A WAVELET-BASED IMAGE DENOISING TECHNIQUE USING SPATIAL PRIORS

Aleksandra PIZURICA [1], Wilfried PHILIPS [1], Ignace LEMAHIEU [1] and Marc ACHEROY [2]

[1] TELIN, Ghent University, Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium
E-mail: Aleksandra.Pizurica@telin.rug.ac.be
[2] Royal Military School, Brussels, Belgium

ABSTRACT

We propose a new wavelet-based method for image denoising that applies the Bayesian framework, using prior knowledge about the spatial clustering of the wavelet coefficients. Local spatial interactions of the wavelet coefficients are modeled by adopting a Markov Random Field model. An iterative updating technique known as iterated conditional modes (ICM) is applied to estimate the binary masks containing the positions of those wavelet coefficients that represent the useful signal in each subband. For each wavelet coefficient a shrinkage factor is determined, depending on its magnitude and on the local spatial neighbourhood in the estimated mask. We derive analytically a closed form expression for this shrinkage factor.

1. INTRODUCTION

Recent research on wavelet based image denoising has demonstrated the advantages of using Bayesian models, exploiting the prior knowledge about the statistical properties of the wavelet coefficients. The techniques of this kind usually assume independent wavelet coefficients and a heavy-tailed probability density function (pdf), such as generalized Laplacian distribution [1]. Nonlinear shrinkage of the wavelet coefficients using Bayes rules is proposed in [2]. In [3], a Bayesian estimator of the wavelet coefficients is described, which incorporates the higher order statistics to estimate the parameters of the pdf model. The list of references to related techniques and the extensive analysis of their performance is given in [4]. A wavelet based denoising method proposed in [5] is related, but different in the sense that it uses *spatial* priors. This method follows the principles of Bayesian image restoration using Markov Random Field models, given in the classical work [6]. In this paper, we propose a related approach, which shows better performance than the method of [5] and which results in a much faster computation. The differences between the two approaches can be summarized as follows: (1) In the method of [5] *random search* is used to estimate the binary mask that indicates the positions of meaningful wavelet coefficients. In our method an *iterative updating technique* is used to find the estimate of the binary mask with far fewer iterations; (2) In [5], the shrinkage factor for the wavelet coefficients equals the marginal probability that the coefficient is noise free, given the set of all observations. This probability is calculated by *counting the number of occurrences* of a given label at a given spatial position in the whole chain of masks that are generated during the random search process. To estimate the probability accurately by this counting procedure a long chain of masks is needed, even if the convergence of the underlying random search process is fast. In our method a *closed form expression* for the shrinkage factor is proposed, depending on the local observation and on the spatial neighborhood in the estimated mask.

2. THE PROPOSED METHOD

We treat the wavelet coefficients as random variables. We will denote random variables by capital letters and their realizations by the corresponding small letters. Boldface capital letters will be used for vectors of random variables and boldface small letters for vectors of realizations. We adopt a simple numbering of pixels by using a raster scanning and assigning a sequence number l to each pixel. The set of all indices l for the given array of pixels is denoted by S. In order to have a more compact representation, we will omit the indices of the wavelet coefficients that indicate the scale and the orientation. Thus, the wavelet coefficients will have only one index that corresponds to their spatial location. The notation $S \setminus l$ will denote the set of all pixels in S except the pixel l, and the notation $\partial(l)$ will denote the neighborhood of the pixel l, except l itself.

With each detail image $\mathbf{w} = \{w_1,\ldots,w_n\}$ we associate a vector $\mathbf{m} = \{m_1,\ldots,m_n\}$ of *measures*; The measure m_l can be the magnitude of the coefficient or the estimate of the local Lipschitz exponent like in [5]. Further on, we will associate with each detail image a vector $\mathbf{x} = \{x_1,\ldots,x_n\}$ of binary labels: $x_l = 0$ if w_l is assumed to originate from noise and $x_l = 1$ if w_l is assumed to represent a useful

signal. The vectors **x** will be called *binary masks*.

To model the spatial dependencies between the wavelet coefficients, we assume that the binary masks **x** are specific configurations of a Markov Random Field (MRF). We search for the estimate $\hat{\mathbf{x}}$ of the binary mask by applying the *maximum a posteriori* (MAP) principle, which maximizes the posterior probability $P(\mathbf{X}=\mathbf{x} \,|\, \mathbf{M}=\mathbf{m})$. For this maximization we use the iterative conditional mode technique (ICM) [7], which converges quickly and guarantees a close to MAP solution. Using the estimated binary mask $\hat{\mathbf{x}}$, we perform a soft modification of the wavelet coefficients as follows

$$w_l^{new} = q(m_l, \tau_{\partial(l)}) \cdot w_l = q_l \cdot w_l, \qquad (1)$$

where q_l is a shrinkage factor that depends on the local measure m_l and on a given quantity $\tau_{\partial(l)}$ derived from the spatial neighbourhood $\partial(l)$ of the pixel l in the estimated binary mask $\hat{\mathbf{x}}$. The shrinkage factor should be high (i.e., close to one) when m_l and/or $\tau_{\partial(l)}$ indicate that w_l is likely to represent a useful signal; in the opposite case q_l should be small. We choose for the shrinkage factor the following marginal probability

