GENERALIZED MULTIVARIATE EXPONENTIAL POWER PRIOR FOR WAVELET-BASED MULTICHANNEL IMAGE RESTORATION

Y. Marnissi[1]*, A. Benazza-Benyahia*[1]*, E. Chouzenoux*[2]*, J.-C. Pesquet*[2]

[1] COSIM Lab., SUP'COM, Carthage Univ.,
Cité Technologique des Communications,
2080, Tunisia
marnissi.yosra@gmail.com
benazza.amel@supcom.rnu.tn

[2] Université Paris-Est,
LIGM and UMR-CNRS 8049,
Champs-sur-Marne, 77454
Marne-la-Vallée, France
first.last@univ-paris-est.fr

ABSTRACT

In multichannel imaging, several observations of the same scene acquired in different spectral ranges are available. Very often, the spectral components are degraded by a blur modelled by a linear operator and an additive noise. In this paper, we address the problem of recovering the image components in a wavelet domain by adopting a variational approach. Our contribution is twofold. First, an appropriate multivariate penalty function is derived from a novel joint prior model of the probability distribution of the wavelet coefficients located at the same spatial position in a given subband through all the channels. Secondly, we address the challenging issue of computing the Maximum A Posteriori estimate by using a Majorize-Minimize optimization strategy. Simulation tests carried out on multispectral satellite images show that the proposed method outperforms conventional techniques.

Index Terms— Multiscale decomposition, multicomponent image, MAP criterion, Majorize-Minimize algorithm

1. INTRODUCTION

Multichannel Images (MCI) are widely used in many application areas such as medical imaging and remote sensing. The multiple components are obtained by imaging a single scene by sensors operating in different spectral ranges. For instance, about a dozen of radiometers may be on-board of remote sensing satellites. Most of the time, MCI are corrupted with noise and blurred during the acquisition process and transmission steps. Therefore, restoring MCI is of primary importance for several applications such as classification, segmentation and object recognition [1]. Several works have been dedicated to MCI denoising especially by using wavelet-based approaches. The challenge is to design efficient multivariate methods that jointly process the spectral components [2, 3, 4, 5]. In this paper, we adopt a Bayesian framework for recovering the wavelet coefficients. The challenge is twofold. At first, it consists of developping a multivariate prior probability distribution that reflects the sparsity of the employed multiscale representation. The second challenge is to design an efficient algorithm in order to optimize the resulting regularized criterion. In this work, we propose a novel multivariate prior model for the distribution of vectors composed of wavelet coefficients at the same spatial position in a given subband through all the channels. Secondly, we propose an algorithm based on recent developments concerning Majorize-Minimize methods in order to derive the optimal estimate.

The paper is organized as follows. In Section 2, we present the observation model and we formulate the problem. In Section 3, we introduce the adopted regularization criterion. In Section 4, experimental results are given and finally, in Section 5, some concluding remarks are drawn.

2. PROBLEM STATEMENT

We are interested in recovering a multicomponent signal with B components $\bar{\mathbf{y}}_1, \ldots, \bar{\mathbf{y}}_B$ in $\mathbb{R}^N$ (the images being columnwise reshaped) from some observations $\mathbf{z}_1, \ldots, \mathbf{z}_B$ which have been degraded by linear operators $\mathbf{H}_1, \ldots, \mathbf{H}_B$ and corrupted by an additive noise:

$$\forall b \in \{1, \ldots, B\}, \qquad \mathbf{z}_b = \mathbf{H}_b \bar{\mathbf{y}}_b + \mathbf{w}_b. \tag{1}$$

Very often, the target images are assumed to have a sparse representation on a set of frame synthesis operators $F_1^*, \ldots, F_b^*$:

$$\forall b \in \{1, \ldots, B\}, \qquad \bar{\mathbf{y}}_b = F_b^* \bar{\mathbf{x}}_b \tag{2}$$

where F_b^* is a linear operator from $\mathbb{R}^Q$ to $\mathbb{R}^N$ with $Q \geq N$. Consequently, the problem amounts to first building estimates $\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_B$ of the unknown coefficients $\bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_B$, and then recovering the estimates $\hat{\mathbf{y}}_1, \ldots, \hat{\mathbf{y}}_B$ as follows

$$\forall b \in \{1, \ldots, B\}, \qquad \hat{\mathbf{y}}_b = F_b^* \hat{\mathbf{x}}_b. \tag{3}$$

