# AN EXTENDED SURE APPROACH FOR MULTICOMPONENT IMAGE DENOISING

*A. Benazza-Benyahia*

Dept MASC,
Ecole Supérieure des Communications,
Tunis, Tunisia
ben.yahia@planet.tn

*J-.C. Pesquet*

IGM and ENST CNRS-UMR 5141,
Université de Marne la Vallée,
France
pesquet@univ-mlv.fr

**ABSTRACT**

Multichannel imaging systems provide several observations of the same scene which are often corrupted by additive noise. In this paper, we are interested in multispectral image denoising in the wavelet domain. We adopt a multivariate approach in order to exploit the correlations existing between the different spectral components. Our main contribution is the application of Stein's principle to build a new estimator for arbitrary multichannel images embedded in Gaussian noise. Simulation tests carried out on multispectral satellite images show that the proposed method outperforms conventional wavelet shrinkage techniques.

## 1. INTRODUCTION

Image recording/transmission systems are not perfect: most of the time, images are contaminated by noise. Hence, denoising is a crucial step before image analysis. Many denoising methods have been reported. The earlier ones are based on spatial filters or order statistic filters. The problem of noise removal may also be viewed as a regression problem, which can be solved by shrinkage methods. Usually, these are applied in a transform domain, the transformation being expected to decorrelate the input data or at least to allow a tractable statistical modelling of them. To this respect, the Wavelet Transform (WT) has been retained in many works [1] since the WT of the signal of interest mostly concentrates its energy in a few low resolution coefficients whereas the WT of the noise spreads out over all coefficients. Wavelet shrinkage is thus effective for signals with sparse representation in the WT domain. In this context, the key issue is the threshold computation. Donoho and Johnstone derived the so-called *universal* threshold from an asymptotic minimax analysis. By exploiting Stein's Unbiased Risk Estimator (SURE), they have subsequently proposed the SUREshrink method [2]. Further improvements of this pioneering work have been realized [3, 4, 5].

In parallel to these works, much attention was paid to noise reduction in multicomponent images. Indeed, in many applications such as remote sensing, imaging systems tend to have increasing spatial and spectral resolution. Operating in different spectral ranges, these instruments deliver several spectral components (corresponding to a multichannel image) for a single sensed area. Two alternatives could be envisaged for denoising multispectral images. The first one consists in processing the spectral components separately whereas the second one attempts to exploit their mutual correlation using multivariate approaches [6]. In this paper, we are interested in developing a new multivariate estimation method based on the WT. It must be pointing out that most of the work done so far about image denoising using wavelets is concerned with single-component images. The problem of multivariate estimation in a wavelet representation has however been lately formulated in [7]. Our main contribution in this paper is the application of Stein's principle [8] to build a new estimator for multichannel images embedded in Gaussian noise. Our approach relies on suitable multivariate Bernouilli-Gaussian (BG) models to reflect the sparseness of the wavelet representation as well as the statistical dependencies existing between the different components.

The paper is organized as follows. In Section 2, we present the relevant statistical background. Section 3 describes the new multivariate SURE approach proposed in this work. In Section 4, experimental results are provided so as to evaluate this method and, finally, some concluding remarks are given.

## 2. STATISTICAL BACKGROUND

### 2.1. Observation model

The "clean" multispectral image consists of $B$ spectral components $s^{(b)}$ (with $b = 1, \ldots, B$). These components are assumed to be corrupted by an additive noise and the observation vector is expressed as $\mathbf{r} = \mathbf{s}+\mathbf{n}$ where the vector noise $\mathbf{n} = (n^{(1)}, \ldots, n^{(B)})^T$ is iid $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma}^{(n)})$, independent of $\mathbf{s} = (s^{(1)}, \ldots, s^{B)})^T$. In the sequel, we treat $\mathbf{\Gamma}^{(n)}$ as known. In order to simplify the notations, the spatial index has been systematically dropped.

