Proc. IEE Conf. On Image Processing and its Applications, IPA99, pp. 216-219, July 1999.

IMAGE DE-NOISING IN THE WAVELET DOMAIN USING PRIOR SPATIAL CONSTRAINTS

A Pizurica[1,2], W Philips[1], I Lemahieu[1] and M Acheroy[2]

[1]University of Ghent, Belgium [2] Royal Military School, Belgium

***Abstract.*- This paper describes an algorithm for the noise removal in the wavelet domain, which takes into account not only the local noise measure, but also prior spatial constraints. These prior spatial or geometrical constraints express the fact that meaningful wavelet coefficients appear in spatially connected clusters, forming edges of given directions. Existing techniques that exploit this kind of prior knowledge are computationally expensive, while the proposed method exploits the spatial constraints in a different and simple manner.**

I. INTRODUCTION

A number of different wavelet based de-noising techniques for the suppression of noise were proposed in recent years. A common approach is to compare the magnitude of the wavelet coefficient with a certain threshold. The coefficients that are below the threshold are usually replaced by zero. The other coefficients are kept unchanged or reduced in absolute value according to some predefined rule. In the wavelet shrinkage technique of Donoho (1) the threshold depends on the estimated noise variance and the number of discrete data points. The method of Xu et al (2) exploits the fact that useful signal features are strongly correlated across scales. On the other hand, for the wavelet coefficients of noise this correlation between scales is very weak. In the approach of Mallat and Hwang (3) the estimate of the local regularity, expressed through Lipschitz exponents, is used to distinguish noise from useful signal transitions. The local Lipschitz exponents are estimated from the decay of the wavelet coefficients through scales. Practically, the method examines only the behavior of the local extrema. Reconstruction of the denoised image from the modified wavelet extrema is performed through an iterative procedure, which makes this method computationally expensive.

The method of Malfait and Rose (4) takes into account not only the evolution of the wavelet coefficients through scales, but also *prior* geometrical constraints. These geometrical constraints are combined with the local regularity measure in a Bayesian probability formulation. For each wavelet coefficient the probability of being noise-free is calculated. The coefficients are then modified according to these probabilities. Computation of the corresponding probabilities in (4) involves a stochastic sampling procedure, which is computationally demanding. This makes the approach of (4) less attractive for those applications that require fast processing.

In this paper, we propose a computationally efficient denoising algorithm in the wavelet domain, which exploits prior spatial constraints in a simpler manner. It is supposed that every wavelet coefficient is affected by noise and a unique modification rule is applied to all the coefficients. The magnitude of each wavelet coefficient is “shrunk” by a certain amount, by multiplying the coefficient with the corresponding *shrinkage factor*, which takes the values between 0 and 1. We propose an analytical expression for the shrinkage factor, which depends on both noise measure and a geometrical measure that are derived from a given neighborhood. We derive the de-noising model in a general form, containing different parameters that can be adjusted according to a specific application. The noise measure can be a local regularity measure or the magnitude of the wavelet coefficients.

The proposed method can be applied to a wide class of natural and synthesized images. We are, however, mostly interested in de-noising infrared images of natural terrain, for the purpose of landmine detection. The regularity properties of the images of natural terrain often vary from point to point and it is important to know the distribution of the singularities in order to analyze their properties. Pointwise measurements of Lipschitz exponents are not possible in this case because of the finite spatial resolution. Therefore, in practical implementations we use the magnitudes of the wavelet coefficients as a criterion for selecting noisy coefficients.

II. THE PROPOSED METHOD

The proposed method combines a criterion, which determines how noisy the wavelet coefficients are with some prior geometrical assumptions. These prior assumptions express the fact that meaningful wavelet coefficients appear in spatially connected clusters at the location of characteristic image features like edges, corners, etc.

