Wavelet domain denoising of single-band and multiband images adapted to the probability of the presence of features of interest

Aleksandra Pižurica*, Wilfried Philips* and Paul Scheunders†

*Ghent University, Dept. Telecommunications and Information Processing, Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium

†University of Antwerp, Vision Lab, Groenenborgerlaan 171, B-2020, Antwerp, Belgium

ABSTRACT

We study an image denoising approach the core of which is a locally adaptive estimation of the probability that a given coefficient contains a significant noise-free component, which we call "*signal of interest*". We motivate this approach within the minimum mean squared error criterion and develop and analyze different locally adaptive versions of this method for color and for multispectral images in remote sensing. For color images, we study two different approaches: (i) using a joint spatial/spectral activity indicator in the RGB color space and (ii) componentwise spatially adaptive denoising in a luminance-chrominance space. We demonstrate and discuss the advantages of both of these approaches in different scenarios. We also compare the analyzed method to other recent wavelet domain denoisers for multiband data both on color and on multispectral images.

Keywords: Wavelets, Bayesian estimation, color, multispectral images.

1. INTRODUCTION

Current trends in image denoising demonstrate that the limits of achievable image quality are still growing, mainly due to the use of multiresolution discontinuity adaptive representations, such as wavelets [1–8], and their successors, like steerable pyramids [9], contourlets [10], curvelets [11, 12], bandelets [13], ... Another interesting research track is development of hardware-friendly denoising methods that achieve high denoising quality, but with a complexity that is suitable for implementations in field programmable gate arrays (FPGA) [14].

Evolution of imaging technology brings an increased use of *multivalued* or *multi-band* images in different application areas. Classical examples are color images. In medical applications, multivalued magnetic resonance images (MRI) are often acquired using different imaging parameters (e.g., different proton density or diffusion, different modes: T1 or T2, ...) Other medical imaging examples include different computer tomography (CT) and nuclear medicine modalities. In remote sensing, the use of multispectral and hyperspectral imagery has proved advantageous for processing tasks such as classification, object recognition and content analysis. For such tasks, image denoising is often a crucial preprocessing step. Compared to greyscale image denoising, much less research has been done on multiresolution denoising of multiband images. Several recent representatives include multiband wavelet thresholding methods [15, 16] and vector-based least square estimators [17, 18].

In this paper we develop a Bayesian wavelet domain image denoising method, which adapts itself to the local spatial and spectral image context. We motivate the proposed estimator by the minimum mean squared error criterion (MMSE) and we take into account uncertainty about the presence of a signal of interest in a noisy observation. This work builds on our recent denoising approach [19], by extending it within the MMSE criterion and by developing it further for multivalued images. In particular, we develop a color image denoising method in the luminance-chrominance space and we show its advantage over the previous RGB-version of the method from [19]. We also give a more elaborate analysis of remote multispectral denoising in remote sensing, where we compare the proposed method with the latest representative multiresolution methods from the literature.

Aleksandra Pizurica (Aleksandra.Pizurica@telin.ugent.be.) is postdoctoral research fellow of FWO (Funds for scientific research in Flanders), Belgium.

The paper is organized as follows. Section 2 describes so-called *ProbShrink* estimator of [19], where we first introduce the assumed noise model and notations (Sec. 2.1) and then we present the prior used (Sec. 2.2) and the three versions of the estimator: subband-adaptive (Sec. 2.3), spatially-adaptive (Sec. 2.4) and spatially-spectrally adaptive estimator (Sec. 2.5). In Section 3, we give a theoretical motivation for this estimator in terms of the minimum mean squared error criterion. Section 4 addresses multispectral image denoising in remote sensing, where we analyze the application of the proposed method to *Landsat* satellite images. Section 5 is devoted to color image denoising, where we analyze two versions of the proposed method: applied in the RGB space (Sec. 5.1) and applied in a luminance-chrominance space (Sec. 5.2). Section 6 concludes the paper.

2. A SHRINKAGE ESTIMATOR BASED ON THE PROBABILITY OF SIGNAL PRESENCE

2.1. Noise model and notation

We assume the input multi-band image is contaminated with additive white Gaussian noise. Let $\beta_{l,s}^b$ denote the noise-free value of the coefficient from the image band b, the wavelet subband s and the spatial position l. The corresponding noisy coefficient value is

$$y_{l,s}^b = \beta_{l,s}^b + \epsilon_{l,s}^b \tag{1}$$

where $\epsilon_{s,l}^b$ are normal random variables $\epsilon_{s,l}^b \sim N(0, \sigma^2)$. For the sake of compactness, in the remainder we suppress the indices of the wavelet coefficients unless in cases where they are explicitly needed.

The probability density function of the noise-free wavelet coefficients in each wavelet subband is well modelled by the *generalized Laplacian* (also called *generalized Gaussian*) density [20, 21]

$$f(\beta) = \frac{\lambda\nu}{2\Gamma(\frac{1}{\nu})}\exp(-\lambda|\beta|^\nu) \tag{2}$$

where $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the Gamma function, $\lambda > 0$ is the *scale* parameter and ν is the *shape* parameter (typically in the range $\nu \in [0, 1]$). The parameters λ and ν are reliably estimated from the noisy coefficient histogram according to the method in [21]. In particular, the variance σ_y^2 and the fourth moment $m_{4,y}$ of the generalized Laplacian signal corrupted by additive white Gaussian noise with standard deviation σ are [21]

$$\sigma_y^2 = \sigma^2 + \frac{\Gamma(\frac{3}{\nu})}{\lambda^2\Gamma(\frac{1}{\nu})}, \quad m_{4,y} = 3\sigma^4 + \frac{6\sigma^2\Gamma(\frac{3}{\nu})}{\lambda^2\Gamma(\frac{1}{\nu})} + \frac{\Gamma(\frac{5}{\nu})}{\lambda^4\Gamma(\frac{1}{\nu})} \tag{3}$$

From the above equations, we find

$$\kappa_y = \frac{\Gamma(\frac{1}{\nu})\Gamma(\frac{5}{\nu})}{\Gamma^2(\frac{3}{\nu})} = \frac{m_{4,y} + 3\sigma^4 - 6\sigma^2\sigma_y^2}{(\sigma_y^2 - \sigma^2)^2} \tag{4}$$

and

$$\lambda = \left((\sigma_y^2 - \sigma^2)\frac{\Gamma(\frac{1}{\nu})}{\Gamma(\frac{3}{\nu})}\right)^{-\frac{1}{2}}. \tag{5}$$

The expression $\Gamma(\frac{1}{\nu})\Gamma(\frac{5}{\nu})/\Gamma^2(\frac{3}{\nu})$ in the left hand side of (4) is a monotonic decreasing function of ν, which simplifies the numerical computation.