The orthonormal 2D separable wavelet expansion of the observations is computed for each component $b$ over $J$ resolution levels. In the vector approach, the wavelet coefficients of all the $B$ channels at the same spatial position in a subband oriented either horizontally ($o = 1$), vertically ($o = 2$) or diagonally ($o = 3$), at a given resolution level $j$, are grouped into $B$-dimensional vectors: $\mathbf{r}_j^{(o)} \triangleq (r_j^{(1,o)}, \ldots, r_j^{(B,o)})^T$, $\mathbf{s}_j^{(o)} \triangleq (s_j^{(1,o)}, \ldots, s_j^{(B,o)})^T$ and, $\mathbf{n}_j^{(o)} \triangleq (n_j^{(1,o)}, \ldots, n_j^{(B,o)})^T$. Hence, in the wavelet domain, the observation model becomes $\mathbf{r}_j^{(o)} = \mathbf{s}_j^{(o)} + \mathbf{n}_j^{(o)}$. The noise coefficients $\mathbf{n}_j^{(o)}$ have a Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_j^{(n,o)})$. Besides, it is easy to show that $\mathbf{\Gamma}_j^{(n,o)} = \mathbf{\Gamma}^{(n)}$ for $j = 1, \ldots, J$.

### 2.2. Prior Model

In the Bayesian framework, the unknown wavelet coefficients $\mathbf{s}_j^{(o)}$ are considered as realizations of random processes. Very often, Bernouilli-Gaussian (BG) priors are retained since they reflect the parcimony of the wavelet representation and they are easily tractable [9, 10, 11, 12]. More precisely, the probability distribution $p_j^{(o)}$ of

$\mathbf{s}_j^{(o)}$ is expressed as:

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

where $g_{\mathbf{0},\mathbf{\Gamma}_j^{(s,o)}}$ is the multivariate Gaussian $\mathcal{N}(\mathbf{0},\mathbf{\Gamma}_j^{(s,o)})$ probability density . The mixture parameter $\epsilon_j^{(o)}$ may be interpreted as the probability that the $\mathbf{s}_j^{(o)}$'s carry useful information. In order to avoid degenerate MAP estimates, this mixture model is coupled with hidden allocation variables $q_j^{(o)}$ which are independent binary random variables defining the following conditional densities

$$p(\mathbf{s}_j^{(o)}/q_j^{(o)}=0) = \delta(\mathbf{s}_j^{(o)}), \quad p(\mathbf{s}_j^{(o)}/q_j^{(o)}=1) = g_{\mathbf{0},\mathbf{\Gamma}_j^{(s,o)}}(\mathbf{s}_j^{(o)}), \quad (2)$$

with $\mathsf{P}(q_j^{(o)}=1) = \epsilon_j^{(o)} \in [0,1]$. Under such statistical modelling, the denoising problem amounts to a classical problem in estimation theory. In [13], we have proposed a MAP estimation procedure of $\mathbf{s}_j^{(o)}$ which is based on this Bayesian approach.

### 2.3. Motivation

However, this MAP estimation procedure relies on a BG distribution for the wavelet coefficients. In practice, this hypothesis is not rigorously satisfied. As a consequence, two main limitations may occur. At first, the model mismatch could render inappropriate the *structure* of the resulting estimator with respect to the considered data. Secondly, the hyperparameters which appear in the expression of the estimator may take suboptimal values since they are deduced from statistical inference based on the BG model assumption. Unfortunately, we cannot expect to alleviate the former problem in the absence of additional prior knowledge on the clean signal. In order to address the later problem, it would be preferable to adjust the hyperparameters directly so as to minimize the risk (e.g. the mean square error). More precisely, the Gaussianity of the noise will allow us to apply Stein's formula so as to realize this task. In the next section, this approach is presented in more detail.

