ELSEVIER

Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT®

Journal of Computational and Applied Mathematics 164–165 (2004) 39–52

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Filtered wavelet thresholding methods

Silvia Bacchelli*, Serena Papi

Department of Mathematics, University of Bologna, Via Sacchi, Cesena I-47023, Italy

Received 20 August 2002; accepted 2 September 2003

Abstract

When working with nonlinear filtering algorithms for image denoising problems, there are two crucial aspects, namely, the choice of the thresholding parameter λ and the use of a proper filter function. Both greatly influence the quality of the resulting denoised image. In this paper we propose two new filters, which are a piecewise quadratic and an exponential function of λ, respectively, arid we show how they can be successfully used instead of the classical Donoho and Johnstone's Soft thresholding filter. We exploit the increased regularity and flexibility of the new filters to improve the quality of the final results. Moreover, we prove that our filtered approximation is a *near-minimizer* of the functional which has to be minimized to solve the denoising problem. We also show that the quadratic filter, due to its shape, yields good results if we choose λ as the Donoho and Johnstone universal threshold, while the exponential one is more suitable if we use the recently proposed H-curve criterion. Encouraging results in extensive numerical experiments on several test images confirm the effectiveness of our proposal.

MSC: 65D; 65Y20; 65F

Keywords: Image denoising; Variational approach; Filtered solution; Wavelet thresholding

1. Introduction

In recent years nonlinear filtering techniques have played an important role for the numerical solution of the image denoising problem. These techniques arise as exact or approximate solutions of the following variational problem: given a positive parameter λ and a noisy image $\bar{f}(x, y)$, defined for (x, y) in some square domain I, find a function g_λ^* that minimizes, over all possible functions g

* Corresponding author.
E-mail address: silvia@csr.unibo.it (S. Bacchelli).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2003.09.005

in a given space Y, the functional

$$\|\bar{f} - g\|_{L_2(I)}^2 + 2\lambda\|g\|_Y, \tag{1}$$

where

$$\|\bar{f} - g\|_{L_2(I)} := \left(\int_I |\bar{f}(x,y) - g(x,y)|^2 \,\mathrm{d}x\,\mathrm{d}y\right)^{1/2} \tag{2}$$

is the $L_2(I)$-norm which measures the root-mean square error between $\bar{f}$ and g and $\|g\|_Y$ is the norm of the approximation g in a smoothness space Y. An important role in the formulation of the problem is played by the positive parameter λ, which determines the relative importance of the smoothness of g and the quality of the approximation to the given data $\bar{f}$.

Since the efficiency of solving problem (1) strongly depends on the computing rate of the two norms $\|\bar{f} - g\|_{L_2(I)}^2$ and $\|g\|_Y$, great advantages both in terms of computational effort and quality of the results can be obtained by working in function spaces Y where the norms of g can be expressed in terms of its wavelet coefficients. Even if there are many possible choices of the space Y, the family of Besov spaces $B_q^\alpha(L_p(I))$ has been revealed to be a good framework when working in this variational context. In fact, these spaces contain functions of different smoothness as are real images; in particular, by varying the values of the parameters α, p and q, it is possible to get the smoothness space where the best representation of any single image can be achieved (for more details the reader is referred to [2]).

In [6] it has been shown that, taking Y as the Besov space $B_1^1(L_1(I))$, the exact minimizer of (1) is obtained by means of the so-called *wavelet shrinkage*, which was previously introduced by Donoho and Johnstone [8]. Wavelet shrinkage discards the wavelet coefficients of $\bar{f}$ whose absolute value is less than the shrinkage parameter λ and shrinks the others by the value of λ towards zero. Hence, it can be seen as a wavelet approximation of $\bar{f}$, which has been filtered with a filter, which turns out to be a piecewise linear function when considered as function of λ. In this work we introduce two new classes of wavelet approximation of $\bar{f}$, which are also filtered versions of the original perturbed image $\bar{f}$, but whose associated filters present an increased regularity as functions of λ, since they are either a piecewise 2-degree polynomial or an exponential function of the parameter λ. The resulting wavelet approximations therefore yield a smoother transition between maintained and discarded wavelet coefficients, and this aspect seems to be particularly useful when dealing with real images. Our main result consists in proving that both the above approximations are *near-minimizers* of functional (1), giving at the same time an upper bound to the constant involved in the definition of near-minimizer.

Moreover, since when working with thresholding algorithms the most important problem is to evaluate an appropriate value for the shrinkage parameter λ, we consider two possible choices of the parameter λ, namely, the universal threshold of Donoho and Johnstone [7,8] and the recently proposed H-curve criterion [10]. We present extensive computations showing how each one of our two new filters is particularly suited for one of the two thresholding choices. More precisely, the parabolic shape of one filter can balance the oversmoothing of the image introduced by the use of the Donoho and Johnstone threshold, while the more regular exponential shape of the other performs better when dealing with the H-curve criterion.

The paper is organized as follows: in Section 2 we briefly recall the basic notions of wavelet theory in the 1-D and 2-D case. In Section 3 we recall the exact solution of the variational problem in the wavelet domain and we introduce our two new filtered approximations. In Section 4 we show that these approximations are near-minimizers of problem (1). The choice of the threshold parameter is discussed in Section 5 together with the analysis of the numerical experiments on several test images, which show the effectiveness of our proposal.