A special case of the prior (2) with $\nu = 1$ also called *Laplacian* or *double exponential*: $f(\beta) = (\lambda/2)\exp(-\lambda|\beta|)$ is often used for simplicity and because it usually does not incur a significant performance loss in denoising and compression [22]. The scale parameter is in this case simply estimated as $\lambda = [0.5(\sigma_y^2 - \sigma^2)]^{-1/2}$. We assume the input noise standard deviation σ in each image channel is known; otherwise it is reliably estimated by the median absolute deviation of the coefficients in the highest frequancy subband, divided by 0.6745 [23].

2.2. A Bernoulli-Laplacian mixture prior

Our assumed prior on noise-free coefficients belongs to a class of *mixture priors* that are superpositions of two distributions, where one distribution models the statistics of "significant" ("high energy") coefficients and the other one models the statistics of "non-significant" coefficients. The mixing parameter is a *Bernoulli* random variable. Related priors are, e.g., mixtures of two Gaussians [24], a mixture of a Gaussian and point mass at zero [18, 25–27] and mixtures of a Laplacian and point mass at zero [22, 28]. Denoting the hypothesis "signal component appears in the observed coefficient with significant energy" by H_1 and the opposite hypothesis by H_0, unifies this class of mixture priors as

$$f(\beta) = P(H_0)f(\beta|H_0) + P(H_1)f(\beta|H_1). \tag{6}$$

Our prior, which also appears in [19, 29, 30], is a mixture of two truncated (generalized) Laplacians. In particular, we start from the genealized Laplacian model $f(\beta)$ from (1) and we define

$$H_0 : |\beta| \leq T \quad \text{and} \quad H_1 : |\beta| > T \tag{7}$$

where T is a threshold which defines the signal of interest. This model yields

$$f(\beta|H_0) = \begin{cases} B_0 \exp(-\lambda|\beta|^\nu), & \text{if } |\beta| \leq T \\ 0, & \text{if } |\beta| > T \end{cases} \tag{8}$$

and

$$f(\beta|H_1) = \begin{cases} 0, & \text{if } |\beta| \leq T \\ B_1 \exp(-\lambda|\beta|^\nu), & \text{if } |\beta| > T \end{cases} \tag{9}$$

with $B_0 = \lambda\nu\Big[2\Gamma(\frac{1}{\nu})\Gamma_{inc}\Big((\lambda T)^\nu, \frac{1}{\nu}\Big)\Big]^{-1}$ and $B_1 = \lambda\nu\Big[2\Gamma(\frac{1}{\nu})\Big[1 - \Gamma_{inc}\Big((\lambda T)^\nu, \frac{1}{\nu}\Big)\Big]\Big]^{-1}$ (for details, see [19]).

2.3. Subband adaptive estimator

A subband-adaptive estimator applies a unique shrinkage function to all the coefficients of a given subband. In [19], a subband-adaptive *ProbShrink* estimator is defined as

$$\hat{\beta} = P(H_1|y)y = \frac{\eta\mu}{1 + \eta\mu} y \tag{10}$$

where $\eta = f(y|H_1)/f(y|H_0)$, $\mu = P(H_1)/P(H_0)$ and the product $\mu\eta$ is called the generalized likelihood ratio [31]. For the assumed noise model, $f(y|H_{0,1})$ is the result of convolving $f(\beta|H_{0,1})$ with the normal density $N(0, \sigma^2)$. For the generalized Laplacian prior, $P(H_1) = 1 - \Gamma_{inc}((\lambda T)^\nu, 1/\nu)$ and for the Laplacian prior ($\nu = 1$) this reduces to $P(H_1) = \exp(-\lambda T)$. It was experimentally shown that this estimator yields a better mean squared error performance compared to the subband-adaptive Bayesian thresholding *BayesShrink* of [32] and compared to MMSE estimator under the normal mixture prior [24].

2.4. Spatially adaptive estimator

A spatially adaptive extension of the estimator (10) uses a *local spatial activity indicator* (LSAI) z_l, as follows:

$$\hat{\beta}_l = P(H_1|y_l, z_l)y_l = \frac{\eta_l\xi_l\mu}{1 + \eta_l\xi_l\mu} y_l \tag{11}$$

where

$$\eta_l = \frac{f(y_l|H_1)}{f(y_l|H_0)}, \quad \xi_l = \frac{f(z_l|H_1)}{f(z_l|H_0)} \quad \text{and} \quad \mu = \frac{P(H_1)}{P(H_0)}. \tag{12}$$

We define LSAI as averaged coefficient magnitude within a small window $\partial(l)$ of size N: $z_l = (1/N)\sum_{k\in\partial(l)} \omega_k$, with $\omega_l = |y_l|$, and simplify that all the coefficients within this small window are equally distributed and *conditionally* independent given H_0 or H_1. In this case, $f(Nz_l|H_{0,1})$ equals N convolutions of $f(\omega_l|H_{0,1})$ with itself, where $f(\omega_l|H_{0,1}) = 2f(y_l|H_{0,1})$ for $\omega_l \geq 0$ and $f(\omega_l|H_{0,1}) = 0$ for $\omega_l < 0$. For window sizes 5×5 or 7×7, this estimator outperforms the subband-adaptive version (10) by up to 1dB in peak signal to noise ratio (PSNR) [19]. Visual improvement is significant: flat areas are better smoothed and sharpness is improved. In the remainder, we denote this spatially adaptive denoiser as *ProbShrink-SP*.

2.5. Spatially and spectrally adaptive estimator

For multiband images, a combined spatial and spectral version of LSAI from [19] is

$$z_{l,s}^b = \frac{1}{NB} \sum_{i=1}^{B} \sum_{k \in \delta(l)} \omega_{k,s}^i \tag{13}$$

where B is the number of image bands. With this definition of LSAI the probability of signal presence is conditioned on the spatial context as well as on information from other image bands. Hence, denoising is performed band-per-band, but taking into account the inter-band correlations. In [19], the neighborhood $\delta(l)$ was reduced to a single pixel, i.e., LSAI included only the coefficients at the same spatial position from different image bands. Its conditional densities were estimated by convolving the corresponding densities of the coefficient magnitudes. It was shown experimentally that this method outperforms multiband wavelet thresholding [16] on multispectral *Landsat* and on color images. In the remainder, we denote this Multi-Band (MB) denoiser as *ProbShrink-MB*.

