1/2}} \right) \\ \chi_{j,j+1,1}^{(o)} & = & \ln\left(\frac{(1-\pi_{j,11}) |\mathbf{\Gamma}_j^{(s,o)} + \mathbf{\Gamma}_j^{(n,o)}|^{1/2}}{\pi_{j,11}|\mathbf{\Gamma}_j^{(n,o)}|^{1/2}} \right) \end{array} , \quad (11)$$

with the following notations:

$$\left\{ \begin{array}{rcl} \pi_{j,11} & \triangleq & \mathsf{P}(q_j^{(o)}=1/q_{j+1}^{(o)}=1) \\ \pi_{j,00} & \triangleq & \mathsf{P}(q_j^{(o)}=0/q_{j+1}^{(o)}=0) \end{array} \right. . \quad (12)$$

3.2 MAP estimation of the wavelet coefficients

A MAP estimate of the signal can be obtained according to the estimated values of $\hat{q}_j^{(o)}$. Indeed, if $\hat{q}_j^{(o)} = 0$, the observation obviously reduces to noise and then $\hat{\mathbf{s}}_j^{(o)} = \mathbf{0}$. At the opposite, if $\hat{q}_j^{(o)} = 1$, a MAP estimator of the signal of interest can be derived:

$$\hat{\mathbf{s}}_j^{(o)} = \arg\max_{\mathbf{u}} p_{\mathbf{s}_j^{(o)}}(\mathbf{u} \mid \mathbf{r}_j^{(o)}, q_j^{(o)}=1). \quad (13)$$

Since the couple of vectors $(\mathbf{r}_j^{(o)}, \mathbf{s}_j^{(o)})$ corresponds to a Gaussian vector when $q_j^{(o)} = 1$, its posterior distribution is also Gaussian such that:

$$\mathrm{E}[\mathbf{s}_j^{(o)} \mid \mathbf{r}_j^{(o)}, q_j^{(o)}=1] = \mathbf{Q}_j^{(o)} \mathbf{r}_j^{(o)} \quad (14)$$

where

$$\mathbf{Q}_j^{(o)} \triangleq \mathbf{\Gamma}_j^{(s,o)} (\mathbf{\Gamma}_j^{(s,o)} + \mathbf{\Gamma}_j^{(n,o)})^{-1}. \quad (15)$$

Finally, we obtain a shrinkage rule that performs a tradeoff between a linear estimation in the sense of a minimum mean square error and a hard thresholding:

$$\hat{\mathbf{s}}_j^{(o)} = \left\{ \begin{array}{ll} \mathbf{Q}_j^{(o)} \mathbf{r}_j^{(o)} & \text{if } \hat{q}_j^{(o)} = 1 \\ \mathbf{0} & \text{otherwise} \end{array} \right. . \quad (16)$$

4. ESTIMATION OF THE HYPERPARAMETERS

The hyperparameters $\mathbf{\Gamma}_j^{(s,o)}$ and $\epsilon_j^{(o)}$ of the Bernouilli-Gaussian priors can be estimated thanks to the method of moments that we have already proposed in [15]. However, the computation of the thresholds $\chi_{j,j+1,0}^{(o)}$ and $\chi_{j,j+1,1}^{(o)}$ also necessitates the estimation of the transition probabilities $\pi_{j,00}^{(o)}$ and, $\pi_{j,11}^{(o)}$. First of all, it is worth noting that:

$$\begin{aligned}\epsilon_j^{(o)} &= \sum_{q_{j+1}^{(o)}} \mathsf{P}(q_j^{(o)}=1/q_{j+1}^{(o)})\mathsf{P}(q_{j+1}^{(o)}) \\ &= \epsilon_{j+1}^{(o)}\pi_{j,11}^{(o)} + (1-\epsilon_{j+1}^{(o)})(1-\pi_{j,00}^{(o)})\end{aligned} . \quad (17)$$

Consequently, it is enough to estimate only one transition probability, e.g. $\pi_{j,11}^{(o)}$. To this respect, we will first denote $|\mathbf{r}_j^{(o)}|$ (resp. $|\mathbf{r}_{j+1}^{(o)}|$) the vector whose components are the magnitudes of the components of $\mathbf{r}_j^{(o)}$ (resp. $\mathbf{r}_{j+1}^{(o)}$). We propose to compute the inter-covariance $\mathbf{A}_j^{(o)}$ of these magnitude vectors:

$$\mathbf{A}_j^{(o)} \triangleq E[|\mathbf{r}_j^{(o)}|\,|\mathbf{r}_{j+1}^{(o)}|^T] - E[|\mathbf{r}_j^{(o)}|]E[|\mathbf{r}_{j+1}^{(o)}|]^T . \quad (18)$$

