Electronic Letters on Computer Vision and Image Analysis 6(2):36-53, 2008

Wavelets and partial differential equations for image denoising

V. Bruni, B. Piccoli and D. Vitulano

Istituto per le Applicazioni del Calcolo "M. Picone", C.N.R., Viale del Policlinico 137, 00161 Rome, Italy

Received 2 October 2006; accepted 18 January 2008

Abstract

In this paper a wavelet based model for image de-noising is presented. Wavelet coefficients are modelled as waves that grow while dilating along scales. The model establishes a precise link between corresponding modulus maxima in the wavelet domain and then allows to predict wavelet coefficients at each scale from the first one. This property combined with the theoretical results about the characterization of singularities in the wavelet domain enables to discard noise. Significant structures of the image are well recovered while some annoying artifacts along image edges are reduced. Some experimental results show that the proposed approach outperforms the most recent and effective wavelet based denoising schemes.

Key Words: Image Restoration, Wavelets, Scale Space Analysis.

1 Introduction

Denoising represents a stimulating challenge in image processing, as proved by the amount of proposals for its solution — see for instance [4, 5, 9, 14, 16, 21, 26, 29, 30, 33, 39]. The main goal consists of recovering a signal f from its noisy observation g, corrupted by an additive zero-mean Gaussian noise ν with variance σ^2, i.e.

$$g(t) = f(t) + \nu(t), \quad t \in \mathbf{R}. \tag{1}$$

One of the objectives of de-noising is the preservation of the main features of the original signal. Following this philosophy, a lot of researchers devoted their study to the construction of more or less sophisticated bases able to catch and well represent image correlations with few coefficients. The better the compaction of the basis, the simpler the restoration strategy — significant coefficients can be retained using a thresholding operation. Nonetheless, it is difficult to detect the significant structures of a signal without having additional information about it. The search of singularities and irregular points is then crucial since they characterize the signal — peaks and jump discontinuities in 1D signals, edges and objects contours in images. The main strategies for providing solutions can be coarsely split into two broad classes:

1. scale space analysis, which exploits the fact that the irregular structures are visible at different resolutions;
2. approximation theory, i.e. the construction of a suitable expansion basis which provides few coefficients with great amplitude in correspondence to singularity points and nearly zero coefficients, in correspondence to flat or regular regions.

Correspondence to: <bruni@iac.rm.cnr.it>

Recommended for acceptance by Jon Sporring
ELCVIA ISSN:1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

The former gives rise to linear and non linear scale space approaches [18, 26, 31, 35, 38], while the latter to linear and non linear approximations in orthonormal bases [6, 20, 23]. The main features of these approaches are briefly described in the next section.

It is now worth observing that the wavelet transform embeds the basics of both scale space theory and non linear approximation because of its intrinsic time-scale structure. In fact, there is a precise link between coefficients at successive scales which is described by the *persistency* property: large/small values of wavelet coefficients tend to propagate across scales [7]. However, a more theoretical contribution has been done by investigating the wavelet transform of signals having singularities of different order. In fact, the decay of modulus maxima of the wavelet transform gives a measure of the kind of signal singularity within a given interval: the *cone of influence* [22].

Nonetheless, it is difficult to build the modulus maxima chains along scales in a deterministic way. In fact, modulus maxima can change their locations and they can assume different appearance whenever the cones of influence of two different singularities overlap. Hence, some empirical constraints have to be used for building the chain, such as the persistency of the sign and the definition of one global maximum in the cone of influence [22, 24] — see [19] for a similar approach in the time domain using the catastrophe theory. This leads to some false alarms or the lack of some important information. A precise estimate of the location becomes more important in case of noisy signals, as it will be clearer in Section 4, since the signal has to be reconstructed from the detected modulus maxima.

The maxima projection algorithm proposed in [24] is useful but it presents three main drawbacks:

- it could not converge to the original signal (counterexample to Mallat conjecture in [20]);
- the convergence of the algorithm requires a minimum distance between two successive maxima [22];
- maxima chains are guaranteed for each scale s only using wavelets which are derivative of a Gaussian kernel (see Theorem 2 in Chapter 1 and Chapter VI of [23]).

These drawbacks can be solved by further characterizing each modulus maximum. A first attempt has been done by Dragotti and Vetterli in [12], who tried to exactly model piecewise polynomial signals. Nonetheless, the distance between two adjacent singularities becomes, again, crucial for distinguishing them. In other words, *footprints* are not able to discriminate two singularities when they interfere.

The main contribution of this paper consists of a theoretical model that allows to estimate the trajectories of wavelet modulus maxima of a signal f. These trajectories model the evolution law of some predefined basic atoms whose superimposition approximates the wavelet transform $w(u, s)$ of the signal f. For each atom (see Fig. 1), the significant maximum is the one having the greatest amplitude [2]. The latter does not disappear along scales but it moves from its initial location whenever its relative atom interferes with an adjacent one. In the case of complete interference the two atoms can generate an only one maximum which takes into account both contributions.

This representation preserves and exploits the singularity characterization of the wavelet transform, the clustering property of wavelet coefficients and their parent-child relationships. Furthermore, it drastically reduces the redundancy of the wavelet representation in correspondence to significant structures of the signal and avoids particular requirements on the smoothness of the original signal. Experimental results show that these properties allow to achieve comparable results to the state of the art of wavelet based denoising approaches, just by modeling images as 1D independent signals.

The paper is organized as follows. Section 2 gives a brief review about the main denoising strategies. Section 3 introduces the evolution law of wavelet atoms and yields their trajectories along scales. Section 4 shows how these laws can be successfully used for de-noising. Some experimental results and comparative studies are then given in Section 5. Finally, Section 6 draws the conclusions.

2 A brief review of image de-noising approaches

Scale space analysis Scale space can be generated using linear and non linear schemes. The former have been introduced by Koenderink [18] and Witkin [38]. They embed the original signal f in a family of functions which are blurred versions of the original signal. More precisely,

$$\begin{cases} I_s = c\Delta I \\ I(t,0) = f(t), \end{cases}$$

where Δ is the Laplacian operator and c is a constant (*conductance term*). It is a linear smoothing and obeys to the heat equation. The unique kernel satisfying the equation is the Gaussian one. Moreover, since the derivatives of I satisfy the heat equation, also Gaussian derivatives generate a scale space.

Wavelets define a scale space since they provide a signal representation at different levels of resolution. In particular, if they are derivative of a Gaussian kernel, they generate the heat flow (chap. VI of [23]).

Linear scale space is totally insensitive to the presence of relevant image features — for example edges. Hence, even if it is able to suppress noise, it also destroys some geometrical information of the image. Approaches based on non linear scale space try to retain these features by making the conductance term non linear. In other words, they guide the smoothing: it is allowed in regular parts while it is inhibited or reduced in correspondence to significant structures of the function. Different conduction terms imply the preservation of specific structures*. For example, the anisotropic diffusion of Perona Malik [26] consists of embedding an edge detection step in a partial differential equation (pde) model which forces smoothing within homogeneous regions, letting their boundaries sharp. The pde which regulates this process is

$$I_s = div(\upsilon(\|\nabla I\|)\nabla I),$$

where υ is the edge detector function and corresponds to a thresholding operator on the gradient magnitude.

It is possible to mention various pdes having slightly different properties, such as the *mean curvature* or the *total variation flow* [27]. Nonetheless, most of them share the fact that there is not a closed form of their solution. Hence, they require numerical methods for being solved. Moreover, high and low frequency components are not completely separated during a non linear diffusion scheme.