3. MOTIVATION IN TERMS OF THE MMSE CRITERION

The minimum mean squared error estimate under the mixture prior (6) is

$$E(\beta|y) = E(\beta|y, H_0)P(H_0|y) + E(\beta|y, H_1)P(H_1|y). \tag{14}$$

The subband-adaptive *ProbShrink* estimator (10) is an approximation of the MMSE estimate above, where $E(\beta|y, H_1)$ is approximated by one and $E(\beta|y, H_0)P(H_0|y)$ is approximated by zero. If we use the exact expressions for $E(\beta|y, H_0)$ and $E(\beta|y, H_1)$ instead, we can directly upgrade this estimator within the MMSE error criterion. Similarly, for the spatially/spectrally adaptive case, the corresponding MMSE estimator is

$$E(\beta|y_l, z_l) = E(\beta|y, H_0)P(H_0|y_l, z_l) + E(\beta|y, H_1)P(H_1|y_l, z_l). \tag{15}$$

which extends the *ProbShrink* estimator from Sec. 2.3 and Sec.2.4 Sec.2.5 within the MMSE criterion. The conditional means are found as

$$E(\beta|y, H_0) = \int_{-\infty}^{\infty} \beta f(\beta|y, H_0) d\beta \tag{16}$$

$$E(\beta|y, H_1) = \int_{-\infty}^{\infty} \beta f(\beta|y, H_1) d\beta \tag{17}$$

with $f(\beta|y, H_0)$ and $f(\beta|y, H_1)$ given in (8) and (9), respectively. In practical computations, numerical integration should be employed, while for the Laplacian prior (with $\nu = 1$) analytical expressions also exist [33]. As we demonstrate next, the performance loss of the *ProbShrink* estimator with respect to the full MMSE estimator is usually negligible in practice.

Table 1 lists experimental PSNR results for several shrinkers under two mixture priors: the analyzed Bernoulli-Laplacian (B-L) prior and the *two-normals* mixture of [24], that we denote here as Bernoulli-Gaussian (B-G): $f(\beta) = P(H_1)\phi(\beta; \sigma_1) + P(H_0)\phi(\beta; \sigma_2)$. A standard wavelet denoising procedure was applied: the noisy images were decomposed using a five-level orthogonal wavelet transform with the Daubechies' wavelet *symmlet* with eight vanishing moments [1]. In each wavelet subband the noise-free wavelet coefficients were estimated using the corresponding shrinkers for each of the analyzed methods. The inverse wavelet transform was applied to reconstruct the denoised image. The parameters of both priors were calculated adaptively for each subband. For the B-L prior we used $T = \sigma$ and for the B-G prior, we used the default choices of the hyperparameters from [24].

Several conclusions can be drawn from the results in Table 1. In most cases, *ProbShrink* yields a minor decrease in PSNR compared to the MMSE estimator under the same, Bernoulli-Laplacian prior. The MAP estimator under the B-L prior yields slightly higher performance loss (up to 0.3dB) with respect to the corresponding MMSE estimate. The results also demonstrate that the analyzed Bernoulli-Laplacian prior yields a significant improvement in PSNR with respect to the two-normals mixture, which is in most cases higher than 1dB.

Table 1. Experimental PSNR[dB] results of several *subband-adaptive* Bayesian wavelet shrinkers and two different mixture priors: *Bernoulli-Laplacian* (B-L) and *Bernoulli-Gaussian* (B-G) prior from [24]. Wavelet transform: orthogonal with five decomposition levels and using Daubechies' least asymetrical wavelet with eight vanishing moments.

		Standard deviation of noise			
Prior	Estimator	10	15	20	25
		BOAT			
	noisy image	28.15	24.62	22.10	20.17
B-L	MAP	32.00	29.92	28.42	27.37
B-L	*ProbShrink*	32.13	30.05	28.63	27.63
B-L	MMSE	**32.23**	**30.14**	**28.70**	**27.70**
B-G [24]	MMSE	31.01	29.04	27.66	26.68
		COUPLE			
	noisy image	28.15	24.60	22.11	20.18
B-L	MAP	31.65	29.26	27.89	26.95
B-L	*ProbShrink*	31.73	29.53	28.20	27.18
B-L	MMSE	**31.81**	**29.61**	**28.29**	**27.32**
B-G [24]	MMSE	30.48	28.36	27.15	26.27
		LENA			
	noisy image	28.13	24.60	22.12	20.16
B-L	MAP	33.43	31.47	30.18	29.23
B-L	*ProbShrink*	**33.69**	**31.70**	**30.39**	**29.41**
B-L	MMSE	33.62	31.65	**30.39**	**29.41**
B-G [24]	MMSE	32.84	30.97	29.66	28.74

4. MULTISPECTRAL IMAGE DENOISING IN REMOTE SENSING

In the satellite systems it may be desirable to perform denoising before the image compression step in order to improve the compression efficiency. Also deconvolution of satellite images is often useful and can be performed as an inverse filtering operation followed by denoising [34]. There are several noise sources in optical satellite images (photonic noise, electronic noise, quantization errors...) and the additive Gaussian noise model is a realistic approximation [16, 35, 36]. We test the performance of the proposed *ProbShrink-MB* method on two different multispectral images shown in Fig. 1.

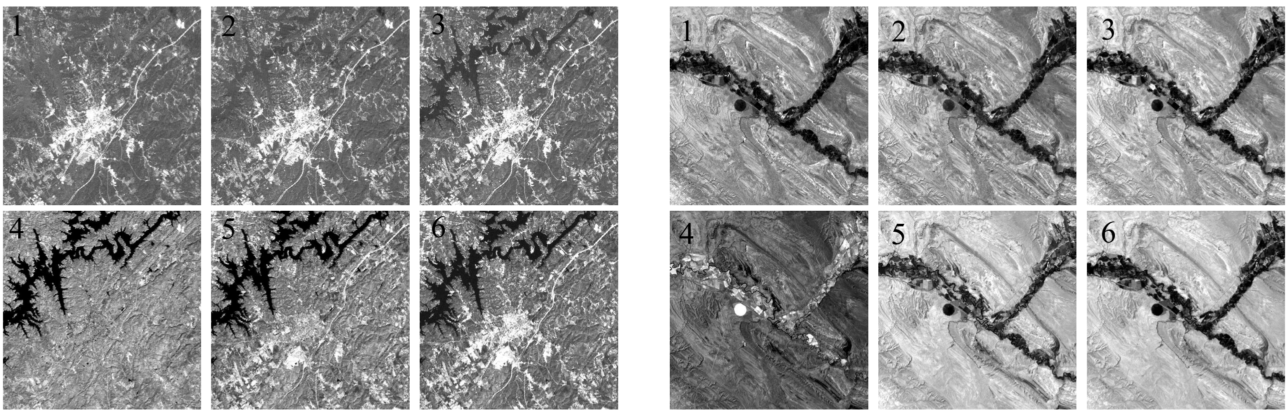

Figure 1. Noise-free test images. Left: multispectral image *MULSC1*. Right: multispectral image *MULSC2*.