By noting that:

$$E[|\mathbf{r}_j^{(o)}|\,|\mathbf{r}_{j+1}^{(o)}|^T] = \sum_{q_j^{(o)},q_{j+1}^{(o)}} E[|\mathbf{r}_j^{(o)}|/q_j^{(o)}]E[|\mathbf{r}_{j+1}^{(o)}|/q_{j+1}^{(o)}]^T \mathsf{P}(q_j^{(0)},q_{j+1}^{(o)}) , \quad (19)$$

we deduce that:

$$\mathbf{A}_j^{(o)} = \sum_{q_j^{(o)},q_{j+1}^{(o)}} E[|\mathbf{r}_j^{(o)}|/q_j^{(o)}]E[|\mathbf{r}_{j+1}^{(o)}|/q_{j+1}^{(o)}]^T \left(\mathsf{P}(q_j^{(o)},q_{j+1}^{(o)}) - \mathsf{P}(q_j^{(o)})\mathsf{P}(q_{j+1}^{(o)})\right) . \quad (20)$$

Besides, as $\mathsf{P}(q_j^{(o)},q_{j+1}^{(o)}) = \mathsf{P}(q_j^{(o)}|q_{j+1}^{(o)})\ \mathsf{P}(q_{j+1}^{(o)})$, Equation (20) reduces to:

$$\mathbf{A}_j^{(o)} = \sum_{q_j^{(o)},q_{j+1}^{(o)}} \mathsf{P}(q_{j+1}^{(o)})[\mathsf{P}(q_j^{(o)}/q_{j+1}^{(o)}) - \mathsf{P}(q_j^{(o)})] E[|\mathbf{r}_j^{(o)}|/q_j^{(o)}]E[|\mathbf{r}_{j+1}^{(o)}|/q_{j+1}^{(o)}]^T . \quad (21)$$

By taking into account the BG model and, after some calculations, $\mathbf{A}_j^{(o)}$ is expressed as follows:

$$\begin{aligned}\mathbf{A}_j^{(o)} = &\left(E[|\mathbf{r}_j^{(o)}|/q_j^{(o)}=1] - E[|\mathbf{r}_j^{(o)}|/q_j^{(0)}=0]\right) \\ &\left((\pi_{j,11}^{(o)}-\epsilon_j^{(o)})\epsilon_{j+1}^{(o)}E[|\mathbf{r}_{j+1}^{(o)}|/q_{j+1}^{(o)}=1]\right. \\ &\left.+(1-\pi_{j,00}^{(o)}-\epsilon_j^{(o)})(1-\epsilon_{j+1}^{(o)})E[|\mathbf{r}_{j+1}^{(o)}|/q_{j+1}^{(o)}=0]\right)^T\end{aligned} . \quad (22)$$

By combining this expression with Equation (17), we obtain:

$$\begin{aligned}\mathbf{A}_j^{(o)} = &(\pi_{j,11}^{(o)}-\epsilon_j^{(o)})\epsilon_{j+1}^{(o)}\left(E[|\mathbf{r}_j^{(o)}|/q_j^{(o)}=1]\right. \\ &\left.-E[\mathbf{r}_j^{(o)}|/q_j^{(0)}=0]\right)\left(E[|\mathbf{r}_{j+1}^{(o)}|/q_{j+1}^{(o)}=1]\right. \\ &\left.-E[|\mathbf{r}_{j+1}^{(o)}|/q_{j+1}^{(o)}=0]\right)^T\end{aligned} . \quad (23)$$

Furthermore, $\mathbf{r}_j^{(o)}$ has a Gaussian distribution $\mathcal{N}(\mathbf{0},\ \mathbf{\Gamma}_j^{(n,o)})$ if $q_j^{(0)}=0$ and $\mathcal{N}(\mathbf{0},\ \mathbf{\Gamma}_j^{(s,o)}+\mathbf{\Gamma}_j^{(n,o)})$ otherwise. Hence, it is straightforward to show that:

$$\begin{aligned}E[|r_j^{(b,o)}|/q_j^{(0)}=0] &= \left(\tfrac{2}{\pi}[\mathbf{\Gamma}_j^{(n,o)}]_{b,b}\right)^{1/2} \\ E[|r_j^{(b,o)}|/q_j^{(0)}=1] &= \left(\tfrac{2}{\pi}[\mathbf{\Gamma}_j^{(s,o)}+\mathbf{\Gamma}_j^{(n,o)}]_{b,b}\right)^{1/2}\end{aligned} . \quad (24)$$