2. Basics on wavelet theory

An orthonormal *multiresolution analysis* of $L_2(\mathbb{R})$ is built using two basis functions: a *scaling function* ϕ and a *wavelet* ψ. This involves a nested sequence of subspaces of $L_2(\mathbb{R})$, $\{V_j\}_{j\in\mathbb{Z}}$, satisfying $\ldots V_{j-1} \subset V_j \subset V_{j+1} \ldots$; the *scaling function* ϕ is the solution of the two-scale equation

$$\phi(x) := \sqrt{2}\sum_k h_k \phi(2x-k). \tag{3}$$

The integer translates of the scaling function $\{\phi(x-k)|k\in\mathbb{Z}\}$, form an orthogonal basis for the closure of their span V_0.

For each $j\in\mathbb{Z}$, the integer translates of the jth diadic dilates of $\phi, \{\phi_{j,k}(x) := 2^{j/2}\phi(2^j x - k),\ k\in\mathbb{Z}\}$ form an orthonormal basis for V_j. It turns out that the wavelet space $W_j = \overline{\text{span}\{\psi_{j,k},\ k\in\mathbb{Z}\}}$ is the orthogonal complement of V_j in V_{j+1}, namely, $V_j \oplus W_j = V_{j+i}$, that $L_2(\mathbb{R}) = \oplus_{j\in\mathbb{Z}} W_j$, and the wavelet ψ satisfies the two-scale equation

$$\psi(x) := \sqrt{2}\sum_k w_k \phi(2x-k). \tag{4}$$

The functions $\{\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)\}$, with $j,k\in\mathbb{Z}$ form an orthogonal basis of $L_2(\mathbb{R})$. Given a function $f\in L_2(\mathbb{R})$, the wavelet expansion of f is therefore given by

$$f = \sum_{j,k\in\mathbb{Z}} c_{j,k}\psi_{j,k}, \tag{5}$$

where $c_{j,k} := \int_{\mathbb{R}} f(x)\psi_{j,k}(x)\,\mathrm{d}x$ and the L_2-norm of f is given in terms of its wavelet coefficients, that is

$$\| f \|^2_{L_2(\mathbb{R})} = \sum_{j,k\in\mathbb{Z}} c^2_{j,k}. \tag{6}$$

The wavelet coefficients $c_{j,k}$ of f can be evaluated by means of a fast algorithm involving the sequences of the two-scale coefficients $\{h_k\}$ and $\{w_k\}$ which are called low-pass and high-pass filters, respectively, and, in the compactly supported case, are finite [5].

The one-dimensional multiresolution analysis can easily be generalized to a 2-D one for $x := (x_1,x_2)\in\mathbb{R}^2$, by defining a 2-D scaling function and the corresponding 2-D wavelets as follows:

$$\begin{aligned} &\phi(x_1,x_2) := \phi(x_1)\cdot\phi(x_2), \qquad \psi^{(1)}(x_1,x_2) := \psi(x_1)\cdot\phi(x_2),\\ &\psi^{(2)}(x_1,x_2) := \phi(x_1)\cdot\psi(x_2), \quad \psi^{(3)}(x_1,x_2) := \psi(x_1)\cdot\psi(x_2). \end{aligned} \tag{7}$$

By setting $\Psi = \{\psi^{(1)}, \psi^{(2)}, \psi^{(3)}\}$, we have that the set of functions

$$\{\psi_{j,k}(x) = 2^j \psi(2^j x - k)\}_{\psi\in\Psi,\ j\in\mathbb{Z},\ k\in\mathbb{Z}^2} \tag{8}$$

forms an orthonormal basis for $L_2(\mathbb{R}^2)$. Hence, for every $f \in L_2(\mathbb{R}^2)$ we have

$$f = \sum_{k\in\mathbb{Z}^2,\ j\in\mathbb{Z},\ \psi\in\Psi} c_{j,k,\psi}\psi_{j,k}\ , \tag{9}$$

with

$$c_{j,k,\psi} := \int_{\mathbb{R}^2} f(x)\psi_{j,k}(x)\,\mathrm{d}x. \tag{10}$$

Moreover, also for the two-dimensional case we have

$$\| f \|^2_{L^2(\mathbb{R}^2)} = \sum_{k\in\mathbb{Z}^2,\ j\in\mathbb{Z},\ \psi\in\Psi} c^2_{j,k,\psi}. \tag{11}$$

Identity (11) expresses the L_2-norm of f in terms of its wavelet coefficients. This can easily be extended to more general spaces, such as the Besov spaces $B^\alpha_q(L_p(I))$, where the norm $\| f \|_{B^\alpha_q(L_p(I))}$ is *equivalent* to a norm of the sequence of wavelet coefficients (for more details, see [2])

$$\| f \|_{B^\alpha_q(L_p(I))} \asymp \left(\sum_j \left(\sum_{k,\psi} 2^{\alpha j p} 2^{j(p-2)} |c_{j,k,\psi}|^p \right)^{q/p} \right)^{1/q}. \tag{12}$$

3. Filtered approximations in the wavelet domain

As has been pointed out in [2], a good choice for the function space Y is the Besov space $B^1_1(L_1(I))$. Problem (1), then, becomes

$$\| \bar{f} - g\|^2_{L_2(I)} + 2\lambda \|g\|_{B^1_1(L_1(I))}. \tag{13}$$

We assume that the observed image $\bar{f}$ in (13) takes the form

$$\bar{f} = f + \varepsilon,$$

where f is the original deterministic signal and ε the noise component, given by the i.i.d. Gaussian Noise with mean 0 and variance σ^2 (denoted as $\mathcal{N}(0, \sigma^2)$).