## 3. EXTENDED SURE APPROACH

### 3.1. Stein's formula in the multivariate case

Let us denote by $\hat{\mathbf{s}}_j^{(o)} = T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)$ an estimator of $\mathbf{s}_j^{(o)}$ where $T_{j,\boldsymbol{\theta}_j^{(o)}}$ is a (weakly) differentiable function from $\mathbb{R}^B$ to $\mathbb{R}^B$, parameterized by a finite dimensional vector $\boldsymbol{\theta}_j^{(o)}$. It is also required that $T_{j,\boldsymbol{\theta}_j^{(o)}}$ fulfills some additional regularity conditions [8]. Our objective is to find the parameters minimizing the quadratic risk $R_j^{(o)}(\boldsymbol{\theta}_j^{(o)}) = E[\|\mathbf{s}_j^{(o)} - T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)\|^2]$. Minimizing $R_j^{(o)}$ obviously amounts to minimize the following criterion:

$$\tilde{R}_j^{(o)}(\boldsymbol{\theta}_j^{(o)}) = E[\|T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)\|^2] - 2E[(\mathbf{s}_j^{(o)})^T\, T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)]. \quad (3)$$

Unfortunately, as the samples $\mathbf{s}_j^{(o)}$ are unknown, it may appear impossible to calculate explicitly $\tilde{R}_j^{(o)}$. However, for Gaussian additive noise, Stein's formula provides an explicit expression of $E[\mathbf{s}_j^{(o)T}\, T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)]$ in terms of the observed data only. More precisely, we have:

$$\begin{aligned} E[T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)(\mathbf{s}_j^{(o)})^T] = \;& E[T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)(\mathbf{r}_j^{(o)})^T] \\ & - E[T'_{j,\boldsymbol{\theta}_j^{(o)}}(\mathbf{r}_j^{(o)})]\;\mathbf{\Gamma}_j^{(n,o)}, \end{aligned} \quad (4)$$

where $T'_{j,\boldsymbol{\theta}_j^{(o)}}$ is the square matrix of size $B$ defined as follows:

$$\forall \mathbf{u} \in \mathbb{R}^B, \quad T'_{j,\boldsymbol{\theta}_j^{(o)}}(\mathbf{u}) = \begin{pmatrix} \left(\boldsymbol{\nabla}[T_{j,\boldsymbol{\theta}_j^{(o)}}(\mathbf{u})]_1\right)^T \\ \vdots \\ \left(\boldsymbol{\nabla}[T_{j,\boldsymbol{\theta}_j^{(o)}}(\mathbf{u})]_B\right)^T \end{pmatrix} \quad (5)$$

with $T_{j,\boldsymbol{\theta}_j^{(o)}}(\mathbf{u}) = ([T_{j,\boldsymbol{\theta}_j^{(o)}}(\mathbf{u})]_1, \ldots, [T_{j,\boldsymbol{\theta}_j^{(o)}}(\mathbf{u})]_B)^T$ and the operator $\boldsymbol{\nabla}$ stands for the gradient operator. Consequently, we deduce the following expression of the risk:

$$\begin{aligned} \tilde{R}_j^{(o)}(\boldsymbol{\theta}_j^{(o)}) = \;& E[\|T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)\|^2] - 2E[(\mathbf{r}_j^{(o)})^T\, T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)] \\ & + 2\,\mathrm{tr}\left(E[T'_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)]\;\mathbf{\Gamma}_j^{(n,o)}\right) \end{aligned} \quad (6)$$

It can be noted that the risk $\tilde{R}_j^{(o)}$ is composed of two parts. The first two terms in the left hand side of (6) correspond to the error resulting from the estimation of $\mathbf{r}_j^{(o)}$ instead of $\mathbf{s}_j^{(o)}$ by $T_{j,\boldsymbol{\theta}_j^{(o)}}\left(\mathbf{r}_j^{(o)}\right)$. The third term compensates the bias related to the above confusion of $\mathbf{s}_j^{(o)}$ with $\mathbf{r}_j^{(o)}$. Finally, an empirical estimator of $\tilde{R}_j^{(o)}(\boldsymbol{\theta}_j^{(o)})$ is easily obtained by invoking the law of large numbers and replacing the expectations in (6) by empirical means.