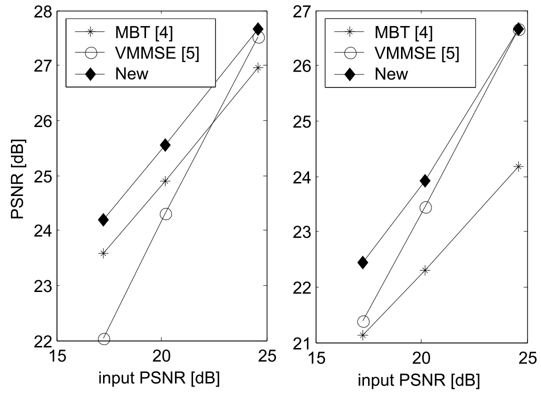

Figure 2. Quantitative performance of the proposed method for band 1 and band 4 from the multispectral image *MULSC1*.

The left-hand multispectral image in Fig. 1 is a Thematic Mapper image from the Huntsville area, Alabama, USA. It contains a river and builded area with several roads. The right-hand test image contains a builded stroke following a river. In both of these images some bands are mutually highly correlated (e.g., bands 1, 2 and 3), while other bands (like band 4) show significantly less correlation with the remaining ones. For our denoising experiments, we added artificial white Gaussian noise to the reference images from Fig. 1, using three different standard deviations of noise: 15, 25 and 35.

Let $\mathbf{u} = [\mathbf{u}^1 ... \mathbf{u}^B]$ denote the reference noise-free multiband image, where each image band is presented as a one-dimensional vector $\mathbf{u}^b = [u_1^b ... u_L^b]$ obtained according to raster scanning. Also, let $\mathbf{g} = \mathbf{u} + \mathbf{n}$ denote the noisy multi-band image, where $\mathbf{n}$ is added white Gaussian noise, and let $\hat{\mathbf{u}}$ denote the estimated noise-free image (i.e., the image denoised by wavelet shrinkage). As a quantitative performance measure we use the peak signal to noise ratio defined as $PSNR = 10log_{10}(255^2/MSE)$ where the mean squared error MSE is averaged over all the spectral channels. The input and the output PSNR are thus calculated respectively as

$$PSNR_{input} = 10log_{10}\frac{255^2}{\frac{1}{LB}\sum_{l=1}^{L}\sum_{b=1}^{B}(g_l^b - u_l^b)^2}$$

$$PSNR_{res} = 10log_{10}\frac{255^2}{\frac{1}{LB}\sum_{l=1}^{L}\sum_{b=1}^{B}(\hat{u}_l^b - u_l^b)^2}.$$

As reference methods we use multiband wavelet thresholding (MBT) [16], and a vector based wavelet domain MMSE estimator [17]. All the methods were implemented using the same *non-decimated* wavelet transform with four decomposition levels and with the Daubechies' wavelet of length four. Table 2 lists peak signal to noise ratio (PSNR) values for the two complete multispectral images from Fig. 1, and Fig. 2 displays the PSNR values calculated over two separate spectral bands. In both cases, the results demonstrate that *ProbShrink-MB* outperforms the two reference ones in terms of PSNR. For some image bands the PSNR gain of the proposed method exceeds 2dB (see Fig. 2). Pictures in Fig. 3 demonstrate also advantage of the proposed method in terms of visual quality. In comparison to multiband thresholding [16], *ProbShrink-MB* preserves better image details and fine structures, while better suppressing noise in flat areas than the vector based MMSE method of [17].

We also analyzed PSNR gains per band. The input and the resulting PSNR in a given image band b, are

$$PSNR_{input}^b = \frac{255^2}{\frac{1}{L}\sum_{l=1}^{L}(g_l - u_l)^2}, \tag{18}$$

$$PSNR_{res}^b = \frac{255^2}{\frac{1}{L}\sum_{l=1}^{L}(\hat{u}_l - u_l)^2}, \tag{19}$$

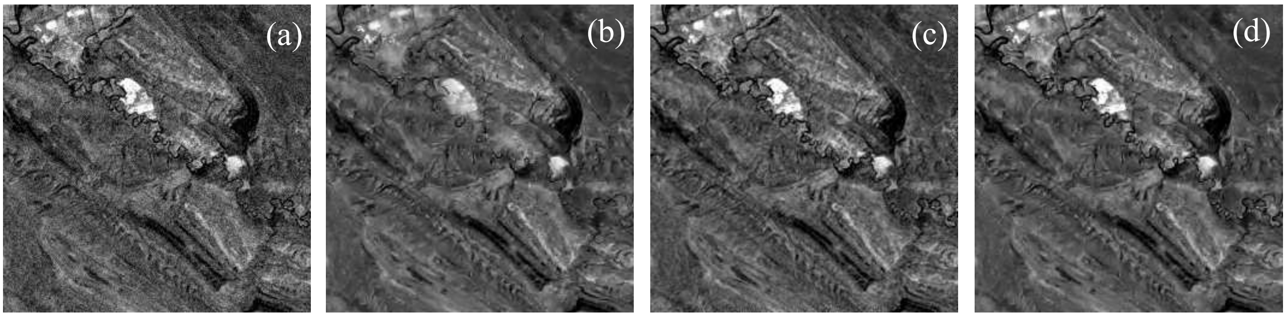

Figure 3. (a) Noisy detail of the 4th band of image *MULSC1*, standard deviation 15. (b) The result of the *MBT* method [16]. (c) The result of the *VMMSE* method [17]. (d) The result of the proposed *ProbShrink-MB* method.

Table 2. PSNR[dB] results for two multispectral images in comparison to multiband wavelet thresholding (MBT) [16] and a wavelet domain vector MMSE estimator (VMMSE) [17]

.