In a recent work [30], it has been shown that a discrete computation step in a non linear diffusion process can be split into three stages: decomposition, regularization and reconstruction. The first and third stages correspond to a two bands filtering system — a decomposition and reconstruction using a basis of wavelets. It turns out that just the high frequency component is regularized by means of a diffusivity function, while low frequency is preserved. This idea follows the same philosophy of the more recent approach presented in [15], where the splitting into low and high pass components is achieved by using the complex domain and combining the diffusion equation with the Schroedinger one. Furthermore, in [37] the authors investigate the connections between 1D discrete schemes for non linear diffusion and shift invariant Haar wavelet shrinkage. In particular, they prove that each diffusivity term corresponds a single spatial level shrinkage function in the wavelet domain and viceversa.

Common drawbacks for these methods are the stopping time in the numerical solution of the pde and the choice of the best diffusivity function for the analysed image. In this paper we will show that it is possible to avoid the direct solution of the pde exploiting a suitable representation of wavelet coefficients and deriving an ordinary differential equation (ode) for describing maxima chains along scales.

*It is similar to the choice of the number of vanishing moments of a wavelet for processing 1D signals, or the choice of suitable 2D basis that is able to preserve geometrical features of the image, as it will be clearer in the following.

Linear and non linear approximation A linear approximation projects f on N vectors, selected a priori, of an orthonormal basis B. An example is the expansion in a Fourier basis when the first N vectors are used for recovering the signal. On the contrary, non linear approximations select the N vectors of B that better correlate f. More precisely, they are the N largest coefficients of the expansion. Non linear approximations are equivalent to apply a thresholding operator to the sorted inner products. It is obvious that the approximation error goes quickly to zero as N increases and if the sorted values have a fast decay [6, 20, 23]. Therefore, the choice of the best expansion basis for a signal is strictly connected to its smoothness.

The literature offers a variety of bases. The well known and widely used are *wavelets* [23], whose expansion coefficients are characterized by the *inter* and *intra-scale* properties and are used for both 1D and 2D signals. More recent 2D geometrical bases are also available, such as: *curvelets* [34], which give an optimal approximation for 2D piecewise smooth functions having C^2 discontinuity curves; *contourlets* [8], that can be considered a flexible discrete evolution of the first ones; *edgeprints* [11], that are a 2D extension of *footprints* and are able to represent 2D piecewise polynomial signals; *bandlets* [13], that follow the *geometric flow* of the image, i.e. local directions in which the image grey levels have regular variation; and the more recent *directionlets* [36], which are anisotropic basis functions with directional vanishing moments.

The construction of these bases can often be complicated and computationally expensive. They are good when denoising can be achieved through a simple thresholding operation by using, for example, the pioneering Donoho *universal threshold* [10]. Moreover, they lose their compaction ability whenever the regularity of the curves in the analysed image does not match with the one they can well represent.

A valid alternative can be the use of simple and fast decomposition bases, like wavelets, and the processing of the corresponding coefficients with more or less sophisticated methods. Most proposals are based on:

- *adaptive thresholding* of coefficients: the threshold is estimated according to signal statistics. For example, in [4] a nearly optimal approximation of the best threshold is achieved by a pixel-wise estimation of the signal variance by means of *context modelling*[†]. Another example is in [1], where two thresholds are used: one for the magnitude of coefficients and the other one for the amount of clustering;
- *adaptive shrinkage*: coefficients are filtered by shrinking their value according to the signal to noise ratio. For instance in [29], a certain Gaussian distribution of wavelet coefficients is assumed and maximum likelihood is employed for getting correlation matrices; in [33] local statistics of the signal are estimated using an adaptive window; or in [32], where a bivariate shrinkage rule using the analysed coefficient, its parent and the local neighborhood is applied.

Previous strategies exploit some of the main three properties of the wavelet decomposition, i.e. amplitude of coefficients, evolution across scales, spatial clustering near image edges. In [28] they are all embedded in a Bayesian framework by further exploiting the characterization of singularities in the wavelet domain. In fact, as proved in [22], the Lipschitz order of a singularity is connected to the decay of the modulus maxima of the wavelet transform for increasing scales. This result allows to discriminate between noise and original signal since noise has negative Lipschitz order. It means that the corresponding coefficients have an opposite behaviour along scales.

A good measure of Lipschitz exponents is related to the choice of the wavelet: the higher the vanishing moments the better the decay estimation [23] and the prediction of modulus maxima chains along scales.

[†]Context modelling technique allows to group pixels of similar nature but not necessarily spatially adjacent, gathering image statistical information from them.

Dragotti and Vetterli[12] proposed to concentrate on a particular class of signals, the piecewise polynomial functions, in order to exploit the time-scale correlation of the corresponding wavelet coefficients. Nonetheless, their model requires a minimum distance between two singularities for getting a good estimation of the original signal.

In this paper we overcome this drawback using a suitable representation of wavelet coefficients. Since it is not possible to know the structure of the signal a priori, the idea is to fix a basis and a reference singularity and then to approximate the analysed signal as superimposition of these basic singularities. This representation preserves and exploits the singularity characterization, the clustering property (intra-scale) and parent child relationships (inter-scale) of the wavelet transform. In fact, maxima chains corresponding to singularity points can be built along scales by modelling the wavelet transform as a combination of interfering atoms, i.e. waves obeying to a precise partial differential equation (pde).

3 Building maxima chains

As previously mentioned, it is important to model the evolution of the information along scale levels. It can be done by characterizing the wavelet transform with its absolute maxima and deriving the maxima chains along scales.

Let ψ be a real and continuous wavelet and set $\overline{\psi}(u,s)$

$$\overline{\psi}(u,s) = \frac{1}{\sqrt{s}}\psi\left(-\frac{u}{s}\right), \tag{2}$$

where $s \in \mathbf{R}^+$ is the scale variable and $u \in \mathbf{R}$ is the time variable. Since the wavelet transform of a function f is

$$w(u,s) = f * \overline{\psi}(u,s),$$

using some algebraic computations [3], the following pde can be written

$$w_s = -\frac{u}{s}w_u - \frac{1}{2s}w + \frac{1}{s}v_u, \tag{3}$$

where v is the wavelet transform of the function $tf(t)$ — see [3] for details. This equation shows two different effects. The first term of the second member guides a sort of transport along the scale s, while the remaining ones guide the decay and the shape of the wavelet transform along scales.

Although the generality of the result, equation (3) is useless in this form since there is not a priori information about the function f.

On the contrary, if we approximate the wavelet transform $w(u,s)$ as superimposition of basic atoms, the equation becomes more manageable. In fact, let us define a *basic atom* at scale s and centered at the location t_1, the one described by the following function

$$F(t_1,u,s) = s\sqrt{s}\left(\int_{\frac{t_1-u}{s}}^{b} t\psi(t)dt - \frac{t_1-u}{s}\int_{\frac{t_1-u}{s}}^{b}\psi(t)dt\right). \tag{4}$$

It corresponds to the wavelet transform of an infinite ramp signal having the singularity located at t_1 with slope $\alpha_1 = 1$. The shape of the atom depends on the adopted wavelet. For simplicity we will use a biorthogonal wavelet which yields the atom as in Fig. 1. The adopted wavelet is nice thanks to its symmetry, its analytical formulation and allows a simple algorithm, as it will be clearer later.