It is worth noting that the Q coefficients are generally grouped into M subbands of size Q_m (corresponding to specific orientations and scales) with $Q = \sum_{m=1}^{M} Q_m$. Hence, we have:

$$\begin{aligned} \forall b \in \{1, \ldots, B\},\ \bar{\mathbf{x}}_b = \quad & (\bar{x}_{b,1,1}, \ldots, \bar{x}_{b,1,Q_1}, \ldots, \\ & \bar{x}_{b,m,1}, \ldots, \bar{x}_{b,m,Q_m}, \ldots, \\ & \bar{x}_{b,M,1}, \ldots, \bar{x}_{b,M,Q_M})^\mathsf{T}. \end{aligned} \tag{4}$$

In this work, we aim at exploiting the cross-component similarities by estimating jointly the coefficients of the different subbands of all the components. In this respect, we define the following vector of multichannel coefficients of dimension B, for every $m \in \{1, \ldots, M\}$ and $q \in \{1, \ldots, Q_m\}$,

$$\bar{\mathbf{x}}_{m,q} = (\bar{x}_{1,m,q}, \ldots, \bar{x}_{B,m,q})^\mathsf{T}. \tag{5}$$

It can be noticed that such vectors can be easily obtained as follows:

$$\bar{\mathbf{x}}_{m,q} = \mathbf{P}_{m,q} \bar{\mathbf{x}} \quad \text{with} \quad \bar{\mathbf{x}} = (\bar{\mathbf{x}}_1^\mathsf{T}, \ldots, \bar{\mathbf{x}}_B^\mathsf{T})^\mathsf{T}, \tag{6}$$

where $\mathbf{P}_{m,q}$ is a $B \times QB$ sparse matrix containing B lines of a permutation matrix. In this paper, we focus on the solution to the following penalized optimization problem:

$$\underset{\mathbf{x} \in \mathbb{R}^{QB}}{\text{minimize}} \quad \left(\phi(\mathbf{H}F^*\mathbf{x} - \mathbf{z}) + \Psi(\mathbf{x})\right), \tag{7}$$

where $\mathbf{z} = (\mathbf{z}_1^\mathsf{T}, \ldots, \mathbf{z}_B^\mathsf{T})^\mathsf{T}$, $\phi : \mathbb{R}^{QB} \longrightarrow \mathbb{R}$ is a data-fidelity term, and $\Psi : \mathbb{R}^{QB} \longrightarrow \mathbb{R}$ is a regularization term. The operators $\mathbf{F}$ and $\mathbf{H}$ are defined as follows:

$$\mathbf{F}^* = \begin{pmatrix} \mathbf{F}_1^* & \mathbf{0} & \ldots & \mathbf{0} \\ 0 & \mathbf{F}_2^* & \mathbf{0} & \mathbf{0} \\ \ddots & \ddots & \ddots & \ddots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{F}_B^* \end{pmatrix}, \tag{8}$$

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ 0 & \mathbf{H}_2 & \mathbf{0} & \mathbf{0} \\ \ddots & \ddots & \ddots & \ddots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{H}_B \end{pmatrix} \tag{9}$$

where the size of the involved matrices $\mathbf{F}_b^*$, $\mathbf{H}_b^*$ and $\mathbf{0}$ is $Q \times Q$. Note that the data-fidelity term is related to the statistical properties of the noise whereas the regularization term depends on the prior knowledge and the constraints on the signal to be recovered. Therefore, the choice of Ψ is of main importance. In the next section, we will address this issue.