	Standard deviation of noise		
Estimator	15	25	35
	image MULSC1		
MBT [16]	26.63	24.37	22.95
VMMSE [17]	26.78	24.70	22.79
ProbShrink-MB	**27.42**	**24.96**	**23.45**
	image MULSC2		
MBT [16]	28.51	26.45	25.10
VMMSE [17]	**29.30**	26.25	24.16
ProbShrink-MB	29.24	**27.03**	**25.63**

respectively. Denoising yields in each band a PSNR gain

$$G^b_{PSNR} = PSNR^b_{res} - PSNR^b_{input} \tag{20}$$

and the PSNR gain summed over all the bands is

$$G_{PSNR} = \sum_{b=1}^{M} G^b_{PSNR}. \tag{21}$$

Compared to the single-band spatially adaptive version of the same method, the multi-band *ProbShrink-MB* method achieves an improved performance on all image bands (see Table 3). Table 3 also shows that in comparison with the *MBT* method of [16] the new method yields a larger total PSNR gain on both multispectral images and for all noise levels. The improvement of the new method wrt. *MBT* is the biggest in case of weakly-to-moderately correlated bands (like band 4 from the two tested multi-band images). This behavior is expected because the *MBT* method of [16] applies the same modification to the corresponding wavelet coefficients from *all* image bands and is thus powerful in case of highly correlated bands, but is less efficient in case of discontinuities which are not present in the majority of bands. This is nicely demonstrated by visual results in Fig. 4: for the band 1 from *MULSC1* image the two methods yield visually similar results. However, in the band 4 from *MULSC2* (the bottom part of Fig. 4) the *MBT* method destroys a river line (because it is not well visible in other bands) while the new method preserves it well.

The new multi-band denoising method performs well both in case of the highly correlated and in case of weakly correlated bands, because it takes into account both the inter-band correlation and the local statistics of different image bands. The visual results in Fig. 5 also show that the proposed method preserves well details and image texture.

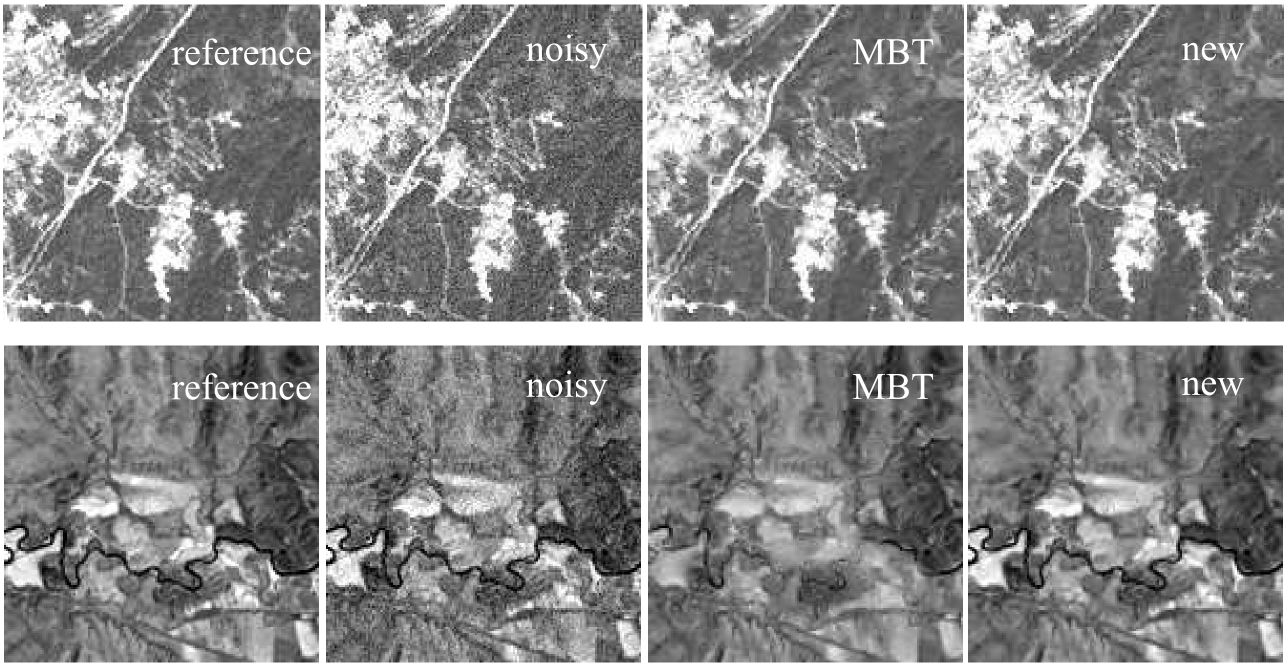

Figure 4. Left to right: parts of noise-free bands, noisy versions $\sigma = 15$, results of *MBT* [16] and the new method *ProbShrink-MB*. Top: band 1 of *MULSC1*. Bottom: band 4 of *MULSC2*. Note the river line in the bottom case.

Table 3. PSNR[dB] gain summed over all image bands.

	Standard deviation of noise		
Estimator	15	25	35
	image MULSC1		
MBT [16]	13.39	25.93	34.89
ProbShrink-SP	14.33	24.98	33.33
ProbShrink-MB	**16.79**	**27.53**	**35.52**
	image MULSC2		
MBT [16]	24.01	38.03	47.35
ProbShrink-SP	24.10	36.66	45.87
ProbShrink-MB	**26.73**	**39.25**	**48.09**

5. COLOR IMAGE DENOISING

One approach to color image denoising is using RGB color space and applying the same approach as discussed in the previous Section. In this Section, we show that such an approach is powerful in case where noise level is equal in all three color channels but is less suitable in case where noise statistics varies from channel to channel. In the latter case better results are achieved by using a luminance-chrominance color space.

5.1. Denoising in the RGB color space

The multiband denoising method from Sec. 2.5 that we named *ProbShrink-MB* assumes uncorrelated noise of *equal variance* in all image channels. Table 4 compares the performance of this method to two reference wavelet based multiband denoisers and to two other versions of the proposed denosing method. The reference methods are multiband wavelet thresholding (MBT) [16] and vector based minimum mean squared error estimator (VMMSE) [17]. Two other versions of the proposed denoiser listed in Table 4 are: *ProbShrink-SP* - spatially