Therefore, if

$$\forall\, s > 0, \quad w(u,s) \sim \sum_{k=1}^{N} \alpha_k F(t_k,u,s), \tag{5}$$

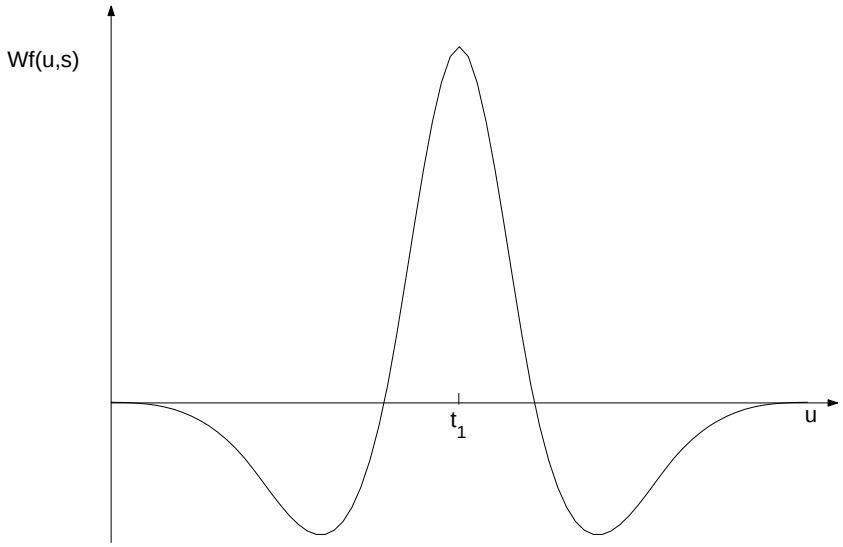

Figure 1: Plot of a basic atom corresponding to a ramp signal — a spline biorthogonal wavelet 3/9 has been used.

where α_k and t_k are respectively the slope and the location of each of the N atoms of the representation, (3) becomes

$$w_s = \frac{\overline{t} - u}{s} w_u + \frac{3}{2s} w + \frac{1}{s} \sum_{k=1}^{N} \left[\alpha_k d_k \int_{\left(\frac{t_k - u}{s}\right)}^{+\infty} \psi(y) dy \right], \tag{6}$$

with $\overline{t} = \frac{\sum_{k=1}^{N} t_k}{N}$ and $d_k = t_k - \overline{t}$. The transport term is equal to the one in eq. (3) while the source term explicitly reveals its influence on the locations of critical points. It is worth noticing that equation (5) is an equality for piecewise linear signals. In the other cases, it can be simply proved that it corresponds to a piecewise approximation through hat functions in the time domain — see [3] for details.

From equation (6) it is not difficult to get modulus maxima trajectories. In fact, we can compute the derivative with respect to u of (6), evaluate it at $u = u(s)$ and combine it with

$$w_{us} + \dot{u} w_{uu} = 0,$$

where $u(s)$ is the analysed maximum chain.

Hence, it results

$$\dot{u} = -\frac{\overline{t} - u}{s} - \frac{1}{s} \frac{\sum_{k=1}^{N} \alpha_k d_k \psi\left(\frac{t_k - u}{s}\right)}{\sum_{k=1}^{N} \alpha_k \psi\left(\frac{t_k - u}{s}\right)}, \tag{7}$$

where $\dot{u}$ is the derivative of $u(s)$ with respect to s.

The equation describes the trajectory for each maximum point of $w(u, s)$ corresponding to a basic atom at scale $s = 1$ centered at location t_h, i.e. $u(1) = t_h, \quad h = 1, \ldots, N$. In particular, if the analysed atom does not interfere with other atoms, i.e. its cone of influence does not intersect the others at scale s, it does not change its initial location. In fact, the last term of the second member of the equation is d_h/s, since $\psi\left(\frac{t_k - u}{s}\right) = 0$ if u does not belong to the support of the wavelet centered at t_k, and then (7) becomes

$$\dot{u} = -\frac{t_h - u}{s},$$

whose solution is

$$u(s) = t_h - (t_h - u(1))s,$$

i.e.

$$u(s) = t_h.$$

On the contrary, when two or more atoms interfere, the location of their maximum changes according to the amplitude of each maximum α_k and their mutual distance d_k. In particular, the greater atom moves from its location slower than the smaller one. In Fig. 2 there is an example of maxima trajectories of a piecewise constant signal. It can be observed that adjacent atoms with different sign show a repulsion while atoms having the same sign show attraction till they become an only atom, i.e. they completely interfere.

The model can be generalized by introducing a decay exponent in the equation of the basic atom, i.e.

$$G(u,s) = \alpha_1 s^{\gamma-1} F(t_1, u, s). \tag{8}$$

This way, the atom amplitude is modulated in agreement with the decay of the analyzed singularity while its shape is approximated with eq. (4).

Even in this form, the atom obeys to a precise evolution law from which it is possible to derive maxima trajectories. Hence, for N interfering atoms located at t_k with growing exponents γ_k, $w(u,s)$ satisfies the following pde

$$\frac{\partial}{\partial s}w = \frac{\overline{t}-u}{s}\frac{\partial}{\partial u}w + \frac{\gamma+1/2}{s}w + \frac{1}{s}\sum_{k=1}^{N}\left[\alpha_k d_k s^{\gamma_k}\int_{\left(\frac{t_k-u}{s}\right)}^{+\infty}\psi(y)dy\right] + \frac{1}{s}\sum_{k=1}^{N}\left[\alpha_k\gamma_k s^{\gamma_k-1}F(t_k,u,s)\right], \tag{9}$$

while atoms trajectories are the solution of the following ode:

$$\dot{u} = -\frac{\overline{t}-u}{s} - \frac{1}{s}\frac{\sum_{k=1}^{N} d_k\alpha_k s^{\gamma_k}\psi\left(\frac{t_k-u}{s}\right)}{\sum_{k=1}^{N}\alpha_k s^{\gamma_k}\psi\left(\frac{t_k-u}{s}\right)} + \\ -\frac{1}{s}\frac{\sum_{k=1}^{N}\gamma_k\alpha_k s^{\gamma_k}\int_{\frac{t_k-u}{s}}^{+\infty}\psi(y)dy}{\sum_{k=1}^{N}\alpha_k s^{\gamma_k}\psi\left(\frac{t_k-u}{s}\right)}. \tag{10}$$

3.1 Parameters estimation

The solution of the ODE (10) is determined by the initial conditions $\{t_k, \alpha_k, \gamma_k\}_{1\le k\le N}$, which respectively are the locations, the slopes and the decay exponents of atoms at $s = 1$.

Hence, the knowledge of $\{t_k, \alpha_k, \gamma_k\}_{1\le k\le N}$ allows to predict wavelet coefficients at all successive scale levels.

α_k and t_k can be estimated at $s = 1$ using the atoms estimation algorithm in [2]. Briefly speaking, atoms slopes are estimated at each scale using a greedy algorithm on the amplitude of modulus maxima, i.e.

$$\alpha_k = \frac{< R_k(u,s), F(t_k,u,s) >}{\|F(t_k,u,s)\|^2}$$

where $R_k(u,s) = w(u,s) - \sum_{h=1}^{k-1} R_h(u,s)$.