3. PROPOSED OPTIMISATION CRITERION

3.1. MEP prior

Recently, it has been observed [6, 7, 8] that a suitable model for the distribution of zero-mean vectors $\bar{\mathbf{x}}_{m,q}$ composed of frame coefficients through multiple channels in a given subband m is the Multivariate Exponential Power (MEP) distribution (also known as the multivariate generalized Gaussian) whose parameters are the shape parameter $\beta_m > 0$, and a symmetric positive matrix $\mathbf{\Sigma}_m$ (the scale matrix). The generic expression of a MEP probability density function (pdf) is defined for every $\mathbf{u}$ in $\mathbb{R}^B$ as follows:

$$f_{\text{MEP}}(\mathbf{u}) = C_B |\mathbf{\Sigma}_m|^{-1/2} g\left(\mathbf{u}^\mathsf{T} \mathbf{\Sigma}_m^{-1} \mathbf{u}; \beta_m\right), \tag{10}$$

where C_B is a normalization constant expressed as

$$C_B = \frac{\Gamma(\frac{B}{2})}{\pi^{\frac{B}{2}} \Gamma(\frac{B}{2\beta_m}) 2^{\frac{B}{2\beta_m}}} |\mathbf{\Sigma}_m|^{-1/2}, \tag{11}$$

$\Gamma(\cdot)$ being the gamma function, and where, for every $t \in \mathbb{R}_+$, $g(t; \beta_m) = \exp\left(-\frac{1}{2} t^{\beta_m}\right)$.
In this work, it appeared useful to consider the following Generalized MEP (GMEP) model, which is defined in the zero-mean case by the following pdf:

$$\forall \mathbf{u} \in \mathbb{R}^B,\ f_{\text{GMEP}}(\mathbf{u}) = C'_B |\mathbf{\Sigma}_m|^{-1/2} g\left(\mathbf{u}^\mathsf{T} \mathbf{\Sigma}_m^{-1} \mathbf{u} + \delta_m; \beta_m\right)$$

where $\mathbf{\Sigma}_m$ and β_m are defined as previously, δ_m is an additional positive constant, and C'_B is the associated normalization constant.

3.2. Regularization term related to GMEP prior

If we assume that the vectors $(\bar{\mathbf{x}}_{m,q})_{1 \leq m \leq M, 1 \leq q \leq Q_m}$ are realizations of mutually independent random vectors, for every $\mathbf{x} \in \mathbb{R}^{QB}$, the regularization term is given by

$$\Psi(\mathbf{x}) = \sum_{m=1}^{M} \sum_{q=1}^{Q_m} \left(\mathbf{x}_{m,q}^\mathsf{T} \mathbf{\Sigma}_m^{-1} \mathbf{x}_{m,q} + \delta_m\right)^{\beta_m}. \tag{12}$$

According to (6), this function can be re-expressed as

$$\Psi(\mathbf{x}) = \sum_{m=1}^{M} \sum_{q=1}^{Q_m} \left((\mathbf{P}_{m,q}\mathbf{x})^\mathsf{T} \mathbf{\Sigma}_m^{-1} (\mathbf{P}_{m,q}\mathbf{x}) + \delta_m\right)^{\beta_m}. \tag{13}$$

3.3. Penalized least squares problem

Let us assume that the noise vector $\mathbf{w} = [\mathbf{w}_1^\mathsf{T}, \ldots, \mathbf{w}_B^\mathsf{T}]^\mathsf{T}$ is a realization of a zero-mean Gaussian random vector with covariance matrix $\mathbf{\Lambda}$. The problem then reduces to a penalized least squares problem and the Maximum A Posteriori solution minimizes the following criterion:

$$\begin{aligned} \mathcal{J}_{\text{GMEP}}(\mathbf{x}) \quad &= (\mathbf{H}F^*\mathbf{x} - \mathbf{z})^\mathsf{T} \mathbf{\Lambda}^{-1} (\mathbf{H}F^*\mathbf{x} - \mathbf{z}) \\ &\quad + \textstyle\sum_{m=1}^{M} \sum_{q=1}^{Q_m} \psi_{\beta_m, \delta_m}(\parallel \mathbf{S}_{m,q}\mathbf{x} \parallel), \end{aligned} \tag{14}$$

where

- $\psi_{\beta_m, \delta_m} : \mathbb{R} \longrightarrow \mathbb{R}$ is the function defined as

$$\forall t \in \mathbb{R}_+, \quad \psi_{\beta_m, \delta_m}(t) = (t^2 + \delta_m)^{\beta_m}; \tag{15}$$

- $\mathbf{S}_{m,q}$ is the matrix $\mathbf{\Sigma}_m^{-1/2} \mathbf{P}_{m,q}$.