### 3.2. Retained structure of the estimator

In what follows, we will derive a tractable structure for the estimator based on the BG model. The idea consists in replacing the MAP estimator by the *a posteriori* conditional mean which corresponds to the optimal Bayesian estimator for a quadratic risk. After some simple calculations, *in the case of the BG model*, the *a posteriori* conditional mean is expressed as:

$$E[\mathbf{s}_j^{(o)}/\mathbf{r}_j^{(o)}] = \gamma_{\epsilon_j^{(o)}}(\mathbf{r}_j^{(o)})\mathbf{Q}_j^{(o)}\mathbf{r}_j^{(o)}, \quad (7)$$

$$\text{where } \gamma_{\epsilon_j^{(o)}} \triangleq \frac{\epsilon_j^{(o)} g_{\mathbf{0},\mathbf{\Gamma}_j^{(s,o)}+\mathbf{\Gamma}_j^{(n,o)}}}{\epsilon_j^{(o)} g_{\mathbf{0},\mathbf{\Gamma}_j^{(s,o)}+\mathbf{\Gamma}_j^{(n,o)}} + (1-\epsilon_j^{(o)}) g_{\mathbf{0},\mathbf{\Gamma}_j^{(n,o)}}} \quad (8)$$

and $\mathbf{Q}_j^{(o)} \triangleq \mathbf{\Gamma}_j^{(s,o)}(\mathbf{\Gamma}_j^{(s,o)} + \mathbf{\Gamma}_j^{(n,o)})^{-1}$. In the sequel, we will retain estimators having the form given by Eqs. (7)-(8). However, in order to take into account the distance between the distribution of the image of interest and the BG model, we will *not* preserve the original meaning of the parameters $\epsilon_j^{(o)}$ and $\mathbf{Q}_j^{(o)}$. By adding degrees of freedom in the choice of these parameters, we aim at building estimators with an increased robustness w.r.t. model mismatch. More precisely, we assume that the parameter-vector $\boldsymbol{\theta}_j^{(o)}$ is formed by $\epsilon_j^{(o)}$ and the elements of $\mathbf{Q}_j^{(o)}$. Then, we propose to adjust these parameters so as to minimize the risk $\tilde{R}_j^{(o)}$ in Eq. (6).

### 3.3. Minimization of the risk

*3.3.1. Preliminary remarks*

Before describing the optimization procedure, the following points should be noted. The matrix $\mathbf{\Gamma}_j^{(s,o)}$ was not included in the parameter vector $\boldsymbol{\theta}_j^{(o)}$ as the optimization of this matrix would be too complex. Hence, an estimator of $\mathbf{\Gamma}_j^{(s,o)}$ must be used as the one provided by the method of moments described in [13]. Furthermore, the proposed method is applicable to *any* differentiable mapping $\gamma_{\epsilon_j^{(o)}}$ from $\mathbb{R}^B$ to $\mathbb{R}$, involving some parameter $\epsilon_j^{(o)}$.

*3.3.2. Closed form expression of the risk*

Introducing the following definitions:

$$\mathbf{B}_j^{(o)} \triangleq E[\gamma_{\epsilon_j^{(o)}}(\mathbf{r}_j^{(o)})\mathbf{r}_j^{(o)}(\mathbf{r}_j^{(o)})^T] - \mathbf{\Gamma}_j^{(n,o)}\big(E[\gamma_{\epsilon_j^{(o)}}(\mathbf{r}_j^{(o)})]\mathbf{I} + E[\boldsymbol{\nabla}\gamma_{\epsilon_j^{(o)}}(\mathbf{r}_j^{(o)})(\mathbf{r}_j^{(o)})^T]\big) \quad (9)$$

$$\mathbf{C}_j^{(o)} \triangleq E[\gamma^2_{\epsilon_j^{(o)}}(\mathbf{r}_j^{(o)})\mathbf{r}_j^{(o)}(\mathbf{r}_j^{(o)})^T] \quad (10)$$

and using Eq. (7), Eq. (6) becomes:

$$\tilde{R}_j^{(o)}(\boldsymbol{\theta}_j^{(o)}) = \begin{array}{l}\mathrm{tr}\big((\mathbf{Q}_j^{(o)} - \mathbf{A}_j^{(o)})\mathbf{C}_j^{(o)}(\mathbf{Q}_j^{(o)} - \mathbf{A}_j^{(o)})^T\big) \\ -\mathrm{tr}\big(\mathbf{A}_j^{(o)}\mathbf{C}_j^{(o)}(\mathbf{A}_j^{(o)})^T\big)\end{array} \quad (11)$$