$$q_l = P(X_l = 1 \,|\, \mathbf{M} = \mathbf{m}, \mathbf{X}_{S\backslash l} = \hat{\mathbf{x}}_{S\backslash l}) \qquad (2)$$

because it satisfies the above requirements and because it can be derived as a closed form expression in terms of the measure m_l and the local neighborhood in the estimated mask $\hat{\mathbf{x}}_{\partial(l)}$. Assuming conditional independence $P(\mathbf{M} = \mathbf{m} \,|\, \mathbf{X} = \mathbf{x}) = \prod_l P(M_l = m_l \,|\, X_l = x_l)$ and using the fact that the label X_l can take only the values $x_l = 0$ or $x_l = 1$, we derive

$$q_l = \frac{\xi_l \cdot \eta_l}{1 + \xi_l \cdot \eta_l}, \qquad (3)$$

where ξ_l represents the ratio of conditional probabilities $P(M_l = m_l \,|\, X_l = 1)/P(M_l = m_l \,|\, X_l = 0)$ and where η_l represents the ratio of the prior probabilities $P(X_l = 1 \,|\, \mathbf{X}_{\partial(l)} = \hat{\mathbf{x}}_{\partial(l)})/P(X_l = 0 \,|\, \mathbf{X}_{\partial(l)} = \hat{\mathbf{x}}_{\partial(l)})$. For the results presented in this paper we have used an autobinomial Markov random field [8], where we assign nonzero and equal potentials only to pairwise cliques. In this case we have

$$\eta_l = \exp(\gamma \cdot \tau_{\partial(l)}), \quad \tau_{\partial(l)} = \sum_{k \in \partial(l)} (2\hat{x}_k - 1). \qquad (5)$$

For the conditional probability model $P(M_l = m_l \,|\, X_l = x_l)$ we have used the model of [5], for which we have derived

$$\xi_l = \begin{cases} \exp(-\alpha), & m_l < (1-\delta)T \\ \exp\left(\dfrac{\alpha}{\delta}\left(\dfrac{m_l}{T} - 1\right)\right), & (1-\delta)T \le m_l \le (1+\delta)T \\ \exp(\alpha), & m_l > (1+\delta)T \end{cases}, \qquad (5)$$

where T is the threshold for the measure m_l, δ determines the width of the transition interval around T, and α is the parameter that controls the influence of the noise measure on the shrinkage factor. The shrinkage factor is now determined by equations (2)–(5).

3. RESULTS AND DISCUSSION

The proposed technique can be applied to different kinds of noise; We will use a test image with artificially added white Gaussian noise (Fig. 1) to compare the performance of the proposed method and the related approach of [5]. The computation time in these two techniques is approximately proportional to the number of iterations that are performed on binary masks (the time required to compute the direct and the inverse wavelet transform is much smaller). It is assumed that one iteration is completed when all binary labels in a mask have been updated. Since both methods involve the same parameters, with equivalent meaning, we use the same set of parameters in both techniques.

In Fig.1(a), an original, noise-free image is shown. The same image with additive white Gaussian noise is shown in Fig.1(b). The result of the approach [5] is given in Fig.1(c), where the number of iterations for calculating the marginal conditional probabilities was $N=10$. The peak signal to noise ratio for this image is PSNR=26.6dB. The result of the proposed approach, with N=5 iterations is shown in Fig. 1(d), and the corresponding PSNR=29.5dB. In our approach $N=5$ iterations suffices to reach the convergence, while in the method of [5] the convergence is reached only after much larger N, but $N=10$ provides a good result. In both cases the magnitude of the wavelet coefficient was the local measure $m_l = |w_l|$, and the parameters were $\alpha=1$, $\beta=1.1$ and $\delta=0.5$.

It can be noted that the result of our method is better as far as the quantitative measure is concerned; in visual appearance it is slightly sharper but also contains more ringing artifacts. We can therefore conclude that the two methods provide denoised images of similar quality, but our method is almost two times faster.

We have also tested the performance of the proposed denoising method on naturally noisy images. An original Synthetic Aperture Radar (SAR) image is shown in Fig. 2.(a) and the result of the proposed technique in Fig. 2(b). In Fig. 3, we show the result of applying the proposed

(a)

(b)

(c)

(d)

Fig 1. (a) The original, noise-free image. (b) The image with additive gaussian noise (zero mean, standard deviation 20). (c) The result of the method of [5], with $N=10$ iterations in each wavelet subband. (d) The result of the proposed method with $N=5$ iterations in each wavelet subband.

technique to an infrared image of a buried landmine. These two examples show that the proposed method can be applied to natural images, where it significantly reduces noise without deteriorating significant edges.