By using relation (9), we can expand $\bar{f}$ and g in terms of their wavelet coefficients as follows:

$$\bar{f} = \sum_{j,k,\psi} \bar{c}_{j,k,\psi}\psi_{j,k} \quad \text{and} \quad g = \sum_{j,k,\psi} d_{j,k,\psi}\psi_{j,k}, \tag{14}$$

where, to ignore all further complications, we shall not precisely specify the domains of the indices of the sums, meaning that the suitable modifications have been carried out [2]. In fact, when dealing with a finite domain I, as is the case for images, some changes must be made to the wavelet basis in order to obtain an orthonormal basis for $L_2(I)$. We refer to [3] for further details.

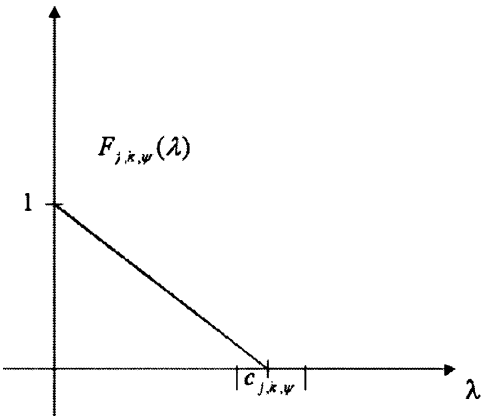

Fig. 1. Soft thresholding filter factor corresponding to $\bar{c}_{j,k,\psi}$ as function of λ.

By replacing $\bar{f}$ and g in (13) with their wavelet expansions (14), and by applying relation (12) with $p = q = 1$ and $\alpha = 1$, we get

$$\sum_{j,k,\psi} |\bar{c}_{j,k,\psi} - d_{j,k,\psi}|^2 + 2\lambda \sum_{j,k,\psi} |d_{j,k,\psi}|, \tag{15}$$

that is, the formulation of the variational problem in the wavelet domain.

This problem is separable; hence the minimum of this functional is obtained by minimizing each term separately over $d_{j,k,\psi}$, namely

$$|\bar{c}_{j,k,\psi} - d_{j,k,\psi}|^2 + 2\lambda |d_{j,k,\psi}|. \tag{16}$$

In [2] the authors have shown that the exact minimizer of problem (15) is given by

$$g_\lambda^* = \sum_{j,k,\psi} d^*_{j,k,\psi,\lambda}\ \psi_{j,k} = \sum_{j,k,\psi} F_{j,k,\psi}(\lambda) \bar{c}_{j,k,\psi} \psi_{j,k} \tag{17}$$

with

$$F_{j,k,\psi}(\lambda) = \frac{(|\bar{c}_{j,k,\psi}| - \lambda)_+}{|\bar{c}_{j,k,\psi}|} = \begin{cases} 1 - \dfrac{\lambda}{|\bar{c}_{j,k,\psi}|} & \text{if } |\bar{c}_{j,k,\psi}| > \lambda, \\ 0 & \text{if } |\bar{c}_{j,k,\psi}| \leqslant \lambda. \end{cases} \tag{18}$$

The solution g_λ^* is none other than the one obtained by wavelet shrinkage, where the wavelet coefficients $\bar{c}_{j,k,\psi}$ whose absolute value is greater than the threshold parameter are shrunk towards zero by an amount of λ to obtain the wavelet coefficients $d^*_{j,k,\psi,\lambda}$ of the exact minimizer g_λ^* (see Fig. 1). In [10] it has been pointed out that the solution g_λ^* given by (17) can be seen as a filtered version of the original corrupted image $\bar{f}$, where the filter factors $F_{j,k,\psi}(\lambda)$, $\psi \in \Psi$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^2$ are obtained by evaluating for $c = \bar{c}_{j,k,\psi}$ the piecewise linear filter function

$$F(c,\lambda) = \frac{(|c| - \lambda)_+}{|c|} = \begin{cases} 1 - \dfrac{\lambda}{|c|} & \text{if } |c| > \lambda, \\ 0 & \text{if } |c| \leqslant \lambda. \end{cases}$$

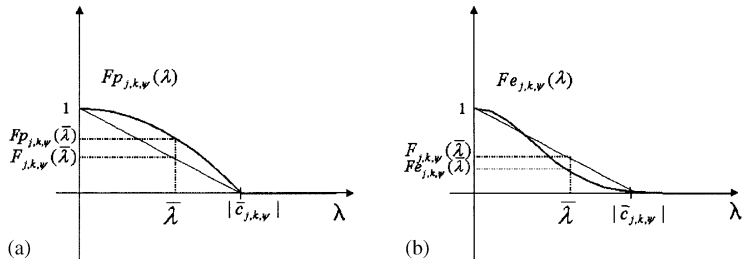

Fig. 2. New kinds of filter factors compared with the Soft thresholding one. (a) quadratic filter factor; (b) an exponential filter factor with $a = 2$ and $h = 4$.