Obviously, the function ψ_{β_m, δ_m} is differentiable, and $t \mapsto \psi_{\beta_m, \delta_m}(\sqrt{t})$ is concave on $\mathbb{R}_+$ provided that $\beta_m \leq 1$. Moreover, ψ_{β_m, δ_m} is convex if $\beta_m \geq 1/2$.
Note that in the case of an additive Gaussian noise, $\mathcal{J}_{\text{GMEP}}$ is an extension of the criterion investigated only in the denoising context (i.e. $\mathbf{H} = \mathbf{I}$) in [9].

3.4. Majorize-Minimize memory gradient algorithm

Iterative methods need to be employed in order to minimize $\mathcal{J}_{\text{GMEP}}$. Basically, they start from an initial guess and build a sequence of updated estimates until an acceptable accuracy is reached. In the case of MCI restoration, we are faced to large scale optimization problems and, consequently, a major concern is to design a fast iterative optimization algorithm providing reliable numerical solutions. In this work, we employ the Majorize-Minimize Memory Gradient (3MG) algorithm [10, 11] which aims at building a sequence $(\mathbf{x}_k)_{k \in \mathbb{N}}$ of $\mathbb{R}^{QB}$ such that

$$\forall k \in \mathbb{N}, \qquad \mathcal{J}_{\text{GMEP}}(\mathbf{x}_{k+1}) \leq \mathcal{J}_{\text{GMEP}}(\mathbf{x}_k). \tag{16}$$

This is performed by defining, for all $k \geq 0$,

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k, \tag{17}$$

with $\mathbf{D}_k = [-\mathbf{g}_k \quad \mathbf{x}_k - \mathbf{x}_{k-1}] \in \mathbb{R}^{QB \times 2}$, where $\mathbf{g}_k$ denotes the gradient of $\mathcal{J}_{\text{GMEP}}$ at $\mathbf{x}_k$, and $\mathbf{u}_k \in \mathbb{R}^2$ is a two-dimensional stepsize that aims at partially minimizing $\varphi_k \colon \mathbf{u} \mapsto \mathcal{J}_{\text{GMEP}}(\mathbf{x}_k + \mathbf{D}_k \mathbf{u})$. In this algorithm, the determination of the stepsize $\mathbf{u}_k$ is based on the Majorization-Minimization (MM) principle. The minimization of φ_k is performed by successive minimizations of tangent majorant functions for φ_k. A function $q_k(\cdot, \mathbf{u}')$ is said to be a tangent majorant for φ_k at $\mathbf{u}'$ if for all $\mathbf{u} \in \mathbb{R}^2$,

$$q_k(\mathbf{u}, \mathbf{u}') \geqslant \varphi_k(\mathbf{u}) \quad \text{and} \quad q_k(\mathbf{u}', \mathbf{u}') = \varphi_k(\mathbf{u}'). \tag{18}$$

Following [11], we propose to employ a convex quadratic tangent majorant function of the form:

$$\begin{aligned} q_k(\mathbf{u}, \mathbf{u}') = \quad &\varphi_k(\mathbf{u}') + \nabla\varphi_k(\mathbf{u}')^\mathsf{T}(\mathbf{u} - \mathbf{u}') \\ &+ \tfrac{1}{2}(\mathbf{u} - \mathbf{u}')^\mathsf{T} \mathbf{B}_{k,\mathbf{u}'}(\mathbf{u} - \mathbf{u}'), \end{aligned} \tag{19}$$