4. CONCLUSION

This paper presents an improvement of the algorithm in [5], which makes it almost two times faster. The proposed method performs well in presence of different kinds of noise. Further research will include more realistic models for the conditional probabilities $P(M_l = m_l \mid X_l = 0)$ and $P(M_l = m_l \mid X_l = 1)$; There models are derived by analyzing the histograms of the local measure m_l in the areas of meaningful edges and in background parts of the image, in the presence of different kinds of noise. We also intend to better exploit the spatial correlation of the wavelet coefficients at different resolution scales, by extending the neighborhood in the prior spatial model across the scales, with appropriately assigned potentials to different intra-scale and inter-scale cliques.

5. REFERENCES

[1] S. Mallat, "A theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. on Pattern Anal. and Machine Intel.*, 11 (7), pp. 674-693, July 1989.
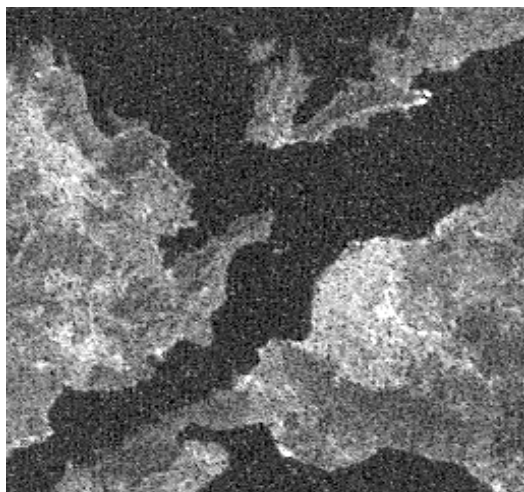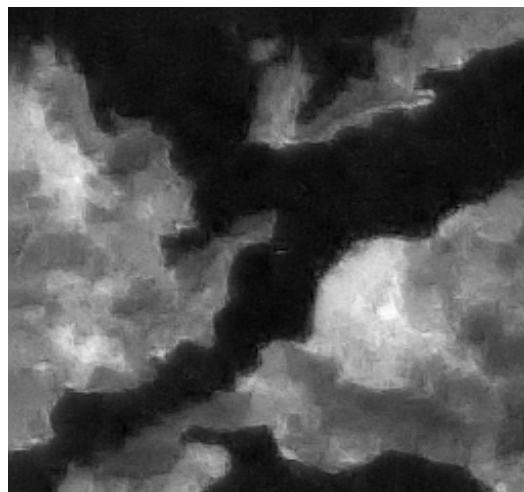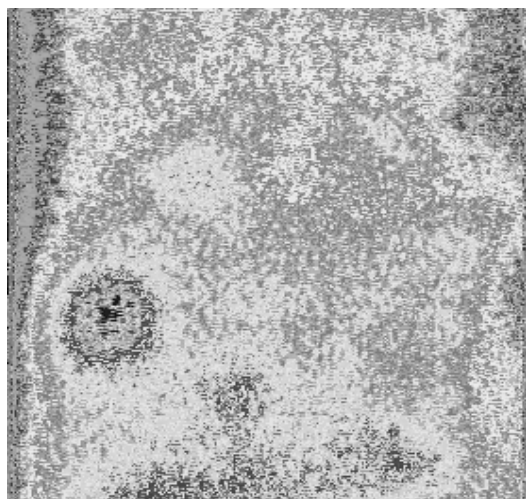
Fig. 2. (a) An original SAR image. (b) The result of the proposed denoising technique.

Fig. 3. (a) An original infrared image of a buried landmine. (b) The result of the proposed denoising technique.

[2] B. Vidakovic, "Nonlinear Wavelet Shrinkage with Bayes Rules and Bayes Factors," Duke Univ., Durham, NC, 1994, preprint, available on http://www.isds.duke.edu

[3] E. P. Simoncelli, E. H. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. IEEE Int. Conf. Image Proc. (ICIP)*, Lausanne, Switzerland, pp. 379-382, Oct. 1996,

[4] P.Moulin, J. Liu, "Analysis of Multiresolution Image Denoising Schemes Using Generalized Gaussian and Complexity Priors," *IEEE, Trans. on Inform. Theory,* 45 (3), pp. 909-919, Apr. 1999.

[5] M. Malfait, D. Roose, "Wavelet-Based Image Denoising Using a Markov Random Field *a Priori* Model", *IEEE, Trans. on Image Proc.*, 6, (4), pp. 549-565, April 1997.

[6] S. Geman, D. Geman, "Stochastic Relaxation, Gibbs Distribu-tions, and the Bayesian Restoration of Images," *IEEE, Trans. on Pattern Anal. and Machine Intel,* PAMI-6 (6), pp. 721-741, Nov. 1984.

[7] S. Z. Li, *Markov Random Field Modeling in Computer Vision*, Computer Science Workbench, Springer, 1995.

[8] G. R. Cross, A. K. Jain, "Markov Random Field Texture Models," *IEEE, Trans. on Pattern Anal. and Machine Intel,* PAMI-5 (1), pp. 25-39, Jan. 1983.

[9] D. L. Donoho, "De-noising by Soft Thresholding," *IEEE Trans on Inform. Theory*, 41 (3), pp. 613-627, May 1995.