Since real images present many different characteristics and hence a fixed choice of the filter function and the threshold level λ can produce a different visual effect on their denoised versions, it is interesting to study which filter function better fits the image features and the choice of the thresholding parameter.

In this work we investigate new kinds of filter functions whose corresponding wavelet approximations $\hat{g}_\lambda$ turn out to be more flexible than the soft thresholding one. These new filters have been chosen taking into account both their capability of adapting to the different types of image and the method used for the choice of the threshold level.

The two new filter functions taken into consideration are the *quadratic* function

$$Fp(c,\lambda) = \begin{cases} -\dfrac{\lambda^2}{c^2} + 1 & \text{if } |c| > \lambda, \\ 0 & \text{if } |c| \leqslant \lambda, \end{cases} \tag{19}$$

which is a piecewise 2-degree polynomial, and the family of *exponential* functions

$$Fe(c,\lambda) = a^{-h\lambda^2/c^2}, \quad a > 1 \text{ and } h > 0, \tag{20}$$

which have C^∞-regularity and whose shape is modelled by acting on parameters a and h.

As previously stated, we obtain the filter factors $Fp_{j,k,\psi(\lambda)}$ and $Fe_{j,k,\psi}(\lambda)$, whose action is to weigh each starting wavelet coefficient $\bar{c}_{j,k,\psi}$, $(\psi \in \Psi,\ j \in \mathbb{Z},\ k \in \mathbb{Z}^2)$ appropriately, by evaluating our new filter functions for $c = \bar{c}_{j,k,\psi}$.

The quadratic filter, by definition, discards the same wavelet coefficients as the soft thresholding, but gives more weight to those that are maintained, shrinking them by a smaller amount towards zero. On the other hand, the exponential filter, while not completely removing any wavelet coefficient, still keeps track of all the high frequencies, although it gives them a rapidly decreasing weight. The plot of both filters, compared with the Soft thresholding one, are shown in Fig. 2. It is easy to note that for the same value $\lambda = \bar{\lambda}$ the corresponding weight is quite different for the three filter functions.

4. Near minimizers of the variational problem

As already mentioned, the well-known Soft thresholding method, when applied to the wavelet coefficients of the perturbed image, yields an approximation of the true image g^*_λ which is the exact minimizer of the variational problem (13). Nevertheless, the authors of [2] pointed out that a near

minimizer can always be found. In fact they proved that the wavelet approximation obtained with the Hard thresholding method is a near minimizer of (15).

In order to investigate the behaviour of the approximations associated with the above introduced new filters, we give a characterization of a near minimizer.

More precisely, we call a family of functions $\hat{g}_\lambda$ a *near minimizer* for (13) if, for a given $\lambda > 0$, a positive constant C not depending on λ or $\bar{f}$ exists such that

$$\|\bar{f} - \hat{g}_\lambda\|^2_{L_2(I)} + 2\lambda\|\hat{g}_\lambda\|_{B^1_1(L_1(I))} \leqslant C \min_{g \in B^1_1(L_1(I))} \|\bar{f} - g\|^2_{L_2(I)} + 2\lambda\|g\|_{B^1_1(L_1(I))} \tag{21}$$

that is,

$$\|\bar{f} - \hat{g}_\lambda\|^2_{L_2(I)} + 2\lambda\|\hat{g}_\lambda\|_{B^1_1(L_1(I))} \leqslant C\|\bar{f} - g^*_\lambda\|^2_{L_2(I)} + 2\lambda\|g^*_\lambda\|_{B^1_1(L_1(I))},$$

where g^*_λ is given by (17).

We are now in a position to prove our main result, namely, to prove that the wavelet approximations of $\bar{f}$ obtained by means of filters (19) and (20) are near minimizers for problem (13) with a small constant C. More precisely, we have

Theorem 4.1. *Let $Fe_{j,k,\psi}$ be the exponential filter family as defined in (20), $\bar{f} = \sum_{j,k,\psi} \bar{c}_{j,k,\psi}\psi_{j,k}$ be a noisy image perturbed with i.i.d. Gaussian noise, and λ be a positive parameter. The function*

$$\hat{g}_\lambda = \sum_{j,k,\psi} \hat{d}_{j,k,\psi_\lambda}\ \psi_{j,k},$$

given by

$$\hat{g}_\lambda = \sum_{j,k,\psi} Fe_{j,k,\psi}(\lambda)\bar{c}_{j,k,\psi}\psi_{j,k} = \sum_{j,k,\psi} a^{-h(\lambda^2/\bar{c}^2_{j,k,\psi})}\bar{c}_{j,k,\psi}\psi_{j,k} \tag{22}$$

is a near-minimizer *of the variational problem (13) with constant $C = 4$.*

Proof. As we have already said, in the wavelet domain problem (15) is separable; hence, its minimum is obtained by minimizing each term of the sum, namely,

$$E(d_{j,k,\psi}) = |\bar{c}_{j,k,\psi} - d_{j,k,\psi}|^2 + 2\lambda|d_{j,k,\psi}| \tag{23}$$

separately over $d_{j,k,\psi}$.

Without loss of generality, we can assume $0 \leqslant d_{j,k,\psi} \leqslant \bar{c}_{j,k,\psi}$.