where $\nabla\varphi_k(\mathbf{u}')$ denotes the gradient of φ_k at $\mathbf{u}'$ and $\mathbf{B}_{k,\mathbf{u}'}$ is a 2×2 symmetric positive semi-definite matrix that ensures the fulfillment of majorization properties (18). The minimization of φ_k

is thus replaced by a sequence of easier two-dimensional quadratic subproblems, corresponding to the following MM update rule:

$$\left\{\begin{array}{l} \mathbf{u}_k^0 = \mathbf{0}, \\ \text{For } j = 1, \ldots, J \\ \quad \left\lfloor \; \mathbf{u}_k^j \in \underset{\mathbf{u}\in\mathbb{R}^2}{\text{Argmin}} \; q_k(\mathbf{u}, \mathbf{u}_k^{j-1}). \right. \end{array}\right. \tag{20}$$

3.5. Construction of the majorizing approximation

For every $m \in \{1, \ldots, M\}$, let us define

$$\forall t \in \mathbb{R}^*, \quad \omega_{\beta_m,\delta_m}(t) = \dot{\psi}_{\beta_m,\delta_m}(t)/t, \tag{21}$$

where $\dot{\psi}_{\beta_m,\delta_m}$ is the derivative of ψ_{β_m,δ_m} (the function $\omega_{\beta_m,\delta_m}$ is extended by continuity at 0). Let $\mathbf{u}' \in \mathbb{R}^2$, let $k \in \mathbb{N}$, and let us define the following matrices:

$$\begin{array}{rl} \mathbf{A}(\mathbf{x}) & = 2(\mathbf{H}F^*)^\intercal \mathbf{\Lambda}^{-1}(\mathbf{H}F^*) \\ & + \sum_{m=1}^{M} \sum_{q=1}^{Q_m} \omega_{\beta_m,\delta_m}(\|\mathbf{S}_{m,q}\mathbf{x}\|) \left(\mathbf{S}_{m,q}^\intercal \mathbf{S}_{m,q}\right) \end{array}$$

$$\begin{array}{rl} \mathbf{B}_{k,\mathbf{u}'} & = \mathbf{D}_k^\top \mathbf{A}(\mathbf{x}_k + \mathbf{D}_k\mathbf{u}')\mathbf{D}_k, \\ & = 2(\mathbf{H}F^*\mathbf{D}_k)^\intercal \mathbf{\Lambda}^{-1}(\mathbf{H}F^*\mathbf{D}_k) \\ & + \sum_{m=1}^{M} \sum_{q=1}^{Q_m} \omega_{\beta_m,\delta_m}(\|\mathbf{S}_{m,q}\mathbf{x}_k + \mathbf{S}_{m,q}\mathbf{D}_k\mathbf{u}'\|) \\ & (\mathbf{S}_{m,q}\mathbf{D}_k)^\intercal(\mathbf{S}_{m,q}\mathbf{D}_k). \end{array}$$

Then, a convex quadratic tangent majorant $q_k(\cdot, \mathbf{u}')$ of φ_k at $\mathbf{u}'$ can be built according to (19).

3.6. Final algorithm

According to (19) and (20), the optimality condition for the choice of the stepsize in the MM iteration is given by: for every $k \in \mathbb{N}$ and $j \in \{1, \ldots, J\}$,

$$\mathbf{B}_{k,\mathbf{u}_k^{j-1}}(\mathbf{u}_k^j - \mathbf{u}_k^{j-1}) + \nabla\varphi_k(\mathbf{u}_k^{j-1}) = \mathbf{0}. \tag{22}$$