The same modification rule is applied at all the levels and orientations in the wavelet decomposition. Therefore, in the following discussion, the coefficients will carry the spatial position as the only index. Hence, w_l is the wavelet coefficient at the spatial position l in the corresponding detail image. The appropriate noise measure m_l is associated with each wavelet coefficient. For this purpose we will use the magnitude of the wavelet coefficient $m_l = |w_l|$. We also associate with each wavelet coefficient a binary label. The binary labels are denoted by x_l, $x_l \in \{0,1\}$ where, $x_l = 0$ if w_l originates from noise and $x_l = 1$ if w_l originates from signal (i.e., belongs to an edge). Finally, spatial constraints will be taken into account through a parameter t_l, which depends on the given neighborhood N_l. Generally, t_l will represent a linear combination of the labels x_k, where $x_k \in N_l$.

We want to derive a modifying rule for the wavelet coefficients, which takes into account both the noise measure (e.g. the magnitude of the coefficient) and the prior spatial constraints. A simple way to accomplish this is to multiply each wavelet coefficient with the corresponding shrinkage factor q_l, $0 \le q_l \le 1$, which depends on both m_l and the neighborhood parameter t_l.

$$w_l^{new} = w_l \cdot q_l = w_l \cdot q(m_l, t_l). \qquad (1)$$

The parameter q_l should be high, i.e., close to 1, when the magnitude of the coefficient m_l is high and/or when t_l is large, i.e., the majority of the coefficients in the neighborhood N_l have high magnitudes. The fulfillment of the latter condition implies that the neighborhood N_l centered at the spatial position l contains an edge segment. More specific geometrical constraints (e.g. edges of given directions in different detail images) can be obtained by choosing the appropriate neighborhood configurations and the appropriate weighting factors for the labels inside the neighborhood.

Let us denote by T an optimal threshold for the noise measure. In classical hard thresholding techniques all the wavelet coefficients for which $m_l \le T$ are set to zero, which corresponds to putting $q_l = 0$ in equation (1), and all the coefficients for which $m_l > T$ are kept unchanged, which corresponds to putting $q_l = 1$. Instead of that, we define a range around the threshold $((1\text{-}\alpha)T, (1+\alpha)T)$, $0 \le \alpha \le 1$, where we suppose that the noise measure m_l itself is not sufficient to classify the coefficients as being clean or noisy. Hence, for $(1\text{-}\alpha)T < m_l < (1+\alpha)T$ the shrinkage factor q_l is in the range $0 < q_l < 1$, and it is monotonically increasing with m_l, but it is also influenced by prior geometrical constraints expressed through t_l. The idea behind this is the following: the magnitudes of those coefficients that are more likely to belong to edges should be less reduced. The coefficient is more likely to belong to an edge if its magnitude m_l is higher and if the most of the neighboring coefficients are supposed to belong to an edge, which is specified by the neighborhood parameter t_l. Only in "extreme" cases when $m_l \le (1\text{-}\alpha)T$, or when $m_l \ge (1+\alpha)T$ geometrical constraints do not influence our decision about the coefficient. Thus, for $m_l \le (1\text{-}\alpha)T$ we have q_l=0 and for $m_l \ge (1+\alpha)T$ we have q_l=1.

With reference to the discussion above, we propose the following form for the shrinkage factor

$$q_l = \begin{cases} 0, & m_l \le (1-\alpha)T \\ \dfrac{\xi_l^{\beta}\eta_l^{\gamma}}{1+\xi_l^{\beta}\eta_l^{\gamma}}, & (1-\alpha)T < m_l < (1+\alpha)T \\ 1, & m_l \ge (1+\alpha)T \end{cases}, \qquad (2)$$

where $\xi_l = \xi(m_l)$ depends on the noise measure in the following way

$$\xi_l = \frac{m_l - (1-\alpha)T}{(1+\alpha)T - m_l}, \qquad (1-\alpha)T < m_l < (1+\alpha)T \qquad (3)$$

and $\eta_l = \eta(t_l) > 0$ expresses the influence of the prior spatial constraints. The exact dependence of the parameter η_l on t_l will be specified later. The proposed forms for q_l and ξ_l ensure that $0 \le q_l \le 1$ and that q_l is monotonically increasing with m_l in the range $(1\text{-}\alpha)T < m_l < (1+\alpha)T$, where $q_l \to 0$ for $m_l \to (1\text{-}\alpha)T$ and $q_l \to 1$ for $m_l \to (1+\alpha)T$, for any given value of η_l. Moreover, when there is no influence of the spatial constraints ($\eta_l^{\gamma} = 1$) the chosen form of $\xi(m_l)$ in equation (3) ensures a linear change of q_l in the range $(1\text{-}\alpha)T < m_l < (1+\alpha)T$, for β=1. The parameters β and γ are introduced in order to allow us to control the influence of the noise measure m_l and the spatial constraint t_l.