As a result, it is easy to compute $E[|r_j^{(b,o)}|/q_j^{(0)}=0]$ whereas $E[|r_j^{(b,o)}|/q_j^{(0)}=1]$ can be computed from an estimate of the hyperparameter $\mathbf{\Gamma}_j^{(s,o)}$. Consequently, in the expression of $\mathbf{A}_j^{(o)}$, the only unknown quantity is the transition probability $\pi_{j,11}^{(o)}$, which appears as the solution of a set of *linear* equations. A least square approach may be adopted to estimate it. It must be pointed out that $q_{j+1}^{(o)}$ is not exactly known and, in practice, it is replaced in Equation (23) by its estimate. Therefore, the whole denoising procedure is a recursive one from coarse to fine resolution. Indeed, at the coarsest level J, the wavelet coefficients of the signal are estimated according to the intrascale MAP procedure described in [15] and corresponding to Equation (4). At the resolution level $J-1$, the hyperparameters are estimated and, in particular, the transition probabilities are derived from the estimation of $q_J^{(o)}$. Then, the hidden variables $q_{J-1}^{(o)}$ are derived from Equation (8) and the MAP estimate of $\mathbf{s}_{J-1}^{(o)}$ is obtained from Equation (16). Then, the procedure is repeated at the susbequent finer levels.

5. EXPERIMENTAL RESULTS

Simulations were carried out on three-component images ($B=3$) of size 512×512 as the SPOT3 image "Tunis" and the RGB color image "Lena". We have artificially corrupted these components by adding a zero-mean Gaussian multivariate process. In our simulations, WT based on Symmlet of order 4 have been used. The performances are evaluated in terms of a Signal-to-Noise Ratio (SNR) averaged over the B components. For a fair comparison, we have also tested up-to-date wavelet-based methods. All of them are applied separately to each spectral component. To this purpose, we have tested the SUREshrink method [16], a MAP estimation method based on Generalized Gaussian (GG) priors [6], the locally bivariate shrinkage method [13]. Only the latter exploits the interscale redundancy. Besides, it is worth noting that in all the considered methods, the noise variance $[\mathbf{\Gamma}_j^{(n,o)}]_{b,b}$ is supposed to be known. The measure of performances is the average Signal-to-Noise Ratio (SNR). Table 1 provides the resulting SNR's for a three-stage decomposition ($J=3$). It is worth noting that the intrascale version of the multivariate BG-MAP estimation is more performant than the other previously developed methods, especially for very low values of the SNR. This result confirms the superiority of such multivariate statistical approach. Furthermore, it is clear that the interscale extension yields improvements of the SNR up to 0.3 dB w.r.t. to the intrascale BG-MAP. Better performances are obtained at deeper decomposition level as indicated in Table 2. For "Tunis", improvements ranging from 0.2dB to 0.3 dB are achieved w.r.t. the intrascale version, and the interscale BG-MAP clearly ouperforms the Bivariate method (about 1dB gain).
Finally, Figure 1 shows a cropped version of denoised images corresponding to the first channel XS1 of "Tunis". Both bivariate shrinkage and interscale BG-MAP attenuates the granular effect of the noise but textures look less blurry with the latter one than with former one.

REFERENCES

[1] I. Pitas, A. Venetsanopoulos, *Nonlinear digital filters*, Kluwer, Dodrecht, 1990.

[2] B. Vidakovic, *Statistical modeling by wavelets*, John Wiley & Sons, New-York, 1999.

[3] A. Pizurika, W. Philips, I. Lemahieu, M. Acheroy, "A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising," *IEEE Trans. on IP*, vol. 11, no. 5, pp. 545-556, May 2002.

[4] S. G. Mallat, *A wavelet tour of signal processing*, Academic Press, San Diego, 1998.

[5] P. Moulin, J. Liu, "Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors," *IEEE Trans. on IT*, vol. 45, no. 3, pp. 909-919, April 1999.

[6] A. Antoniadis, D. Leporini, J.-C. Pesquet, "Wavelet thresholding for some classes of non-Gaussian noise," *Statistica Neerlandica*, vol. 56, issue 4, pp. 434-453, December 2002.

[7] P. Ishwar, P. Moulin, "On the equivalence of set-theoretic and maxent MAP estimation," *IEEE Trans. on SP*, vol. 51, no. 3, pp. 698-713, March 2003.

[8] H.A. Chipman, E.R. Kolaczyk, R.E. McCulloch, "Adaptive Bayesian wavelet shrinkage," *JASA*, vol. 92, no. 440, pp. 1413-1421, December 1997.