This yields the explicit stepsize formula

$$\mathbf{u}_k^j = \mathbf{u}_k^{j-1} - \mathbf{B}_{k,\mathbf{u}_k^{j-1}}^{-1} \nabla\varphi_k(\mathbf{u}_k^{j-1}), \tag{23}$$

where $\mathbf{B}_{k,\mathbf{u}_k^{j-1}}^{-1}$ is the pseudo-inverse of $\mathbf{B}_{k,\mathbf{u}_k^{j-1}} \in \mathbb{R}^{2\times 2}$. The resulting 3MG algorithm reads:

$$\left\{\begin{array}{l} \mathbf{x}_0 \in \mathbb{R}^{QB}, \\ \text{For } k = 0, \ldots \\ \quad \left\lfloor \begin{array}{l} \mathbf{u}_k^0 = \mathbf{0}, \\ \text{For } j = 1, \ldots, J \\ \quad \left\lfloor \begin{array}{l} \mathbf{B}_{k,\mathbf{u}_k^{j-1}} = \mathbf{D}_k^\intercal \mathbf{A}(\mathbf{x}_k + \mathbf{D}_k\mathbf{u}_k^{j-1})\mathbf{D}_k, \\ \mathbf{u}_k^j = \mathbf{u}_k^{j-1} - \mathbf{B}_{k,\mathbf{u}_k^{j-1}}^{-1} \mathbf{D}_k^\intercal \nabla\mathcal{J}_{\text{GMEP}}(\mathbf{x}_k + \mathbf{D}_k\mathbf{u}_k^{j-1}), \end{array} \right. \\ \mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{D}_k\mathbf{u}_k^J. \end{array} \right. \end{array}\right. \tag{24}$$

The convergence of the sequence $(\mathbf{x}_k)_{k\in\mathbb{N}}$ generated by the 3MG algorithm to a critical point of $\mathcal{J}_{\text{GMEP}}$ can be deduced from the results in [11]. In practice, the algorithm is run until the fulfillment of the stopping criterion

$$\|\nabla\mathcal{J}_{\text{GMEP}}(\mathbf{x}_k)\|/\sqrt{N} \leq \epsilon, \tag{25}$$

where, typically, $\epsilon = 10^{-4}$.

4. EXPERIMENTAL RESULTS

Synthetic GMEP data We have first generated realizations of a GMEP random vector thanks to the Metropolis-Hastings algorithm with the following parameters: $B = 3$, $M = 1$, $\mathbf{\Sigma}_1 = \mathbf{I}$, $\beta_1 = 0.3$, and $\delta_1 = 10^{-4}$. We have corrupted them by adding a multivariate i.i.d. Gaussian noise with zero-mean and variance $\sigma_b^2 = 100$. The noise variance was adjusted so as to correspond to an initial SNR = 11.2 dB. We plot in Fig. 1, the variations of $\left(\mathcal{J}_{\text{GMEP}}(\mathbf{x}_k) - \mathcal{J}_{\text{GMEP}}(\hat{\mathbf{x}})\right)_k$ for various values of J (number of MM steps), with respect to the computational time, when performing tests on an Intel Core i3CPU @ 2.39 GHz using a Matlab 7.12 implementation. Note that the optimal solution $\hat{\mathbf{x}}$ has been precomputed using a large number of iterations. It appears that the best results in terms of convergence profile are obtained with $J = 1$. Consequently, this setting has been retained in the remaining experiments.

Natural images We have considered several multispectral images. In this paper, we report the results for two test images of size 512 $\times$ 512. The first one is a three-component SPOT multispectral test image (called "Tunis") It corresponds to a scene depicting the city of Tunis ($B = 3$). The second one is a four-component Quickbird image (called "Zaghouan") corresponding to a rural area in Tunisia ($B = 4$). The spectral components have been corrupted with a multivariate zero-mean Gaussian noise having a diagonal covariance matrix. The quality of the denoised images is evaluated in terms of SNR. All the subsequent reported results are related to the 4-resolution wavelet orthonormal decomposition using symmlet wavelet of order 8. Note that simulations with other wavelet bases have led to similar results. The hyperparameters of the priors are adjusted empirically. The proposed method is compared with state-of-the-art denoising methods, namely the spatial 2D Wiener filter, the spectral Wiener filter, the non local means method (NLMeans) [12], Visushrink [13], SUREshrink [14], the BG-MAP method [16], and the E-SURE estimate [3]. We have also made a comparison with the locally bivariate shrinkage method which is a competitive technique based on the exploitation of the dependence between coefficients across scales [15]. Tables 1 and 2 provide the output SNR achieved by the different denoising techniques for the two test images. It can be observed that the proposed method outperforms the other methods. For the tested images,the average gain relatively to the E-SURE method is around 0.2 dB. Besides, the MM algorithm converges in no more than 100 iterations (2 iterations/s for $B = 3$, and 1 iteration/s for $B = 4$). Fig. 2 shows the denoised image corresponding to the first channel of "Tunis". The GMEP-based method attenuates the granular effect of the noise. This is confirmed by the perceptual measures of visual similarity (SSIM) indicated in the figure caption.