The decay exponents can be estimated by solving (9) in a suitable interval $[1, 1+\Delta s]$. Δs has to be quite small for guaranteeing that the interference between atoms does not still affect the locations of their maxima. Under this assumption, each atom can be considered isolated and then the equation (9) can be solved for each of them, yielding

$$w_s^{(k)} = \frac{\gamma_k + 1/2}{s}w^{(k)},$$

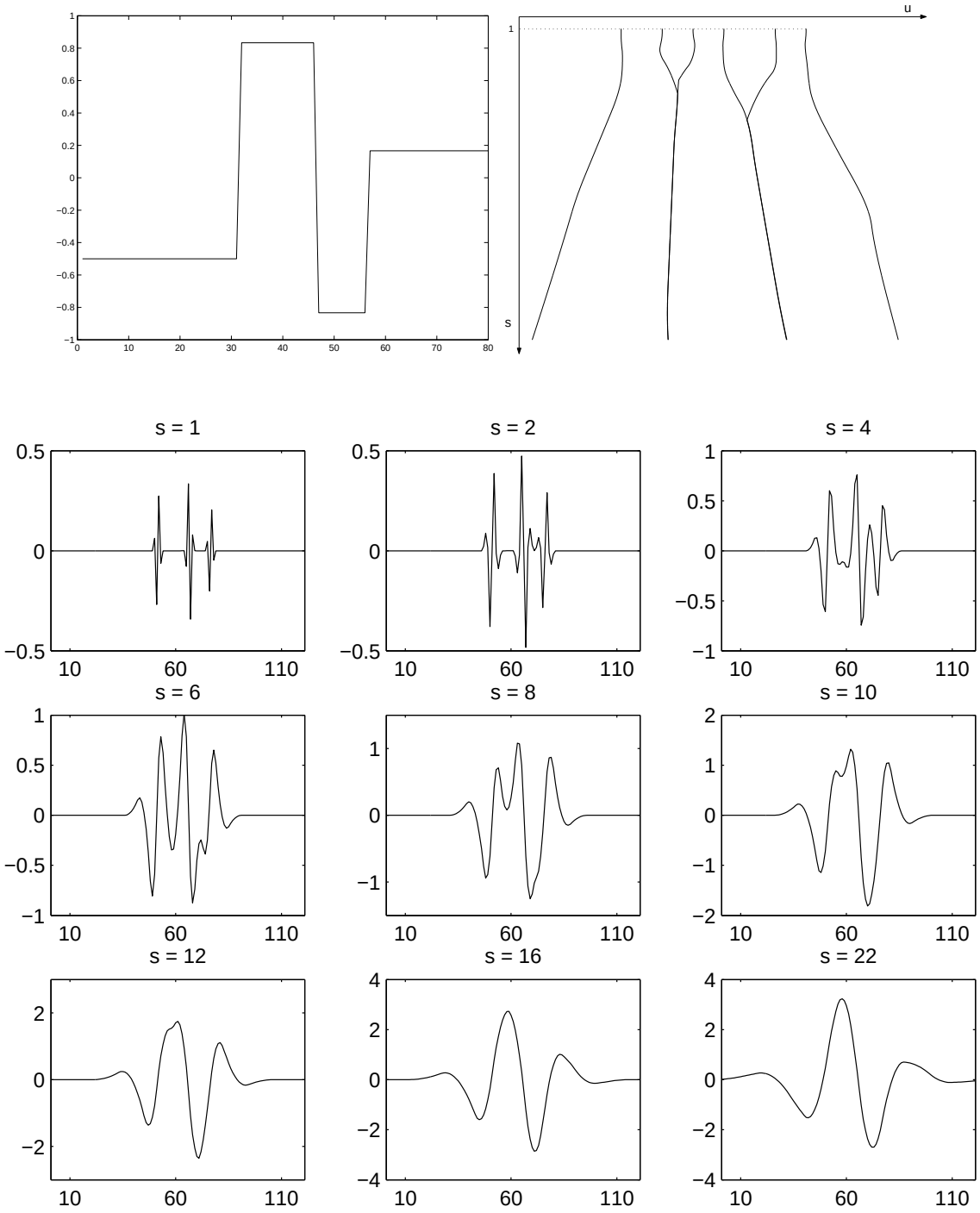

Figure 2: Piecewise constant signal having three jump discontinuities *(topleft)* . Each of them is composed of two atoms having different sign. Trajectories along scales *(topright)* of atoms modulus maxima of its wavelet transform *(bottom)* computed at scales $s = 1, 2, 4, 6, 8, 10, 12, 16, 22$. Notice that the six initial atoms interfere and produce four atoms at the coarsest considered scale.

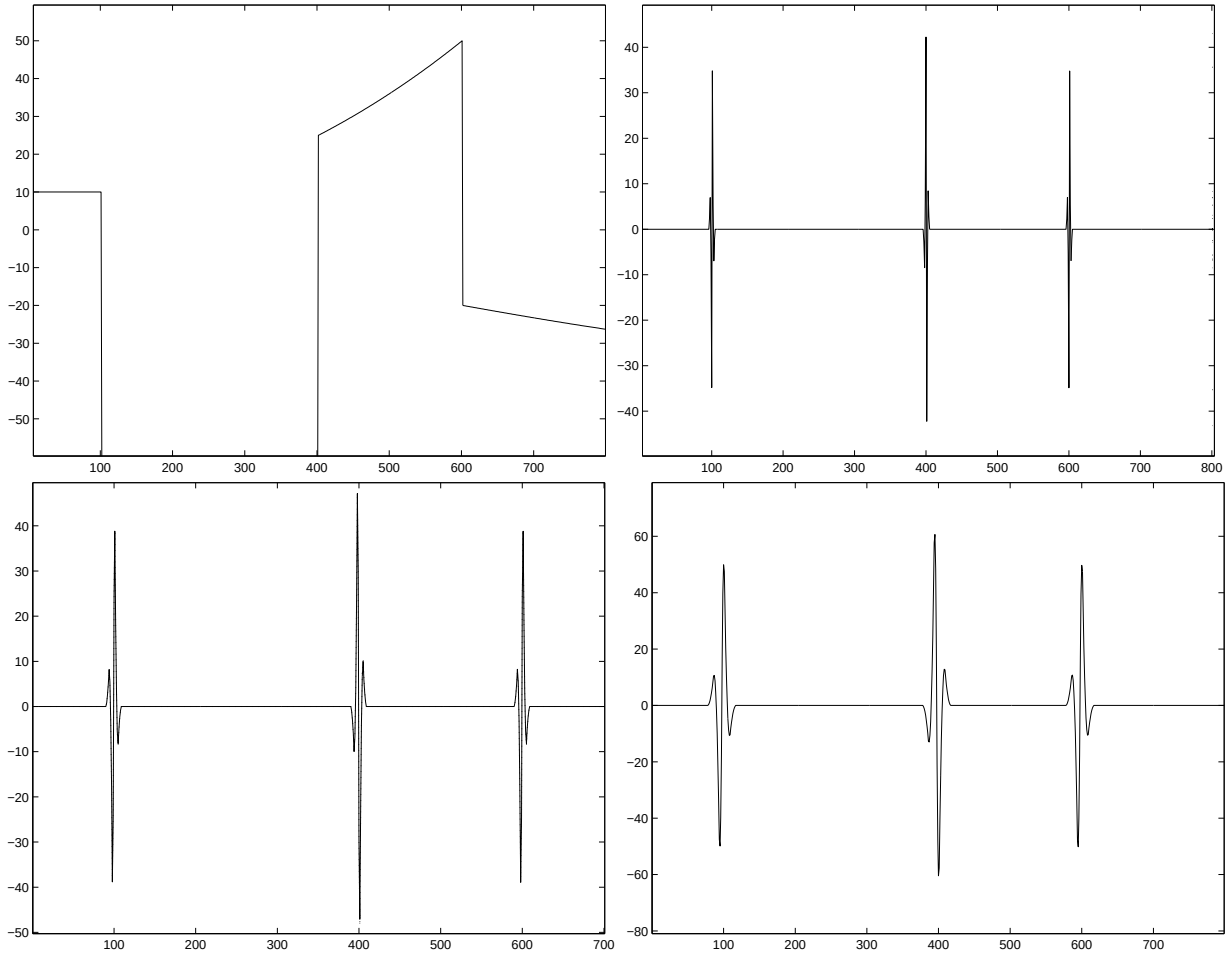

Figure 3: *Topleft)* Portion of piecewise polinomial signal. *Top-bottom left-right* Wavelet details up to third scale level reconstructed using atoms evolution law. The recovered signal gives PSNR = 60.41db.