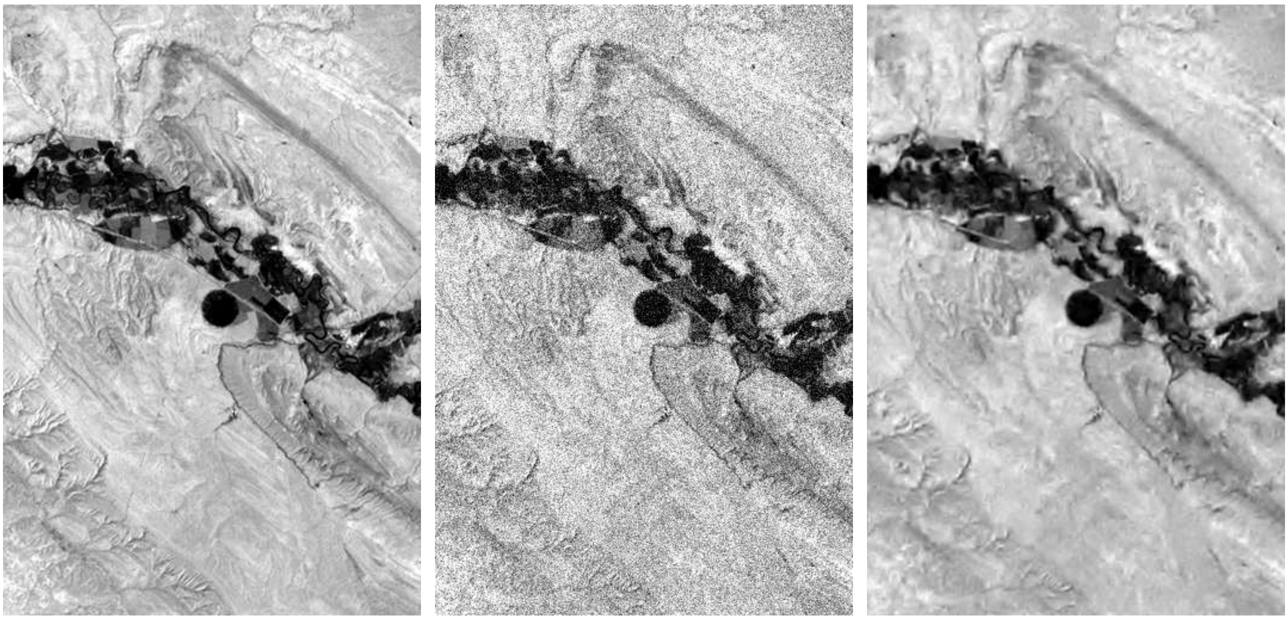

Figure 5. Left to right: noise-free, noisy and denoised band 6 from *MULSC2* using *ProbShrink-MB*.

adaptive method from Section 2.4 applied to each of the RGB channels separately and *ProbShrink-SP-YUV* the same spatially adaptive method but applied in a luminance-chrominance space, as detailed in the next Section. We implemented all the methods with a *non-decimated* wavelet transform, with four decomposition levels and with the Daubechies' wavelet of length four.

The results in Table 4 demonstrate that for uncorrelated noise of equal variance in each channel the proposed approach yields best performance when applied with the local spatial/spectral activity indicator from Section 2.5 in the RGB space. This version of the method, *ProbShrink-MB*, outperforms separate spatially adaptive channel denoising both in the RGB space (*ProbShrink-SP*) and in the luminance-chrominance space *ProbShrink-SP-YUV*. The proposed method compares favorably to the reference ones [16, 17]. On a very textured image *Baboon*, the vector based VMMSE method [17] yielded slightly better PSNR then *ProbShrink-MB*, but for other test images *ProbShrink-MB* yields an improvement of up to almost 2dB, which increases with the increase of noise level. The same behavior was noted on *Landsat* multispectral images in Section 4 (see Fig. 2).

5.2. Denoising in the luminance-chrominance space

Television broadcasting uses color spaces in terms of one luminance and two chrominance components. A standard PAL broadcasting system defines YUV model, where Y stands for the luminance component (the brightness) and U and V are the chrominance (color) components. YUV signals are created from an original RGB (red, green and blue) source. The weighted values of R, G and B are added together to produce a single Y signal, representing the overall brightness, or luminance, of that spot: $Y = 0.299R + 0.587G + 0.114B$. Two chrominance signals are formed as: $U = 0.492(B - Y) = -0.147R - 0.289G + 0.436B$, and $V = 0.877(R - Y) = 0.615R - 0.515G - 0.100B$. In a matrix form

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{22}$$

and the inverse is

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.396 & -0.581 \\ 1 & -2.029 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix}. \tag{23}$$

Table 4. PSNR[dB] results for color 512x512 images with *equal* noise levels in RGB channels. $ProbShrink - MB$ (multiband denoising) and $ProbShrink - SP$ (componentwise denoising using the spatially adaptive single band estimator) are implemented in the RGB space and $ProbShrink - SP - YUV$ in a luminance-chrominance space. All the methods are implemented using a non-decimated wavelet transform with four decomposition levels and Daubechies' wavelet of length four.

	Standard deviation of noise			
Method	10	15	20	25
	BABOON			
MBT [16]	28.50	26.78	25.53	24.56
VMMSE [17]	**30.68**	**28.24**	**26.63**	**25.36**
ProbShrink-MB	30.17	27.83	26.38	25.27
ProbShrink-SP	29.80	27.34	25.79	24.60
ProbShrink-SP-YUV	28.17	26.80	25.70	24.84
	LENA			
MBT [16]	33.84	32.29	31.14	30.15
VMMSE [17]	34.02	31.89	30.24	28.88
ProbShrink-MB	**34.60**	**33.03**	**31.92**	**31.04**
ProbShrink-SP	34.19	32.46	31.22	30.29
ProbShrink-SP-YUV	34.21	32.90	31.87	31.01
	PEPPERS			
MBT [16]	31.19	30.22	29.45	28.77
VMMSE [17]	33.12	31.13	29.67	28.45
ProbShrink-MB	**33.44**	**32.05**	**31.12**	**30.35**
ProbShrink-SP	33.20	31.65	30.61	29.75
ProbShrink-SP-YUV	32.56	31.37	30.60	29.40

Table 5. PSNR[dB] results for images with *nonequal* noise level in R, G and B channels: $\sigma_R = 10$, $\sigma_G = 20$ and $\sigma_B = 35$ for three versions of the proposed method from Table 4.

BABOON		LENA		PEPPERS	
Input	20.54	*Input*	20.54	*Input*	19.18
ProbShrink-MB	23.73	*ProbShrink-MB*	30.32	*ProbShrink-MB*	28.17
ProbShrink-SP	24.85	*ProbShrink-SP*	30.98	*ProbShrink-SP*	**29.66**
ProbShrink-SP-YUV	**25.47**	*ProbShrink-SP-YUV*	**31.76**	*ProbShrink-SP-YUV*	29.53

We investigate componentwise denoising in the YUV space by applying the spatially adaptive method from Sec. 2.4 to each channel separately, and we denote this method as *ProbShrink-SP-YUV*. The results in Table 4 demonstrate that for uncorrelated noise of equal variance in each of the RGB channels, this approach is inferior to multiband denoising in the RGB space. This is also confirmed by the visual results in Fig. 6. However, as we show next, the performance comparison is different in case when noise variances in R,G and B channels are not equal to each other.