5. CONCLUSION

In this paper, we have proposed a wavelet-based variational method to restore multichannel images. The novelty of this work firstly relies on the new multivariate model adopted for the prior distribution. Secondly, an iterative Majorize-Minimize optimization algorithm have been applied to derive the optimal estimator. Experiments carried out on several multispectral satellite images have shown the good performance of the proposed approach with respect to state-of-the-art restoration methods. Several issues could be investigated in our future work such the ability of the proposed framework to exploit inter-scale dependencies in addition to the cross-channel ones.

Table 1. "Tunis" image, performance in terms of SNR (in dB) in the presence of a white noise.

Initial	Spatial Wiener	Spectral Wiener	NLMeans [12]	Visu shrink [13]	SURE shrink [14]	Bivariate [15]	BG MAP [16]	E-SURE [3]	GMEP MAP
15.08	16.84	16.61	17.13	13.26	17.74	17.96	18.59	19.16	**19.36**
14.08	16.54	15.75	16.72	12.72	17.07	17.26	17.93	18.50	**18.70**
13.08	16.19	14.87	16.28	12.21	16.41	16.59	17.30	17.86	**18.04**
12.08	15.78	14.01	15.83	11.71	15.77	15.94	16.65	17.23	**17.41**
11.08	15.32	13.14	15.35	11.22	15.17	15.32	16.06	16.61	**16.79**
10.08	14.79	12.15	14.86	11.18	14.16	14.66	14.89	15.46	**16.19**
9.074	14.24	11.43	14.36	10.34	13.98	14.13	14.87	15.42	**15.60**
8.076	13.60	10.58	13.84	9.950	13.47	13.57	14.34	14.87	**15.04**
7.079	12.90	9.759	13.31	9.572	12.86	13.00	13.80	14.30	**14.47**
6.079	12.18	8.947	12.83	9.202	12.35	12.47	13.29	13.77	**13.91**

Table 2. "Zaghouan" Quickbird image, performances in terms of SNR (in dB) when a white noise is added ($B = 4$).

Initial	Spatial Wiener	Spectral Wiener	Visu shrink [13]	SURE shrink [14]	Bivariate [15]	BG MAP [16]	E-SURE [3]	GMEP MAP
21.64	12.98	22.44	15.87	22.13	22.16	22.58	23.01	**23.13**
18.65	12.84	19.95	13.53	19.47	19.50	20.25	20.72	**20.86**
16.64	12.68	18.37	12.15	17.76	17.79	18.79	19.27	**19.43**
13.64	12.32	16.07	10.35	15.36	15.41	16.68	17.19	**17.39**
11.64	11.93	14.55	9.33	13.90	13.94	15.33	15.88	**16.10**
8.65	11.07	12.25	8.03	11.87	11.91	13.37	13.97	**14.20**
6.64	10.26	10.73	7.31	10.63	10.73	12.20	12.80	**13.02**