where $\mathbf{A}_j^{(o)} \triangleq \mathbf{B}_j^{(o)}\big(\mathbf{C}_j^{(o)}\big)^{-1}$ and the matrix $\mathbf{C}_j^{(o)}$ is supposed to be invertible. Therefore, the problem is to find the matrix $\mathbf{Q}_j^{(o)}$ and the real $\epsilon_j^{(o)}$ which minimize the expression obtained in Eq. (11).

*3.3.3. Optimization of the risk*

We notice that for any matrix $\mathbf{M}$, $\mathrm{tr}\big(\mathbf{M}\mathbf{C}_j^{(o)}\mathbf{M}^T\big)^{1/2}$ is the Frö-benius norm weighted by the definite-positive matrix $\mathbf{C}_j^{(o)}$. Thus, setting $\mathbf{Q}_j^{(o)}$ to $\mathbf{A}_j^{(o)}$ leads to the minimization of the risk $\tilde{R}_j^{(o)}$. We replace $\mathbf{Q}_j^{(o)}$ by $\mathbf{A}_j^{(o)}$ in the expression (11) and, we obtain:

$$\tilde{R}_j(\boldsymbol{\theta}_j^{(o)}) = -\mathrm{tr}(\mathbf{B}_j^{(o)}\big(\mathbf{C}_j^{(o)}\big)^{-1}\mathbf{B}_j^{(o)T}). \quad (12)$$

Hence, it is enough to find the real $\epsilon_j^{(o)}$ within the interval $[0, 1]$ that maximizes $\mathrm{tr}(\mathbf{B}_j^{(o)}\big(\mathbf{C}_j^{(o)}\big)^{-1}\mathbf{B}_j^{(o)T})$ where $\mathbf{B}_j^{(o)}$ and $\mathbf{C}_j^{(o)}$ are given by Eqs. (9) and (10). Once the optimal value of $\epsilon_j^{(o)}$ is determined by using simple one-variable optimization search, the matrix $\mathbf{Q}_j^{(o)}$ is computed as $\mathbf{Q}_j^{(o)} = \mathbf{B}_j^{(o)}\big(\mathbf{C}_j^{(o)}\big)^{-1}$.

### 3.4. Robustness of the estimator

In very noisy environments, the contribution of the noise components to the risk may dominate, especially in case of great sparsity of the wavelet representation of the underlying signal. Therefore, in this case, the computed value of the risk may become unreliable. To alleviate such a problem, a robust estimator can be derived from a hybrid scheme. To this respect, we adopt a strategy similar to that proposed in [2]. More precisely, the wavelet coefficients of the observations $\mathbf{r}_j^{(o)}$ are considered as unreliable if the power level of the "clean" data is below a given threshold $\lambda_j^{(b)}$ and, then, they are discarded from the estimation of the risk. In this case, it is recommended to use another estimator. For instance, we can envisage to apply a MAP estimator [13] or universal thresholding.

## 4. EXPERIMENTAL RESULTS AND CONCLUSIONS

We have used 2 sets of SPOT multispectral test images of size 512 $\times$ 512. The first set ("Tunis") is a part of a SPOT3 scene depicting the city of Tunis ($B = 3$). The second set ("Kairouan") consists of a SPOT4 scene ($B = 4$) representing the rural area near the city of Kairouan. We have added realizations of a multivariate Gaussian noise and performed Monte-Carlo simulations. The performance of a denoising method is measured in terms of the final average SNR $\mathrm{snr}_{\mathrm{final}}$. All the subsequent reported results are related to the Daubechies s8 symmlet but simulations with other wavelet bases have given similar results. The proposed method (denoted E-SURE) is compared to some existing denoising methods, namely the spatial 2D Wiener filter, the spectral Wiener filter, the Visushrink [1], the SUREshrink [2], the BayesShrink [14] and the BG-MAP methods [13]. We have also made a comparison with the locally bivariate shrinkage method which is a very competitive method based on the exploitation of the dependence between coefficients across scales [15].