with the following initial condition $w^{(k)}(u(1),1) = \alpha_k F(t_1,t_1,1)$. Hence, $w^{(k)} = C_0 s^{\gamma_k+1/2}$, that is $\alpha_k(s)s\sqrt{s} = \alpha_k(1)s^{\gamma_k+1/2}$, and then

$$\gamma_k = \log_s(\alpha_k(s)/\alpha_k(1)) + 1. \tag{11}$$

It is worth outlining that the estimation of γ_k depends on the precision of the estimation of the corresponding slope α_k. We can then iterate the algorithm used for slopes estimation — see [3] for details.

In Figs. 3 and 4 it is possible to see the reconstruction of the wavelet coefficients of two signals using the model. Notice that for a piecewise constant signal, the reconstruction is somewhat perfect — the recovered coefficients (dashed line) are covered by the original ones (solid line). On the other hand, for a more complicated signal like that in Fig. 4, the estimated coefficients give a quite faithful recovering of the original ones.

4 Denoising

It is worth spending some words about the importance of maxima chains in denoising. Fig. 5 depicts the noisy wavelet transform of a simple ramp signal. Noise makes various ambiguous maxima points around the atom location in the wavelet domain. It turns out that an acceptable recovering of clean coefficients requires the selection of the right maximum (indicated by the arrow) for estimating its contribution (slope) in the corresponding domain (*cone of influence*)— in Fig. 5, a wrong estimation would give an atom having a sign opposite to the

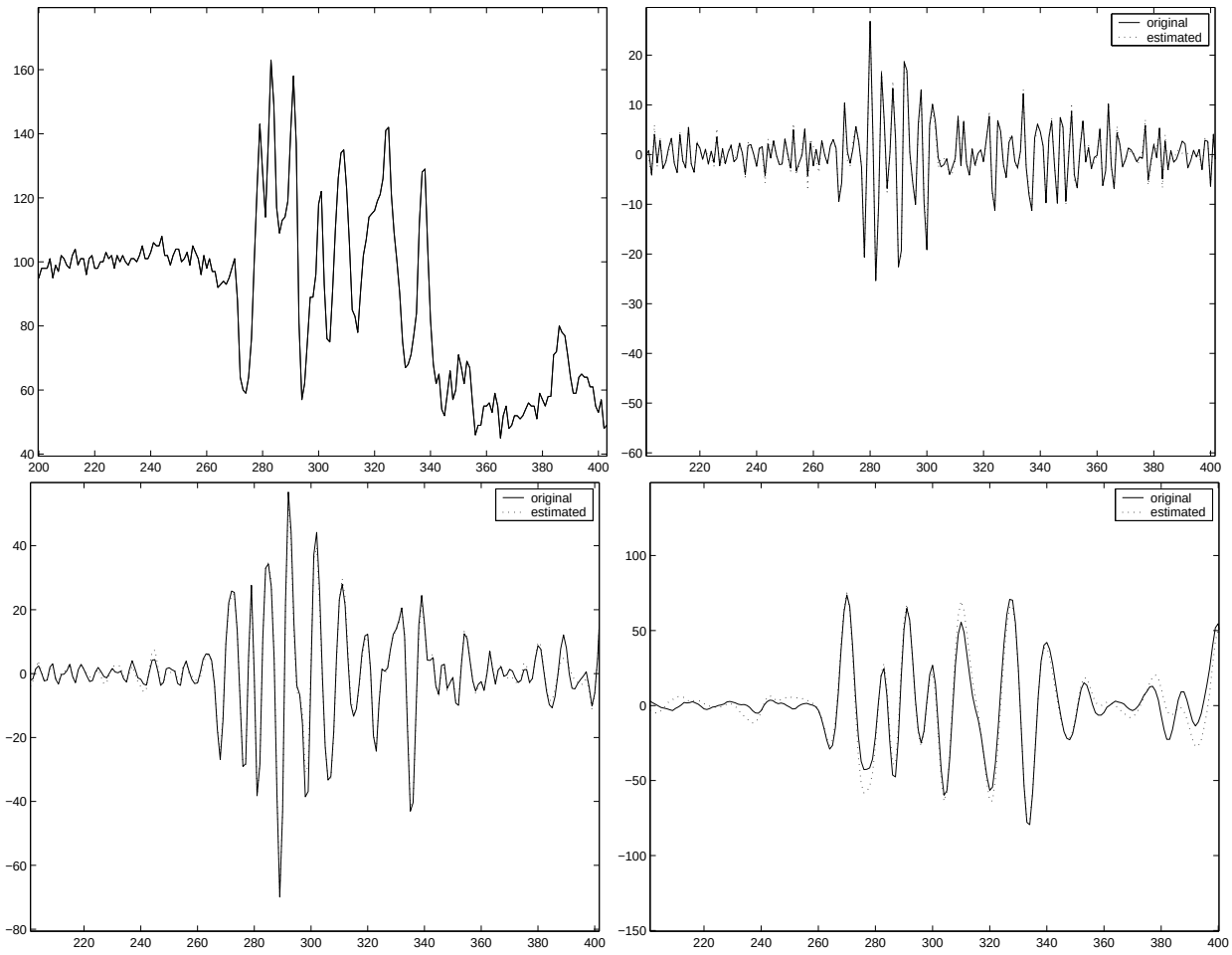

Figure 4: *Topleft)* Portion of row no. 100 of lena image. *Top-bottom left-right* Wavelet details up to third scale level reconstructed using atoms evolution law. The recovered signal gives PSNR = 45.57 db.