We remark that, if $|d_{j,k,\psi}| \leqslant |\bar{c}_{j,k,\psi}|/2$ then the greatest term in (23) is $|\bar{c}_{j,k,\psi} - d_{j,k,\psi}|^2$ and therefore $E(d_{j,k,\psi}) \geqslant |\bar{c}_{j,k,\psi}|^2/4$; on the other hand, if $|d_{j,k,\psi}| \geqslant |\bar{c}_{j,k,\psi}|/2$, then the greatest term in (23) is $2\lambda|d_{j,k,\psi}|$ and hence $E(d_{j,k,\psi}) \geqslant 2\lambda|\bar{c}_{j,k,\psi}|/2$. Therefore, for each choice of $d_{j,k,\psi}$, a lower bound for $E(d_{j,k,\psi})$ is given by $\min(|\bar{c}_{j,k,\psi}|^2/4, 2\lambda|\bar{c}_{j,k,\psi}|/2)$.

If we now choose $d_{j,k,\psi}$ as

$$\hat{d}_{j,k,\psi} = Fe_{j,k,\psi}(\lambda)\bar{c}_{j,k,\psi} = (a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})\bar{c}_{j,k,\psi}, \quad a > 1 \text{ and } h > 0, \tag{24}$$

we find that, if $|\hat{d}_{j,k,\psi}| \leqslant |\bar{c}_{j,k,\psi}|/2$, then $E(\hat{d}_{j,k,\psi}) \geqslant 4(a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})^2|\bar{c}_{j,k,\psi}|^2/4$, while if $|\hat{d}_{j,k,\psi}| \geqslant |\bar{c}_{j,k,\psi}|/2$, then $E(\hat{d}_{j,k,\psi}) \geqslant 4\lambda(a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})|\bar{c}_{j,k,\psi}|/2$.

In order to estimate the lowest value of the constant C which satisfies (21), we take

$$\begin{aligned}
E(\hat{d}_{j,k,\psi}) &= \min(4(a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})^2\bar{c}^2_{j,k,\psi}/4, 4\lambda(a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})|\bar{c}_{j,k,\psi}|/2)\\
&= \min((a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})^2\bar{c}^2_{j,k,\psi}, 2\lambda(a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})|\bar{c}_{j,k,\psi}|)\\
&\leqslant 4\min((a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})^2\bar{c}^2_{j,k,\psi}/4, 2\lambda(a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}})|\bar{c}_{j,k,\psi}|/2)\\
&\leqslant 4\min(\bar{c}^2_{j,k,\psi}/4, 2\lambda|\bar{c}_{j,k,\psi}|/2)\\
&\leqslant 4\left(\min_{d_{j,k,\psi}} E(d_{j,k,\psi})\right) = 4E(d^*_{j,k,\psi}),
\end{aligned}$$

since the quantity $(a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}}) \leqslant 1$. Hence solution (22) is a *near-minimizer* of problem (13) with constant $C = 4$. □

In a similar way, we can prove the following result regarding the quadratic filter.

Theorem 4.2. *Let $Fp_{j,k,\psi}$ be the quadratic filter as defined in* (19), $\bar{f} = \sum_{j,k,\psi}\bar{c}_{j,k,\psi}\psi_{j,k}$ *be a noisy image perturbed with i.i.d. Gaussian noise, and λ be a positive parameter. The function*

$$\hat{g}_\lambda = \sum_{j,k,\psi}\hat{d}_{j,k,\psi,\lambda}\ \psi_{j,k},$$

given by

$$\hat{g}_\lambda = \sum_{j,k,\psi} Fp_{j,k,\psi}(\lambda)\bar{c}_{j,k,\psi}\psi_{j,k} \tag{25}$$

is a near minimizer *of the variational problem* (13) *with constant $C = 4$.*

5. The choice of the threshold level and numerical results

In this section we test the proposed new filters on several test images. In our numerical simulations we use interval adapted Daubechies orthonormal wavelet bases with 4 vanishing moments [3]. The quality of the results is estimated by means of two classical parameters, the root mean square error (RMSE) and the peak signal to noise ratio (PSNR), defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\|f - \hat{g}_\lambda\|^2_{L_2}}{N}} \quad \text{and} \quad \mathrm{PSNR} = 20\log_{10}\frac{255}{\mathrm{RMSE}},$$

where f is the original non-corrupted image, $\hat{g}_\lambda$ is the denoised image, and N is the number of image pixels. We consider the 256 grey level images shown in Fig. 3, with dimensions 256×256 and 512×512. The test images have been perturbed with i.i.d. Gaussian Noise with different values of standard deviation ($\sigma = 15, 25, 35$).

The experimentation is performed using two different methods for the choice of the thresholding parameter λ. Actually, both the exact minimizer (given by the Soft thresholding) and our approximate minimizers (given by quadratic or exponential thresholding) are strongly dependent on the choice

Fig. 3. Test images: (a) 512 × 512 Lena; (b) 256 × 256 Church of San Vitale.