In the first part of the simulations, we have considered a white Gaussian noise characterized by a diagonal covariance matrix. A four-stage wavelet transform is applied ($J = 4$). Tables 1 and 2 provide the output SNR achieved by the different considered techniques. The linear methods have been discarded because they present the lowest values of SNR's: an improvement of more than 2.5 dB is obtained with the SUREshrink method compared to the spectral Wiener filter in the case of the image "Tunis". As expected, by taking into account the inter-scale relationships between the coefficients, the bivariate method outperforms most methods. However, it exhibits lower performances than the BG-MAP and the E-SURE methods. For "Tunis", the E-SURE method provides an average gain around 1.27 dB over the bivariate method.

In the second part of the simulations, we have considered a colored noise with the following covariance matrix $\mathbf{\Gamma}^{(n)}$:

$$\mathbf{\Gamma}^{(n)} = (\sigma^{(n)})^2 \begin{pmatrix} 1 & 0.9 & 0.9 \\ 0.9 & 1 & 0.9 \\ 0.9 & 0.9 & 1 \end{pmatrix}, \quad (13)$$

where $\sigma^{(n)}$ is adjusted so as to get the desired value of $\mathrm{snr}_{\mathrm{init}}$. Figure 1 shows the performances of all the considered methods. Roughly speaking, we can draw the same conclusions as in the case of a white noise.

Whatever the noise covariance matrix $\mathbf{\Gamma}^{(n)}$ is, we have noted that it is enough to stop the decomposition at the third stage since there is no significant improvement (no more than 0.02 dB) by going from a three-stage to a four-stage decomposition for both the MAP-BG and the E-SURE methods.

In conclusion, we have proposed a new multivariate approach for estimating multichannel images corrupted by Gaussian noise. According to simulations carried out on multispectral satellite images, this method leads to improved performances compared with state-of-the-art wavelet-based denoising techniques. However, it is more computationally intensive than basic thresholding methods. Several directions could be investigated to extend this work. In particular, it would be interesting to take into account interscale dependencies between spectral components in addition to intrascale

dependencies. For example, multivariate extensions of [15] could be looked for.

## 5. REFERENCES


[1] D.L. Donoho, I.M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425-455, 1994.

[2] D.L. Donoho, I.M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. of the Amer. Stat. Ass.*, vol. 90, no. 432, pp. 1200-1224, 1995.

[3] N. Weyrich, G.T. Warhola, "Wavelet shrinkage and generalized cross validation for image denoising," *IEEE Trans. on Image Proc.*, vol. 7, no. 1, pp. 82-90, Jan. 1998.

[4] H. Krim, D. Tucker, S. Mallat, D. Donoho, "On denoising and best signal representation," *IEEE Trans. on Info. Theory*, vol 5, no. 7, pp. 2225-2238, November 1999.

[5] M.K. Mihçak, I. Kozintsev, K. Ramchandran, P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Proc. Letters*, vol. 6, no. 12, pp. 300-303, December 1999.

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

[7] A. K. Fletcher, V. K. Goyal, K. Ramchandran, "On multivariate estimation by thresholding," *Proc. of the ICIP*, Barcelona, Spain, 4-17 Sept. 2003.

[8] C. Stein, "Estimation of the mean of a multivariate normal distribution," *Annals of Stat.*, vol. 10, pp. 1040-1053, 1981.

[9] H.A. Chipman, E.R. Kolaczyk, R.E. McCulloch, "Adaptive Bayesian wavelet shrinkage," *J. of the Amer. Stat. Assoc.*, vol. 92, no. 440, pp. 1413-1421, Dec. 1997.

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

[11] 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.

[12] M.S. Crouse, R.D. Nowak, R.G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. on Signal Proc.*, vol. 46, no. 4, pp. 886-902, April 1998.