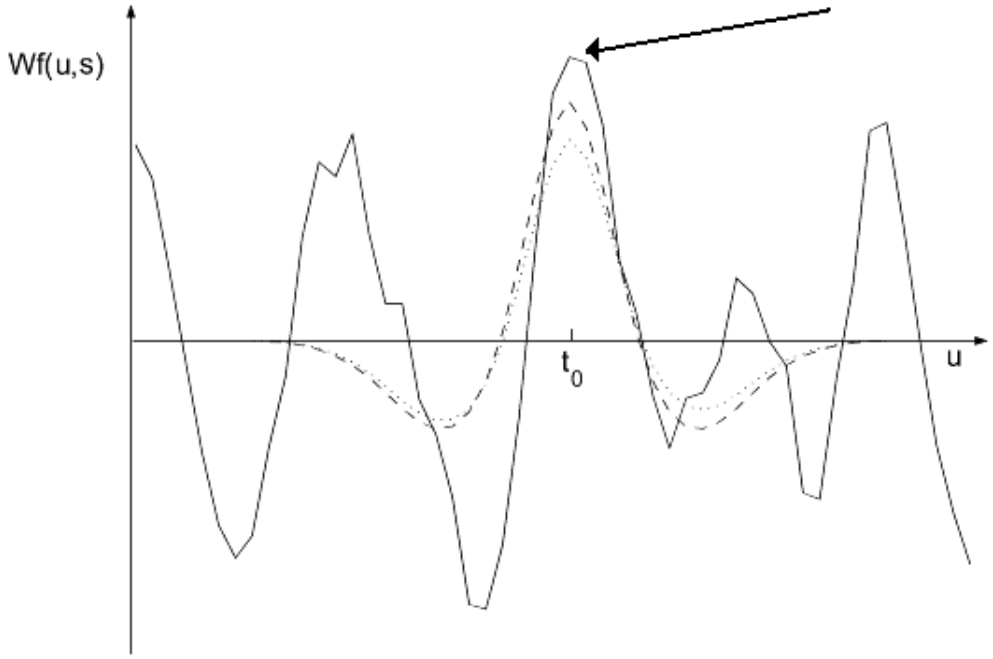

Figure 5: Wavelet transform of a noisy infinite ramp signal (solid line). The arrow indicates the correct location of atom maximum corresponding to the signal singularity. The wavelet transform of the clean signal is depicted with a dashed line while the estimated atom using the proposed model is indicated by a dotted line.

original one. Equations (9) and (10) allow us to predict the contribution of the adopted atoms in their cone of influence at different scales, even when they completely interfere. Equation (11) provides a faithful estimation of the decay exponents γ_k and allows to discard the noisy initial conditions, i.e. atoms having negative decay.

Using the preserved atoms, the location of their modulus maxima at successive scales can be estimated and the corresponding slope value can be refined. In fact, at coarser scales the noise flattens while the estimation domains become wider, since the dilation property of the wavelet transform. It turns out that the least squares used for slopes estimation are more precise. Atoms slopes are estimated at each scale using a greedy algorithm on the amplitude of modulus maxima (from the highest to the smallest), i.e.

$$\alpha_k = \frac{< R_k(u,s), F(t_k,u,s) >}{\|F(t_k,u,s)\|^2}$$

where $R_k(u,s) = Wg(u,s) - \sum_{h=1}^{k-1} R_h(u,s)$ and Wg is the wavelet transform of the noisy signal. It corresponds to a least squares estimation in a suitable domain by imposing a function model and then providing a regularization of the noisy data (see Fig.5 and [2] for details).

It is worth noticing that the atomic representation, as in eq. (5), intrinsically preserves the correlation between adjacent coefficients of the wavelet decomposition. This property also allows the recovering of coefficients under threshold and avoids artifacts due to the rough cut off of information in the selection based approaches.

4.1 The Algorithm

Let us consider an overcomplete wavelet decomposition [23] of g. The overcomplete representation is employed to avoid the distortion of atom shape caused by the decimation. Images are considered as independent 1D signals. It is just a first attempt to directly apply the results to signals in more than one dimension. Moreover, any 2D structure has been imposed but only a local modelling of the image is given.

In the following, the mono dimensional denoising algorithm is described.

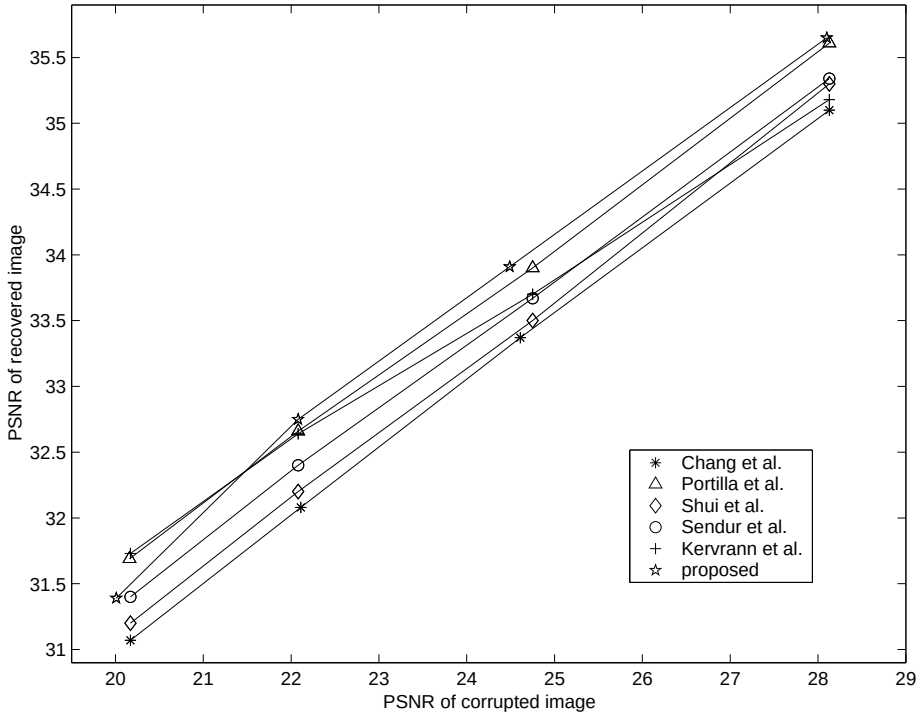

Figure 6: Comparisons in terms of PSNR between the proposed model and the models proposed in [29, 4, 33, 32, 17]. The test image is Lena and the noise standard deviations respectively are $\sigma = 10, 15, 20, 25$.

1. Perform the undecimated wavelet transform of g up to J^{th} scale level.
2. Perform the continuous wavelet transform on g at scales $s \in [1, 2]$ using the step $\Delta s = .05$.
3. Estimate the parameters $\{t_k, \alpha_k, \gamma_k\}$ using WISDOW [2] slope estimation algorithm three times for getting α_k, and t_k at scale $s = 1$ and (11) for estimating γ_k.
4. Eliminate atoms having $\gamma_k < 0$.
5. Compute atoms trajectories by solving (10) using a 4^{th} order Runge Kutta method and extract the solution at dyadic scales $s = 2^j, \quad j = 1, \ldots, J$.
6. At scale $s = 2^j, \quad j = 1, \ldots, J$, sort selected maxima in decreasing order with respect to their absolute value and estimate the corresponding slope α_k. The data to use in the least squares estimation are weighted proportionally to the ratio between the analysed maximum and the ones in its cones of influence, which have been predicted by the law (eq. (10)).
7. Invert the undecimated wavelet transform using the recovered detail bands.

5 Experimental Results

Many experiments have been done for testing the performances of the proposed model. Both 1D signals and images have been processed having different kinds of smoothness. A biorthogonal wavelet 3/9 associated to an over-complete multi-resolution decomposition computed up to 3^{rd} scale has been adopted in all tests, while the integration step h for solving the ode in the step 5 of the denoising algorithm, has been set to 0.05.