Table 1
Soft thresholding and quadratic filter are compared: in both cases the universal choice λ_u for the threshold has been considered

		Soft thresholding		Quadratic filter	
	λ_u	RMSE	PSNR	RMSE	PSNR
Lena 512 × 512					
$\sigma = 15$	74.92	12.34	26.29	10.47	27.72
$\sigma = 25$	124.88	15.46	24.34	13.20	25.71
$\sigma = 35$	174.83	17.87	23.08	15.23	24.47
San Vitale 256 × 256					
$\sigma = 15$	70.64	21.52	21.47	18.52	22.77
$\sigma = 25$	117.74	26.67	19.60	23.35	20.76
$\sigma = 35$	164.83	30.23	18.52	26.29	19.73

of λ. We have, therefore, considered two of the possible approaches known in the literature and, for each parameter choice, we have compared the results obtained using our new filters with those given by the soft thresholding method.

As first choice for the thresholding parameter, we consider the universal threshold proposed by Donoho and Johnstone [8] and given by $\lambda_u = \sigma\sqrt{2 \ln N}$, where N is the original number of wavelet coefficients and σ^2 the variance of the noise. Even if this value seems to be a good choice for the denoising of several 1-D signals [8], when working with images characterized by sharp edges, it can lead to oversmoothing, if the Soft thresholding technique is used. On the other hand, if we use this threshold with the quadratic filter, which gives more importance to the kept coefficients, we obtain better results, since the undersmoothing produced by this new filter "balances" the possible oversmoothing coming from the use of the Donoho and Johnstone threshold. In Table 1 we show the numerical results obtained using the universal threshold and both the classical Soft Thresholding linear filter function and the new quadratic one. It is evident that, when dealing with this choice of threshold, the new quadratic filter outperforms the Soft thresholding both in terms of PSNR and visual quality, as shown by Figs. 5(c,d) and 6(c,d).

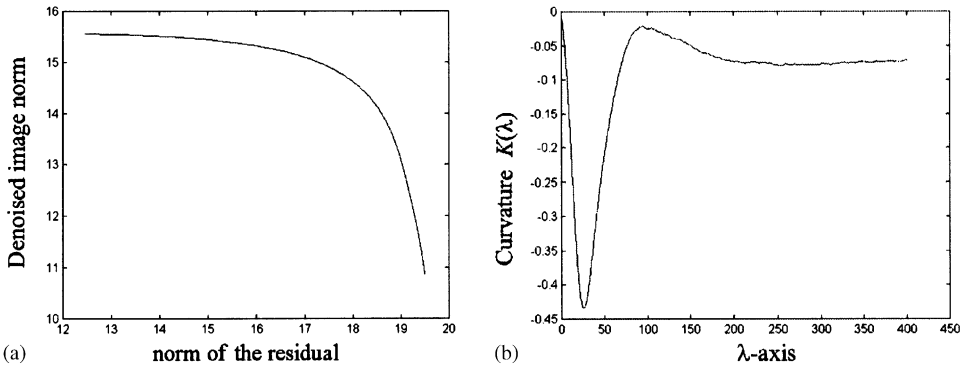

Fig. 4. (a) Parametric plot of the H-curve; (b) curvature of the H-curve as function of λ.

Next, we consider the H-curve criterion which has been shown to be a promising technique for the evaluation of a "good" thresholding parameter in [10]. This method arises from a deterministic approach to the denoising problem and takes into account not only the approximation error, but also the norm of the recovered image. A convenient way to do this is to plot, in the logarithmic scale, the norm of the filtered solution versus the norm of the corresponding residual, obtaining a curve parametrized by λ. This curve is a concave *Hook-shaped curve*, with a well localized maximum of the curvature absolute value. The H-curve method consists in choosing the threshold parameter λ corresponding to this maximum. This technique, created in a wavelet context where the exact solution was the one obtained by applying the Soft thresholding filter, can easily be generalized by considering our new filter functions.

As the C^∞-regularity of the exponential filter family allows us to analyze the shape of the curvature of the H-curve very easily, we consider this filter. Let $\rho(\lambda)$ and $\eta(\lambda)$ be defined as

$$\rho(\lambda) := \| \bar{f} - \hat{g}_\lambda \|^2_{L_2(I)} \quad \text{and} \quad \eta(\lambda) := \|\hat{g}_\lambda\|_{B^1_1(L_1(I))}, \tag{26}$$

where $\hat{g}_\lambda$ is the filtered approximation defined by relation (22). If we set

$$\hat{\rho}(\lambda) = \log \rho(\lambda), \quad \hat{\eta}(\lambda) = \log \eta(\lambda), \tag{27}$$

the plot of $\hat{\eta}(\lambda)$ versus $\hat{\rho}(\lambda)$ is shown in Fig. 4(a). By substituting $\bar{f}$ and $\hat{g}_\lambda$ with their wavelet expansions and using relations (11) and (12) we have

$$\rho(\lambda) = \sum_{j,k,\psi} [(1 - Fe_{j,k,\psi}(\lambda))\bar{c}_{j,k,\psi}]^2 \tag{28}$$

and

$$\eta(\lambda) = \sum_{j,k,\psi} |Fe_{j,k,\psi}(\lambda)\bar{c}_{j,k,\psi}|, \tag{29}$$

where $Fe_{j,k,\psi}$ are the exponential filter factors defined in relation (20).