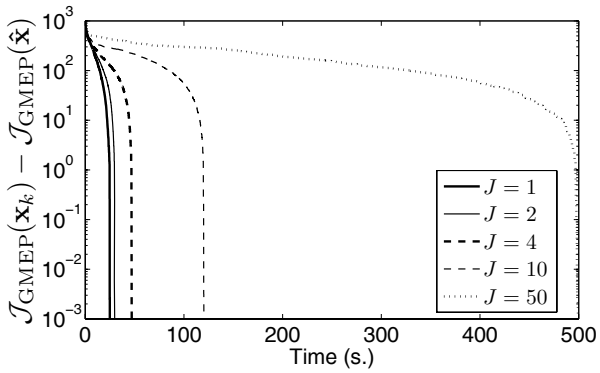
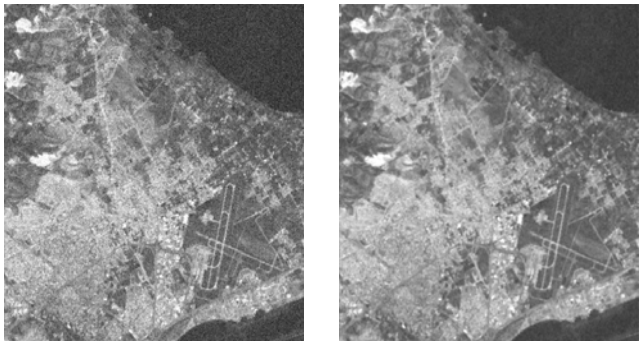


Fig. 1. Performance of 3MG algorithm in terms of criterion decrease for different values of J.



(a) (b)

Fig. 2. (a) Noisy version of the first component of "Tunis" (SNR = 6.08 dB, SSIM= 0.4387), (b) denoised version (SNR =13.91 dB, SSIM = 0.7314).

6. REFERENCES

[1] C. Lee and D. Landgrebe, "Image compression using block truncation coding," *IEEE Trans. Geosci. Remote Sens.*, vol. 317, no. 4, pp. 388–400, Apr. 1993.

[2] P. Scheunders, "Wavelet thresholding of multivalued images," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 411–416, Apr. 2004.

[3] A. Benazza-Benyahia and J.-C. Pesquet, "Building robust wavelet estimators for multicomponent images using Stein's principle," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1814–1830, Nov. 2005.

[4] A. Pižurica and W. Philips, "Estimating probability of presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 654–665, Mar. 2006.

[5] A. Duijster, P. Scheunders, and S. De Backer, "Wavelet-

based em algorithm for multispectral-image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3892 – 3898, Nov. 2009.

[6] N. Khelil and A. Benazza-Benyahia, "A wavelet-based multivariate approach for multispectral image indexing," in *SPIE Conf. on Wavelet Applications in Industrial Processing*, Philadelphia, USA, 24-28 Oct. 2004, vol. 5607.

[7] G. Verdoolaege, S. De Backer, and W. Philips, "Estimating probability of presence of a signal of interest in multiresolution single- and multiband image denoising," in *IEEE Internat. Conf. on Image Processing (ICIP)*, San Diego, CA, USA, 12-15 Oct. 2008, pp. 169–170.

[8] R. Kwitt, P. Meerwald, and A. Uhl, "Color-image watermarking using multivariate power-exponential distribution," in *IEEE Internat. Conf. on Image Processing (ICIP)*, Cairo, Egypt, 7-10 Nov. 2009, pp. 4245–4248.

[9] S. Zozor and C. Vignat, "Some results on the denoising problem in the elliptically distributed context," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 134–150, Jan. 2010.

[10] E. Chouzenoux, J. Idier, and S. Moussaoui, "A majorize-minimize subspace strategy for subspace optimization applied to image restoration," *IEEE Trans. Image Process.*, vol. 20, no. 18, pp. 1517–1528, Jun. 2011.

[11] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, and H. Talbot, "A majorize-minimize subspace approach for $l2 - l0$ image regularization," *SIAM J. Imag. Sci.*, 2011, to appear.

[12] A. Buades, B. Coll, and J. Morel, "A non-local algorithm for image denoising," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Washington, DC, USA, 2005, pp. 60–65.

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

[14] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Am. Stat. Assoc.*, vol. 90, pp. 1200–1224, 1995.

[15] L. Şendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.

[16] A. Benazza-Benyahia and J.-C. Pesquet, "Wavelet-based multispectral image denoising with bernoulli-gaussian model," in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP)*, Grado, Italy, 8-11 Jun. 2003.