In order to completely specify the shrinkage factor q_l it still remains to define the function $\eta(t_l)$. To ensure the desired influence of the spatial constraints on q_l the parameter η_l should be greater than one if it is likely that the coefficient w_l belongs to an edge ($x_l = 1$), given t_l. If it is more probable that w_l does not belong to an edge ($x_l = 0$) η_l should be smaller than one. When both cases are equally probable, η_l should not influence the shrinkage factor q_l, and thus η_l should be equal to one in this case. A reasonable choice for η_l is therefore

$$\eta_l = \frac{P(x_l = 1 \mid t_l)}{P(x_l = 0 \mid t_l)}. \qquad (4)$$

To estimate the conditional probailities in equation (4) we need a model for the interaction between neighboring labels. A set of binary labels $\{x_l\}$ at a given level and orientation in the wavelet decomposition will be called *mask*. The masks are obtained by comparing the magnitude of the wavelet coefficients with a certain threshold. We assume that the probability of the binary mask can be modeled as a Markov random field. In this case, according to Cross and Jain (5), probability of the binary label x_l conditioned on the neighboring labels can be expressed as

$$P(x_l \mid t_l) = \frac{\exp(x_l t_l)}{1+\exp(t_l)}, \qquad (5)$$

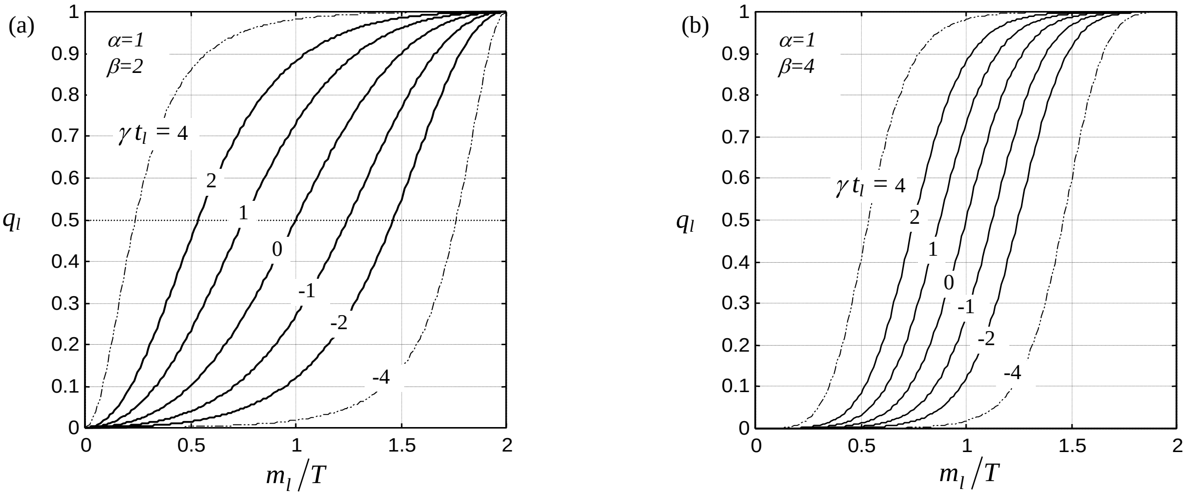

Figure 1: The shrinkage factor q_l versus normalized magnitude of the coefficient m_l /T. The parameter is the product γt_l (a) $\alpha = 1$, $\beta = 2$ and (b) $\alpha = 1$, $\beta = 4$.