As the computation of the curvature of the parametric curve $(\hat{\rho}(\lambda), \hat{\eta}(\lambda))$ involves the first and the second derivatives of $\hat{\rho}(\lambda)$ and $\hat{\eta}(\lambda)$, we exploit the C^{∞}-regularity of the exponential filter functions to compute the analytical expression of this curvature as a function of λ. In fact, we have

$$\rho'(\lambda) := \frac{\mathrm{d}\rho(\lambda)}{\mathrm{d}\lambda} = -2\sum_{j,k,\psi}(1 - Fe_{j,k,\psi}(\lambda))\frac{\mathrm{d}Fe_{j,k,\psi}(\lambda)}{\mathrm{d}\lambda}\,\bar{c}^2_{j,k,\psi}, \tag{30}$$

$$\eta'(\lambda) := \frac{\mathrm{d}\eta(\lambda)}{\mathrm{d}\lambda} = \sum_{j,k,\psi}\frac{\mathrm{d}Fe_{j,k,\psi}(\lambda)}{\mathrm{d}\lambda}\,|\bar{c}_{j,k,\psi}|. \tag{31}$$

By recalling relation (20), we can rewrite relations (30) and (31) as follows:

$$\rho'(\lambda) = 4h\lambda \ln(a)\left(\sum_{j,k,\psi}(a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}} - a^{-2h\lambda^2/\bar{c}^2_{j,k,\psi}})\right), \tag{32}$$

$$\eta'(\lambda) = -2h\lambda \ln(a)\sum_{j,k,\psi}\frac{a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}}|\bar{c}_{j,k,\psi}|}{\bar{c}^2_{j,k,\psi}}. \tag{33}$$

We therefore obtain

$$\rho''(\lambda) := \frac{\mathrm{d}\rho'(\lambda)}{\mathrm{d}\lambda} = 4h\ln(a) \times \left(\sum_{j,k,\psi}\frac{4h\lambda^2\ln(a)a^{-2h\lambda^2/\bar{c}^2_{j,k,\psi}} - 2h\lambda^2\ln(a)a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}} + \bar{c}^2_{j,k,\psi}a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}} - \bar{c}^2_{j,k,\psi}a^{-2h\lambda^2/\bar{c}^2_{j,k,\psi}}}{\bar{c}^2_{j,k,\psi}}\right) \tag{34}$$

Table 2
Soft thresholding and the exponential filter with $a = 2$ and $h = 4$ are compared: in both cases the H-curve threshold λ_{H} has been considered

	Soft thresholding			Exponential filter		
	λ_{H}	RMSE	PSNR	λ_{H}	RMSE	PSNR
Lena 512×512						
$\sigma = 15$	17.88	7.71	30.38	19.66	7.32	30.83
$\sigma = 25$	26.5	11.71	26.75	30.28	10.71	27.52
$\sigma = 35$	34.5	16.13	23.97	42.09	13.88	25.28
San Vitale 256×256						
$\sigma = 15$	26.30	13.12	25.76	26.99	12.74	26.02
$\sigma = 25$	33.51	16.14	23.96	36.20	15.91	24.09
$\sigma = 35$	39.08	19.76	22.21	46.55	19.06	22.52

Fig. 5. Test on image Lena 512×512: (a) original noncorrupted image; (b) noisy image $\sigma = 25$; (c) denoised image using DB4 soft thresholding and the universal threshold; (d) denoised image using DB4, quadratic filter and the universal threshold; (e) denoised image using DB4, soft thresholding and the H-curve threshold; (f) denoised image using DB4, exponential filter and the H-curve threshold.

and

$$\eta''(\lambda) := \frac{\mathrm{d}\eta'(\lambda)}{\mathrm{d}\lambda} = 2h \ln(a) \left(\sum_{j,k,\psi} \frac{a^{-h\lambda^2/\bar{c}^2_{j,k,\psi}} (2h\lambda^2 \ln(a) - \bar{c}^2_{j,k,\psi}) |\bar{c}_{j,k,\psi}|}{\bar{c}^4_{j,k,\psi}} \right). \tag{35}$$