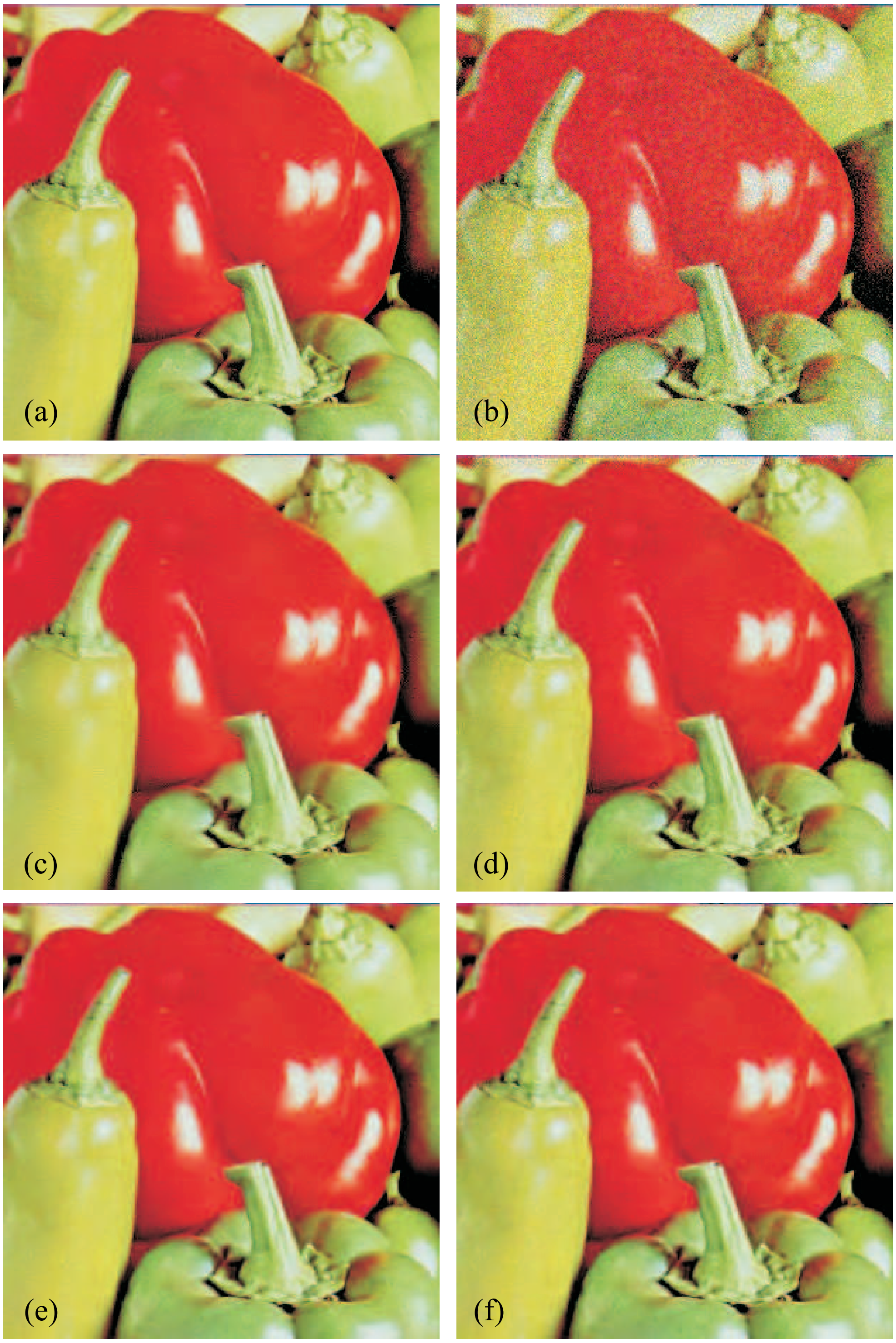

Figure 6. (a) A part of the noise-free image *Peppers*. (b) image with additive white Gaussian noise noise $\sigma_R = \sigma_G = \sigma_B = 25$. (c) *MBT* method [16] (PSNR=28.77dB). (d) *VMMSE* method [17] (PSNR=28.45dB). (e) *ProbShrink-MB* method (PSNR=30.35dB). (f) *ProbShrink-SP-YUV* method (PSNR=29.40dB).

Figure 7. Top left: noise-free image. **Top right**: image with additive white Gaussian noise noise σ_R=55, σ_G=25 and σ_B = 10. (PSNR=17.17dB). **Bottom left**: *ProbShrink-MB* (PSNR=25.67dB). **Bottom right**: *ProbShrink-SP-YUV* (PSNR=29.54dB).

Now we experiment with *nonequal* noise variance in R, G and B channels. In particular, we form noisy images by adding white Gaussian noise with standard deviations $\sigma_R = 10$, $\sigma_G = 20$ and $\sigma_B = 35$ to R, G and B channels, respectively. Table 5 lists the performance of three methods: a multiband spectrally adaptive denosier in the RGB space (*ProbShrink-MB*), componentwise spatially adaptive denoising in the RGB space (*ProbShrink-SP*) and componentwise spatially adaptive denoising in the YUV space (*ProbShrink-SP-YUV*). Table 5 shows that for nonequal noise levels, *ProbShrink-MB* shows the worst performance among the three methods. On one out of three tested images componentwise denoising in the RGB space yielded slightly better results compared to denoising in the YUV space, but for other two images YUV space offered a significant improvement over the RGB space for denoising. A poor performance of the *ProbShrink-MB* method in this case can be explained by the fact that the information from “bad” (noisier) channels is used to denoise the less noisy ones. Hence, in case of different noise levels in different channels a direct application of *ProbShrink-MB* is not suitable, but a smarter extension is sought, where for example spectral information from less noisy channels is used to denoise the noisier ones, but not viceversa. A simpler alternative is spatially adaptive componentwise denoising in a luminance chrominance space, as demonstrated in this Section. This approach performs well both in terms of PSNR and visually, as demonstrated in Fig. 7.

6. CONCLUSION

We analyzed a Bayesian wavelet domain denoising method for single- and for multiband images, which adapts itself to the probability of the presence of a signal of interest at each position. In particular, we motivated the approach of [19] within the minimum mean squared error criterion and we developed further and analyzed different variants of this estimator for spatially and spectrally adaptive denoising of multispectral and color images. Our results demonstrated that for color images with unequal noise in RGB channels the use of the color space transformation is crucial. In this case, the analyzed approach should be implemented in a luminance-chrominance space to achieve the best denoising results.