Images are split into independent 1D signals. As mentioned above, it is just a first attempt to use the evolution laws for images and a significant test for measuring the potentialities of the model. In fact, experimental results

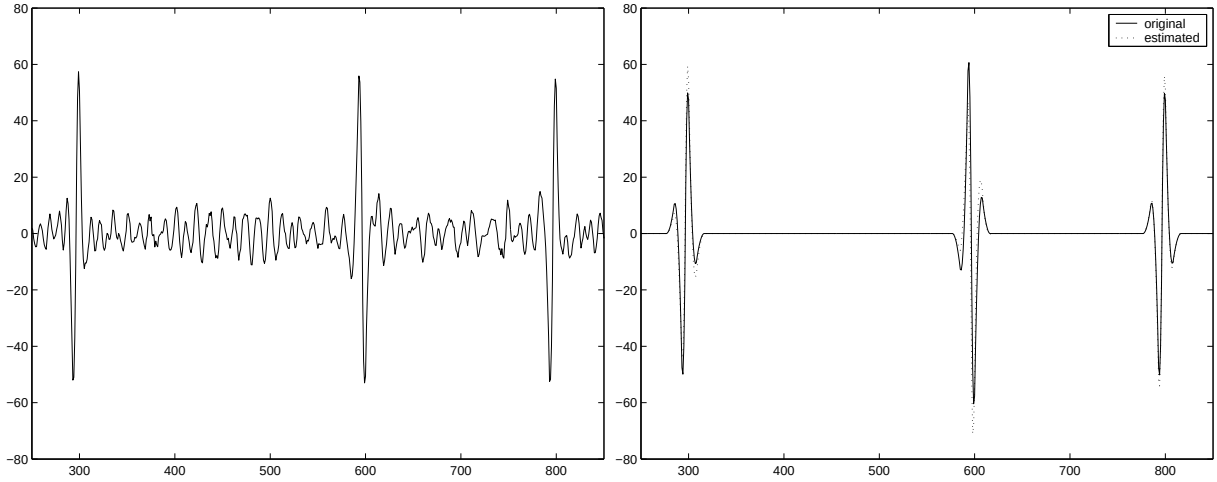

Figure 7: *Left)* Noisy wavelet coefficients ($\sigma = 5$) at 3^{rd} scale level of the piecewise polynomial signal in Fig. 3. *Right)* Recovered coefficients using the proposed model (dotted line). They are compared with the original one (solid line).

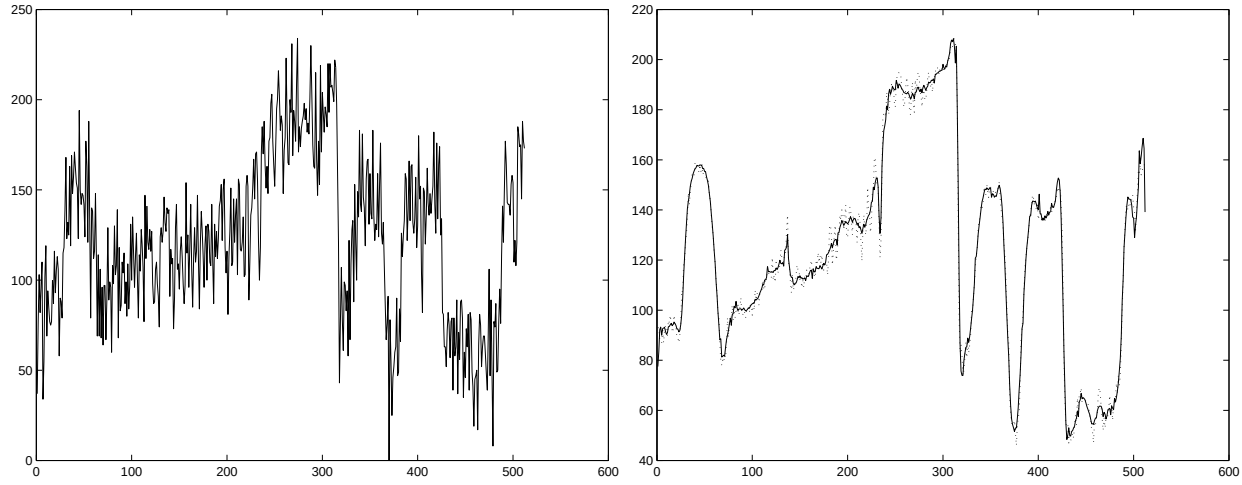

Figure 8: *Left)* Noisy row no. 100 of Lena image (PSNR = 22.06 db). *Right)* Denoised signal (solid line) using the proposed approach (PSNR = 32.75 db). The original signal is the dotted line.

Figure 9: Denoised Lena image using the proposed approach (Noisy image (PSNR = 22.06 db) Recovered PSNR = 32.75 db).

Figure 10: Denoised Barbara image using the proposed approach (Noisy image (PSNR = 20.10 db) Recovered PSNR = 28.40 db).

Figure 11: Zoom of the denoised pictures in Figs. 9 and 10.

show that the proposed denoising algorithm gives performances comparable to the most effective wavelet based denoising approaches even though just one direction has been used for image analysis.

In particular, Fig. 6 compares the proposed model with the Gaussian mixture estimator presented in [29], the adaptive Bayesian thresholding using context modelling contained in [4], the local Wiener filtering using elliptic directional windows for different subbands in [33], the bivariate shrinkage rule in [32] using the complex wavelet transform and the adaptive regularization algorithm with patch-based weights and variable window sizes in [17]. The latter work is not wavelet based. Nonetheless, it has been considered for comparisons since it simulates a multiscale approach. The $512 \times 512 \times 8bits$ Lena test image has been used for comparisons, while four levels of noise have been considered — $\sigma = 10, 15, 20, 25$. Presented results derive from the ability of the evolution law in establishing a precise link between corresponding coefficients at different scales and in modelling the interference between singularities even at coarser scales (spatial correlation). This guarantees an almost faithful reconstruction of the original signal, avoiding constraints on the minimum distance between them, as in [12], or thresholds tuning.

Fig. 7 shows the recovered wavelet details of the piecewise polynomial signal of Fig. 3, while Fig. 8 depicts the estimation for the row no. 100 of Lena image.

For a visual evaluation of the results, in Figs. 9 and 10 the denoised $512 \times 512 \times 8$ Lena and Barbara test images are depicted (PSNR = 32.75 and PSNR = 28.40) — the noise standard deviations respectively are $\sigma = 20$ (PSNR = 22.06) and $\sigma = 25$ (PSNR = 20.14). It is worth noticing that edges are well recovered with a drastic reduction of both ringing effects around edges and isolated spikes. To better appreciate the results, a zoom of the recovered images is also shown in Fig. 11.

As regards the computational effort of the algorithm, it is linear with respect to the number of atoms used for the approximation, while it requires additional computations for the solution of the ode in (10). In fact, the integration step h has to be lower than 0.1 for getting good prediction for atoms locations at dyadic scales.

6 Conclusions

In this paper a wavelet based model for image denoising has been presented. It describes the wavelet transform of a generic signal as superimposition of predefined basic atoms, whose evolution law along scale can be modelled. This law allows to build modulus maxima chains of the wavelet transform and then to faithfully estimate the contribution of each atom at each scale. We have shown an application to image denoising but the model can be successfully exploited also in image compression or image segmentation, while a possible application to

super-resolution is under investigation. The results achieved in denoising are satisfying in terms of both mean square error and visual quality — edges and image contours are well recovered without introducing ringing artifacts.

Future research will be oriented to the reduction of the computational cost along with the 2D extension of the definition of a basic atom.

References

[1] E. J. Balster, Y.F. Zheng, R.L. Ewing, "Feature-Based Wavelet Shrinkage Algorithm for image Denoising", *IEEE Transactions on Image Processing*, Vol. 14, No. 12, December 2005.

[2] V. Bruni, D. Vitulano, "Wavelet based Signal Denoising via Simple Singularities Approximation", *Signal Processing*, Elsevier Science, Vol. 86, pp. 859-876, April 2006.

[3] V. Bruni, B. Piccoli, and D. Vitulano, "Scale space atoms for signals and image de-noising," in *IAC Report*, 2006.