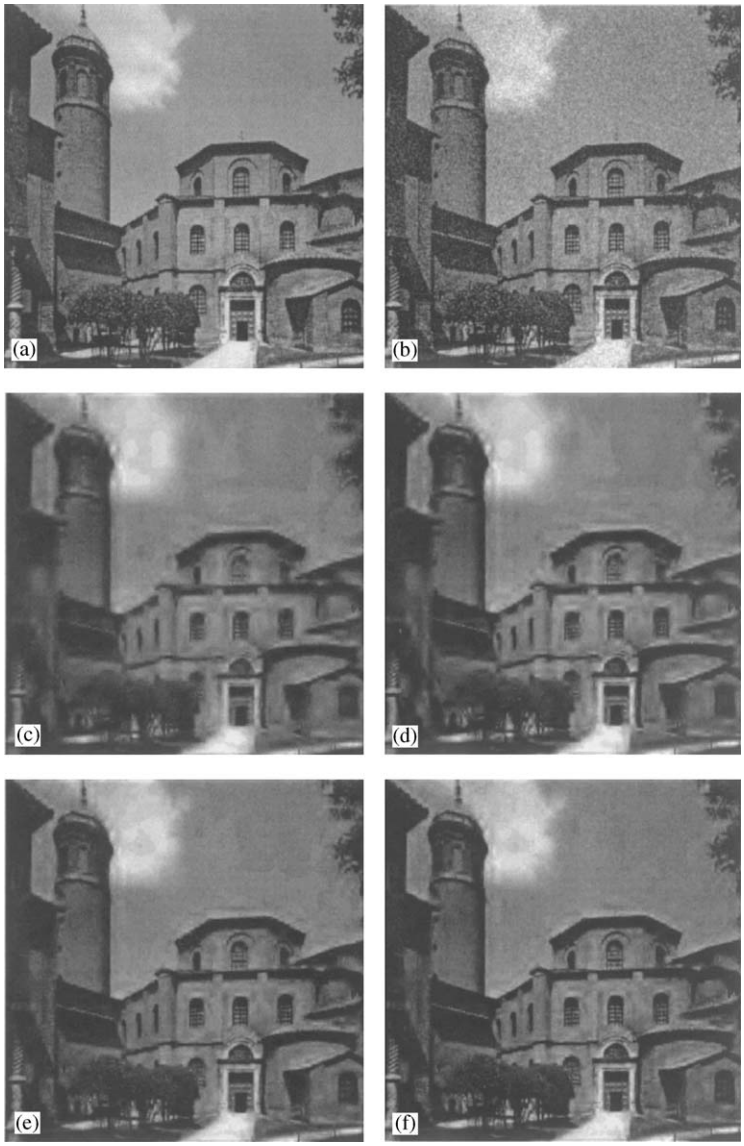

Fig. 6. Test on image San Vitale 256×256: (a) original noncorrupted image; (b) noisy image $\sigma = 15$; (c) denoised image using DB4 Soft thresholding and the universal threshold; (d) denoised image using DB4, quadratic filter and the universal threshold; (e) denoised image using DB4, soft thresholding and the H-curve threshold; (f) denoised image using DB4, exponential filter and the H-curve threshold.

By remembering that

$$\hat{\eta}'(\lambda) = \frac{\mathrm{d}\hat{\eta}(\lambda)}{\mathrm{d}\lambda} = \frac{\eta'(\lambda)}{\eta(\lambda)}, \quad \hat{\rho}'(\lambda) = \frac{\mathrm{d}\hat{\rho}(\lambda)}{\mathrm{d}\lambda} = \frac{\rho'(\lambda)}{\rho(\lambda)}$$

and

$$\hat{\eta}''(\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\eta'(\lambda)}{\eta(\lambda)} = \frac{\eta''\eta - (\eta')^2}{\eta^2},$$

$$\hat{\rho}''(\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\rho'(\lambda)}{\rho(\lambda)} = \frac{\rho''\rho - (\rho')^2}{\rho^2},$$

and recalling the expression of the algebraic curvature given by

$$K(\lambda) = \frac{\hat{\rho}'\hat{\eta}'' - \hat{\rho}''\hat{\eta}'}{[(\hat{\rho}')^2 + (\hat{\eta}')^2]^{3/2}}, \tag{36}$$

we obtain the analytic expression of our curvature function. Although it presents a quite intricate structure, it can easily be implemented in Matlab language. Moreover, even if it is difficult to study the signum of the curvature function, due to its complexity, it can be experimentally verified that the H-curve is concave and that its curvature has a well localized minimum as shown in Fig. 4(b).

As the H-curve criterion consists in choosing the parameter λ which corresponds to this minimum, in our numerical experiments we found this value by means of the well-known Golden Section algorithm [10].

In Table 2 we again compare the numerical results obtained using the H-curve threshold value for both the classical Soft thresholding linear filter and the exponential filter (20) with $a = 2,\ h = 4$. Also these experiments highlighted a better performance of the new filter with respect to the Soft Thresholding both in terms of PSNR and visual quality as shown by Figs. 5(e, f) and 6(e, f).

6. Uncited references

[1,4,9]

References

[1] J.-P. Antoine, P. Carrette, R. Murenzi, B. Piette, Image analysis with two-dimensional continuous wavelet transform, Signal Process. 31 (3) (1993) 241–272.
[2] A. Chambolle, R.A. DeVore, N. Lee, B.J. Lucier, Nonlinear wavelet image processing: variational problem, compression and noise removal through wavelet shrinkage, IEEE Trans. Image Process. 7 (1998) 319–335.
[3] A. Cohen, I. Daubechies, P. Vial, Wavelets on the interval and fast wavelet transforms, Appl. Comput. Harmonic Anal. 1 (1993) 54–81.
[4] A. Cohen, R.A. DeVore, P. Petrushev, H. Xu, Nonlinear Approximation and the Space BV(R2), 1997.
[5] I. Daubechies, Ten lectures on wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 61, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.
[6] R.A. Devore, B.J. Lucier, Fast Wavelet Techniques for Near-Optimal Image Processing, IEEE Press, New York, 1992, pp. 1129–1135.
[7] D. Donoho, Wavelet shrinkage and W.V.D: a 10 minute tour, Progress in Wavelet Analysis and Applications, Editiones Frontieres, Dreuk, 1992, pp. 109–128.
[8] D. Donoho, I. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika 81 (1994) 425–455.
[9] D. Donoho, I. Johnstone, G. Kerkyacharian, D. Picard, Wavelet shrinkage: asymptopia?, J. Roy. Statist. Soc. Ser. B 57 (1995) 301–369.
[10] L.B. Montefusco, S. Papi, A variational approach to wavelet shrinkage, to appear on BIT.