where t_l represents a linear combination of the labels in the neighborhood of x_l. It then follows that the parameter η_l is equal to

$$\eta_l = \exp(t_l)\,. \qquad (6)$$

In this paper, we choose different neighborhoods for the subbands of different orientations and we calculate the neighborhood parameter at the position l simply as

$$t_l = \sum_{k \in N_l} (2x_k - 1)\,. \qquad (7)$$

This form ensures that $t_l > 0$ when most of the labels x_k in the neighborhood N_l are equal to one. In other words, a positive t_l means that most of the coefficients in the neighborhood of l are assumed to belong to an edge. It is then likely that the coefficient at the position l also belongs to an edge. According to equation (6), $\eta_l > 1$ in this case, as it was required. The value $t_l = 0$ in equation (7) indicates that equal number of the wavelet coefficients in N_l have labels 0 and 1. In this case, it is equally likely that the coefficient w_l belongs to an edge or not and the corresponding η_l should be equal to one as it is provided by equation (6).

The shrinkage factor $q_l = q(m_l,t_l)$ is determined by equations (2), (3) and (6). Dependence of q_l on both the magnitude of the wavelet coefficient m_l and the neighborhood parameter t_l is presented in Figure 1. If the magnitude of the coefficient is equal to the chosen threshold ($m_l/T = 1$) and the neighborhood N_l contains the same number of edge and non-edge labels ($t_l = 0$), then the coefficient is shrunk to one half of its value. It can be seen that depending on the neighborhood a coefficient can be significantly reduced even if its magnitude is highly above the threshold. Also, a coefficient might be only slightly reduced even when it is well bellow the threshold if the neighborhood parameter is high. A steepest change of q_l with m_l is obtained when parameter β is higher, as can be seen by comparing Figure 1(a) and Figure 1(b).

III. RESULTS

The images shown in Figure 2(a) and Figure 2(b) are original infrared images of buried land mines. Results obtained with the proposed de-noising technique applied to these images are shown in Figure 2(c) and Figure 2(d), respectively. We used the redundant wavelet transform, with two detail images at each scale as in the approach of Mallat and Hwang (3). For each of these detail images a universal threshold of Donoho (1) was calculated. The neighborhood parameter was calculated from 5x3 and 3x5 neighborhoods for horizontal and vertical subbands, respectively.

Obviously, the proposed method facilitates the interpretation of the analyzed infrared images and as such could be an important help to a human observer. The proposed method could also be useful as a pre-processing step to automated target detection applications, where it could improve the detection performance because it removes noise without deteriorating the sharpness of the image. However, additional efforts are needed in order to confirm the usefulness of the proposed method in the latter application

IV CONCLUSION

We proposed a de-noising method in the wavelet domain, which incorporates prior spatial assumptions. These prior assumptions are general and express the fact that meaningful wavelet coefficients are spatially clustered, forming edges of given directions in different detail images. We combine this prior knowledge with a criterion that describes how noisy the coefficients are. Each wavelet coefficient is multiplied with a corresponding shrinkage factor, which depends both on the noise measure and on the prior spatial assumptions. An