REFERENCES

1. I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia: SIAM, 1992.
2. S. Mallat, *A wavelet tour of signal processing*, Academic Press, London, 1998.
3. M. Malfait and D. Roose, “Wavelet-based image denoising using a Markov random field a priori model,” *IEEE Trans. Image processing* **6**, pp. 549–565, Apr. 1997.
4. S. G. Chang, B. Yu, and M. Vetterli, “Spatially adaptive wavelet thresholding with context modeling for image denoising,” in *Proc. IEEE Internat. Conf. on Image Proc.*, (Chicago, IL, Oct. 1998.).
5. M. K. Mihçak, I. Kozintsev, K. Ramchandran, and P. Moulin, “Low-complexity image denoising based on statistical modeling of wavelet coefficients,” *IEEE Signal Proc. Lett.* **6**, pp. 300–303, Dec. 1999.
6. J. Romberg, H. Choi, and R. Baraniuk, “Bayesian tree structured image modeling using wavelet-domain hidden Markov models,” *IEEE Trans. Image Processing.* **10**(7), pp. 1056–1068, 2001.
7. L. Şendur and I. W. Selesnick, “Bivariate shrinkage with local variance estimation,” *IEEE Signal Proc. Letters* **9**, pp. 438–441, Dec. 2002.
8. A. Pižurica, W. Philips, I. Lemahieu, and M. Acheroy, “A versatile wavelet domain noise filtration technique for medical imaging,” *IEEE Trans. Medical Imaging* **22**, pp. 323–331, Mar. 2003.
9. J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, “Image denoising using Gaussian scale mixtures in the wavelet domain,” *IEEE Trans. Image Proc.* **12**, pp. 1338–1351, Nov. 2003.
10. E. J. Candes and D. L. Donoho, “Curvelets - a surprisingly effective nonadaptive representation for objects with edges,” in *Curves and Surfaces*, e. a. L. L. Schumaker, ed., Vanderbilt University Press,Nashville, TN., 1999.
11. J. Starck, E. Candes, and D. Donoho, “The curvelet transform for image denoising,” *IEEE Trans. Image Proc.* **11**, pp. 670 –684, June 2002.
12. D. D. Y. Po and M. N. Do, “Directional multiscale modeling of images using the contourlet transform,” *IEEE Trans. Image Proc.* , 2005 (in press).
13. E. LePennec and S. Mallat, “Bandelet image approximation and compression,” *SIAM Journal of Multiscale Modeling and Simulation* , 2005 (in press).
14. M. Katona, A. Pižurica, N. Teslić, V. Kovačević, and W. Philips, “FPGA design and implementation of a wavelet-domain video denoising system,” in *Proc. Int’l Conf. Advanced Concepts for Intelligent Vision Systems ACIVS’05*, (Antwerp, Belgium), Sept. 2005.
15. A. Fletcher, V. Goyal, and K. Ramchardan, “On multivariate estimation by thresholding,” in *Proc. of the IEEE Internat. Conf. on Image Proc. ICIP*, (Barcelona, Spain), Sept. 2003.

16. P. Scheunders, "Wavelet thresholding of multivalued images," *IEEE Trans. Image Proc.* **13**, pp. 475–483, Apr. 2004.
17. P. Scheunders and D. J., "Least squares interband denoising of color and multispectral images," in *Proc. IEEE Internat. Conf. Image Proc. ICIP*, (24-27 Oct, Singapore, 2004).
18. A. Benazza-Benyahia and J.-C. Pesquet, "Building robust wavelet estimators for multicomponent images using Steins' principle," *IEEE Trans. Image Processing* , 2005 (in press).
19. A. Pižurica and W. Philips, "Estimating probability of presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Trans. Image Processing (In press)* .
20. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. and Machine Intel.* **11**(7), pp. 674–693, 1989.
21. E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. IEEE Internat. Conf. Image Proc. ICIP*, (Lausanne, Switzerland, 1996).
22. M. Hansen and B. Yu, "Wavelet thresholding via MDL for natural images," *IEEE Trans. Inform. Theory* **46**, pp. 1778–1788, Aug. 2000.
23. D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Amer. Stat. Assoc.* **90**, pp. 1200–1224, Dec. 1995.
24. H. A. Chipman, E. D. Kolaczyk, and R. E. McCulloch, "Adaptive Bayesian wavelet shrinkage," *J. of the Amer. Statist. Assoc* **92**, pp. 1413–1421, 1997.
25. F. Abramovich, T. Sapatinas, and B. Silverman, "Wavelet thresholding via a Bayesian approach," *J. of the Royal Statist. Society B* **60**, pp. 725–749, 1998.
26. M. Clyde, G. Parmigiani, and B. Vidakovic, "Multiple shrinkage and subset selection in wavelets," *Biometrika* **85**(2), pp. 391–401, 1998.
27. B. Vidakovic, "Wavelet-based nonparametric Bayes methods," in *Practical Nonparametric and Semiparametric Bayesian Statistics*, D. D. Dey, P. Müller, and D. Sinha, eds., *Lecture Notes in Statistics* **133**, pp. 133–155, Springer Verlag, New York, 1998.
28. I. Johnstone and B. Silverman, "Empirical Bayes selection of wavelet thresholds," *Annals of Statistics* **33**, 2005 (in press).
29. M. Jansen, *Noise Reduction by Wavelet Thresholding*, Springer-Verlag, New York, 2001.
30. A. Pižurica, W. Philips, I. Lemahieu, and M. Acheroy, "A joint inter- and intrascale statistical model for wavelet based Bayesian image denoising," *IEEE Trans. Image Proc* **11**, pp. 545–557, May 2002.
31. D. Middleton and R. Esposito, "Simultaneous optimum detection and estimation of signals in noise," *IEEE Trans. Inform. Theory* **14**, pp. 434–443, May 1968.
32. S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Proc.* **9**, pp. 1532–1546, Sept. 2000.
33. A. Pižurica and W. Philips, "On least squares estimators under Bernoulli-Laplacian priors," *Technical Report 09/05* , Ghent University, 2005 (available at: http://telin.ugent.be/∼sanja).
34. J. Kalifa and S. Mallat, "Minimax restoration and deconvolution," in *Bayesian Inference in Wavelet-Based Models*, P. Müller and B. Vidakovic, eds., *Lecture Notes in Statistics*, pp. 115–138,, Springer Verlag, 1999.
35. B. Corner, R. Narajanan, and S. Reichenbach, "Noise estimation in remote sensing imagery using data masking," *Internat. J. Remote Sensing* **24**(4), pp. 689–702, 2003.
36. M. Abrams and S. Cain, "Sampling, radiometry and image reconstruction for polar and geostationary meteorological remote sensing systems," in *Proc. SPIE, Image Reconstruction from Incomplete Data II*, P. Bone, M. Fiddy, and R. Milane, eds., **4792**, pp. 207–215.