[9] F. Abramovich, T. Sapatinas, B.W. Silverman, B.W., "Wavelet thresholding via a Bayesian approach," *J. of the Royal Stat. Soc.*, Series B, vol. 60, pp. 725-749, 1998.

[10] D. Leporini, J.-C. Pesquet, H. Krim, "Best basis representations based on prior statistical models," in *Bayesian Inference in Wavelet Based Models*, P. Müller and B. Vidakovic editors, Springer Verlag, 1999.

[11] J.K. Romberg, H. Choi, R.G. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," *IEEE Trans. on IP*, vol. 10, no. 7, pp. 1056-1068, July 2001.

[12] M. Malfait, D. Roose, "Wavelet-based image denoising using a Markov random field a priori model," *IEEE Trans. on IP*, vol. 6, no. 4, pp. 549-565, April 1997.

[13] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. on SP*, vol. 50, no. 11, pp. 2744-2756, November 2002.

[14] K. Tang, J. Astola, "Nonlinear multivariate filtering techniques," *IEEE Trans. on Image Proc.*, vol. 4, no. 6, pp. 788-797, June 1995.

[15] A. Benazza-Benyahia, J.-C. Pesquet, "Wavelet-based multispectral image denoising with Bernouilli-Gaussian model," *Proc. of the NSIP'03*, Grado, Italy, 8-11 June 2003.

[16] D.L. Donoho, I.M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *JASA*, vol. 90, no. 432, pp. 1200-1224, 1995.

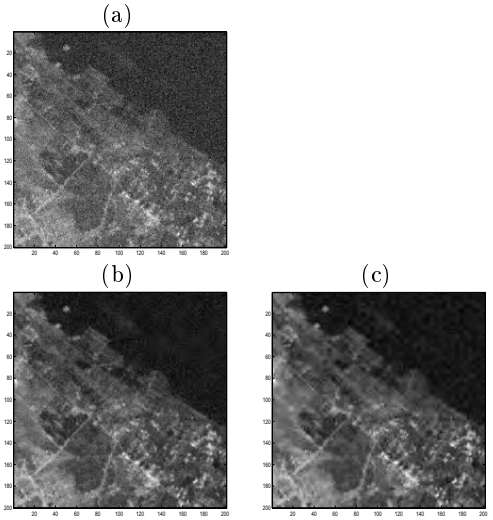

Figure 1: "Tunis" image, denoised red component, cropped version, (a): Noisy version, (b): bivariate method, (c): interscale BG-MAP ($J = 3$).

Table 1: Performances in terms of SNR (in dB) of three-stage wavelet-based denoising methods.

"Tunis"

Initial	GG MAP	Sure shrink	Bivariate inter	BG MAP intra	BG MAP inter
8.08	13.07	13.32	13.32	14.16	**14.29**
9.08	13.55	13.83	13.89	14.71	**14.87**
10.07	14.08	14.40	14.48	15.27	**15.44**
11.08	14.63	15.02	15.12	15.92	**16.11**
12.08	15.22	15.64	15.77	16.51	**16.72**
13.08	15.85	16.27	16.43	17.13	**17.37**
14.08	16.50	16.95	17.13	17.81	**18.09**
15.08	17.19	17.65	17.85	18.51	**18.80**

"Lena"

Initial	GG MAP	Sure shrink	Bivariate inter	BG MAP intra	BG MAP inter
7.84	14.49	15.01	15.23	15.76	**15.88**
8.34	14.78	15.36	15.55	15.99	**16.15**
8.84	14.99	15.55	15.83	16.21	**16.36**
9.84	15.48	16.01	16.42	16.66	**16.84**
10.84	15.97	16.50	16.97	17.12	**17.31**
11.84	16.46	17.16	17.51	17.57	**17.74**
12.84	16.96	17.69	18.06	18.00	**18.20**
13.84	17.45	18.24	18.61	18.49	**18.65**
14.82	17.97	18.83	**19.15**	18.95	19.10

Table 2: "Tunis" image, influence of the number of resolution levels J on the performances of the BG-MAP methods.

init	intra $J = 3$	inter $J = 3$	intra $J = 4$	inter $J = 4$
8.08	14.16	14.29	14.20	14.41
9.08	14.71	14.87	14.75	14.97
10.07	15.27	15.44	15.30	15.53
11.08	15.92	16.11	15.94	16.20
12.08	16.51	16.72	16.53	16.80
13.08	17.13	17.37	17.14	17.43
14.08	17.81	18.09	17.82	18.13
15.08	18.51	18.80	18.52	18.84