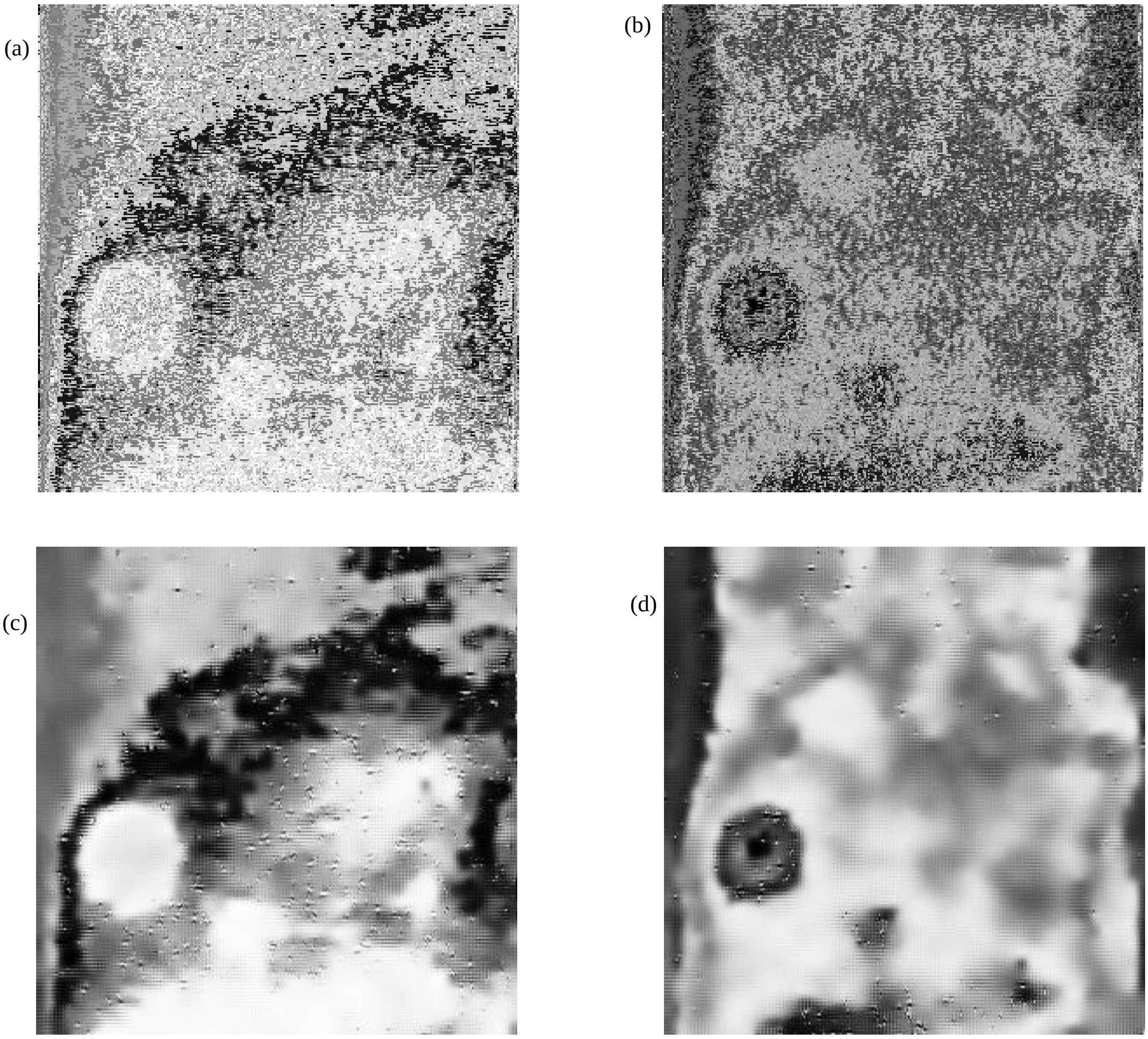

Figure 2: (a) and (b) original infrared images of buried land mines.
(c) and (d) results of the de-noising applied to the images (a) and (b), respectively.

analytical expression for the shrinkage factor is proposed. The presented results show that the proposed method can significantly reduce the noise while preserving the significant image features. Other important advantages of this method are speed and simplicity. It is computationally much more efficient then some other algorithms that use prior spatial constraints, like the method of (4) where a stochastic sampling procedure is needed to compute shrinkage factors for the wavelet coefficients.

REFERENCES

1. Donoho D L, 1995, “De-noising by Soft Thresholding,” Trans on Inform. Theory, 41. 613-627.
2. Xu Y, Weaver J B, Healy D M, Lu J, 1994, “Wavelet Transform Domain Filters: A Spatially Selective Noise Filtration Technique,” Trans. on Image Proc., 3, 747-758.
3. Mallat S, Hwang W L, 1992, “Singularity Detection and Processing with Wavelets,” Trans on Inform. Theory, 38, 617-643.
4. Malfait M, Roose D, 1997, “Wavelet-Based Image Denoising Using a Markov Random Field *a Priori* Model”, Trans. on Image Proc., 6, 549-565.
5. Cross G R, Jain A K, 1983, “Markov Random Field Texture Models,” Trans. Pattern Anal. Machine Intell., PAMI-5, 25-39.