[4] S.G. Chang, Bin Yu and M. Vetterli, "Adaptive Wavelet Thresholding for Image Denoising and Compression", *IEEE Transactions on Image Processing*, Vol. 9, No. 9, pp. 1532-1546, September, 2000.

[5] H. Choi and R. Baraniuk, "Analysis of Wavelet - Domain Wiener Filters",*Proc. IEEE - SP International Symposium on Time-frequency and Time-scale Analysis*, October 1998.

[6] A. Cohen, I. Daubechies, O. Guleryuz. and M.T. Orchard, "On the importance of Combining Wavelet-Based Nonlinear Approximation With Coding Strategies",*IEEE Transactions on Information Theory*, Vol. 48, No. 7, pp. 1895-1921, July 2002.

[7] M. S. Crouse, R. D. Nowak and R. G. Baraniuk, "Wavelet-based Statistical Signal Processing using Hidden Markov Models", *IEEE Transactions on Signal Processing*, Vol. 46, No. 4, pp. 886-902, April 1998.

[8] M. N. Do and M. Vetterli, "Contourlets: A New Directional Multiresolution Image Representation", *Proc. 36th Asilomar Conf. on Signals Systems and Computers*, Vol.1, November 2002, pp. 497-501.

[9] D. L. Donoho and I. M. Johnstone, "Ideal Spatial Adaptation via Wavelet Shrinkage", *Biometrika*, Vol. 81, pp. 425-455, 1994.

[10] D. L. Donoho, " Denoising by soft thresholding", *IEEE Transactions on Information Theory*, Vol. 41, No. 3, pp. 613-627, May 1995.

[11] P.L. Dragotti and M. Vetterli, "Footprints and Edgeprints for Image Denoising and Compression", *Proceedings of IEEE International Conference on Image Processing (ICIP)*,Thessaloniki, Greece, October 2001.

[12] P.L. Dragotti and M. Vetterli, "Wavelet Footprints: Theory, Algorithms and Applications", *IEEE Transactions on Signal*, Vol. 51, No. 5, pp. 1306-1323, May 2003.

[13] E. Le Pennec and S. Mallat, "Non linear image approximation with bandelets", *Tech. Rep. CMAP/ Ecole Polytechnique*, 2003.

[14] G. Fan and X. Xia, "Image Denoising using a Local Contextual Hidden Markov Model in the Wavelet Domain", *IEEE Signal Processing Letters*, Vol. 8, No. 5, pp. 125-128, May 2001.

[15] G. Gilboa, N. Sochen and Y.Y. Zeevi, "Image Enhancement and Denoising by Complex Diffusion Processes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 8, pp. 1020-1036, August 2004.

[16] M. Kazubek, "Wavelet Domain Image Denoising by Thresholding and Wiener Filtering", *IEEE Signal Processing Letters*, Vol. 10, No. 11, pp. 324-326, November, 2003.

[17] C. Kervrann and J. Boulanger, "Unsupervised Patch-Based Image Regularization and Representation", *Proc. of European Conf. Comp. Vision (ECCV'06)*, Graz, Austria, May 2006.

[18] J. Koenderink, "The structure of images", *Biol. Cybern.*, Vol. 50, pp. 363-370, 1984.

[19] A. Kuijper and L.M.J. Florack, "Using Catastrophe Theory to Derive Trees from Images", *Journal of Mathematical Imaging and Vision*, Vol. 23 (3), pp. 219-238, November 2005.

[20] S. Jaffard, Y. Meyer and R. D. Ryan, *Wavelets: Tools for Science and Technology*, SIAM 2001.

[21] X. Lin and M. T. Orchard, "Spatially Adaptive Image Denoising under Overcomplete Expansion", *Proceedings IEEE International Conference on Image Processing*, pp. 300-303, September 2000.

[22] S. Mallat and W.L. Hwang, "Singularity Detection and Processing with Wavelets", *IEEE Transactions on Information Theory*, Vol. 38, No. 2, March 1992.

[23] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[24] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges", *IEEE Trans. On Pattern Analysis Machine Intelligence*, Vol. 14, pp. 710-732, 1992.

[25] M.K. Mihcak, I. Kozintsev, K. Ramchandran and P. Moulin, "Low-complexity Image Denoising based on Statistical Modeling of Wavelet Coefficients", *IEEE Signal Processing Letters*, Vol. 6, No. 12, pp. 300-303, December, 1999.

[26] P. Perona and J. Malik, "Scale-Space and Edge Detection Using Anisotropic Diffusion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, pp. 629-639, July 1990.

[27] A. Petrovic, O.D. Escoda and P. Vandergheynst, "Multiresolution Segmentation of Natural Images: From Linear to Nonlinear Scale-Space Representations", *IEEE Transactions on Image Processing*, Vol. 13, No. 8, August 2004.

[28] A. Pizurica, W. Philips, I. Lemanhieu, M. Acheroy, "A Joint Inter- and Intrascale Statistical Model for Bayesian Wavelet Based Image Denoising", *IEEE Transactions on Image Processing*, Vol. 11, No. 5, May 2002.

[29] J. Portilla, V. Strela, M. Wainwright and E. Simoncelli, " Image Denoising using Scale Mixtures of Gaussians in the Wavelet Domain", *IEEE Transactions on Image Processing*, Vol. 12, No. 11, pp. 1338-1351, November 2003.

[30] A.C. Shih, H.M. Liao and C. Lu, "A New Iterated Two-Band Diffusion Equation: Theory and Its Application", *IEEE Transactions on Image Processing*, Vol. 12, No. 4, pp. 466-476, April 2003.

[31] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis", *IEEE Transactions on Comput.*, Vol. C-20, pp. 562-569, May 1971.

[32] L. Sendur, I.W. Selesnick, "Bivariate Shrinkage with Local Variance Estimation", *IEEE Signal Processing Letters*, Vol. 9, No. 12, December 2002.

[33] P. L. Shui, ”Image Denosing Algorithm via Doubly Local Wiener Filtering with Directional Windows in Wavelet Domain”, *IEEE Signal Processing Letters*, Vol. 12, No. 10, pp. 681-684, October 2005.

[34] J. L. Starck, E. J. Candes and D. L. Donoho, ”The Curvelet Transform for Image Denoising”, *IEEE Transactions on Image Processing*, Vol. 11, No. 6, pp. 670-684, June 2002.

[35] S. Teboul, L. Blanc-Feraud, G. Aubert and M. Barlaud, ”Variational Approach for Edge-Preserving Regularization Using Coupled PDE’s”, *IEEE Transactions on Image Processing*, Vol. 7, No. 3, pp. 387-397, March, 1998.

[36] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, P.L. Dragotti, ”Directionlets: Anisotropic Multidirectional Representation with Separable Filtering”, *IEEE Transactions on Image Processing*, Vol. 15(7), pp. 1916-1933, 2006.

[37] P. Mrazek, J. Weickert and G. Steidl, ”Diffusion-Inspired Shrinkage Functions and Stability Results for Wavelet Denoising”, *International Journal of Computer Vision*, Vol. 64 (2/3), pp. 171-186, 2005.

[38] A. Witkin, ”Scale-space filtering”, *International Joint Conf. Artificial Intelligence*, Karlsruhe, West Germany, pp. 1019-1021, 1983.

[39] H. Zhang, A. Nosratinia and R.O. Wells, ”Image Denoising via Wavelet-Domain Spatially Adaptive FIR Wiener Filtering”, *Proc. IEEE International Conf. on Acoustic Speech and Signal Processing 2000*, Vol. 5, Istanbul, Turkey, pp. 2179-2182, June 2000.