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

[14] S.G. Chang, B. Yu, M. Vetterli, "Spatially adaptive wavelet thresholding with context modelling for image denoising," *IEEE Trans. on Image Proc.*, vol. 9, no. 9, pp. 1522-1531, Sep. 2000.

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


**Table 1**. "Tunis" image, performances in terms of SNR (in dB), a four-stage wavelet decomposition is used ($J = 4$) and a white noise is added.

| Initial | Visu shrink | Bayes Shrink | SURE shrink | Bi-variate | MAP BG | E-SURE |
|---|---|---|---|---|---|---|
| 15.08 | 13.27 | 17.28 | 17.75 | 17.96 | 18.65 | 19.17 |
| 14.58 | 12.99 | 16.96 | 17.40 | 17.61 | 18.27 | 18.84 |
| 14.07 | 12.72 | 16.62 | 17.06 | 17.26 | 17.93 | 18.50 |
| 13.58 | 12.46 | 16.30 | 16.72 | 16.92 | 17.61 | 18.17 |
| 13.08 | 12.20 | 15.99 | 16.39 | 16.59 | 17.28 | 17.85 |
| 12.58 | 11.94 | 15.69 | 17.14 | 16.26 | 16.98 | 17.53 |
| 12.08 | 11.70 | 15.39 | 15.77 | 15.93 | 16.66 | 17.21 |
| 11.58 | 11.46 | 15.09 | 15.46 | 15.61 | 16.34 | 16.89 |
| 11.07 | 11.22 | 14.81 | 15.14 | 15.31 | 16.05 | 16.60 |
| 10.58 | 10.99 | 14.53 | 14.84 | 15.01 | 15.75 | 16.30 |
| 10.08 | 10.77 | 14.26 | 14.53 | 14.70 | 15.45 | 15.99 |
| 9.58 | 10.55 | 13.99 | 14.25 | 14.41 | 15.16 | 15.70 |
| 9.08 | 10.34 | 13.74 | 13.98 | 14.12 | 14.88 | 15.41 |

**Table 2**. "Kairouan" image, performances in terms of SNR (in dB), a four-stage wavelet decomposition is used ($J = 4$) and a multivariate white noise is added.

| Initial | Visu shrink | Bayes Shrink | SURE shrink | Bi-variate | MAP BG | E-SURE |
|---|---|---|---|---|---|---|
| 4.93 | 6.71 | 9.42 | 10.04 | 10.23 | 11.23 | 11.95 |
| 3.92 | 6.35 | 8.94 | 9.55 | 9.68 | 10.71 | 11.42 |
| 2.93 | 6.01 | 8.50 | 9.01 | 9.16 | 10.21 | 10.91 |
| 1.93 | 5.70 | 8.08 | 8.49 | 8.66 | 9.74 | 10.42 |
| 0.93 | 5.40 | 7.67 | 8.06 | 8.17 | 9.29 | 9.95 |
| -0.06 | 5.12 | 7.28 | 7.62 | 7.70 | 8.85 | 9.48 |
| -1.06 | 4.87 | 6.93 | 7.13 | 7.25 | 8.42 | 9.04 |
| -2.06 | 4.64 | 6.58 | 6.72 | 6.81 | 8.00 | 8.60 |
| -3.06 | 4.43 | 6.25 | 6.38 | 6.39 | 7.62 | 8.18 |

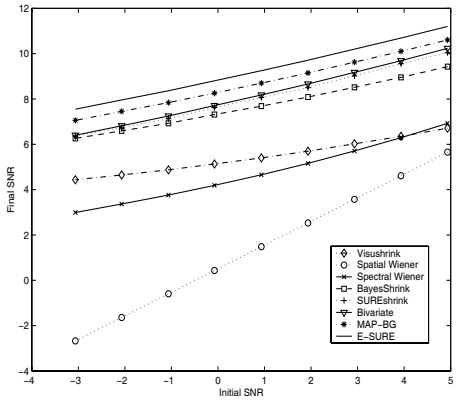


**Fig. 1**. Performances in terms of SNR (in dB) with a four-stage decomposition of "Kairouan" in the presence of colored noise.