Scale space atoms for signals and image de-noising

V. Bruni, B. Piccoli, D. Vitulano

Istituto per le Applicazioni del Calcolo
"M. Picone" - C.N.R.
Viale del Policlinico 137,
00161 Rome, Italy
Tel. +39-6-884709, Fax +39-6-4404306
E-mail: {bruni,piccoli,vitulano}@iac.rm.cnr.it

Abstract

In this paper the dependencies of wavelet coefficients along scales are investigated and used for removing noise from degraded signals. Wavelet coefficients are split into *overlapping atoms*, that correspond to basic singularity points in the original signal. This model reduces the redundancy of the wavelet representation in correspondence to significant structures of the signal, while preserving correlation between adjacent coefficients. Since, the atoms are defined both in time and scale, it is possible to predict their evolution through scales via a first order partial differential equation. The latter enables to compute the trajectories of atoms global maxima in scale-space. The trajectories are solutions of a non linear ordinary differential equation that allows to precisely link significant wavelet coefficients at successive scales using general kernels. This results are successfully used for de-noising corrupted signals, as shown by the experimental results.

1 Introduction

Denoising is a widely investigated topic in Image Processing, as shown by the variety of approaches that can be found in literature. It consists of recovering a signal f from its noisy version g, where

$$g(t) = f(t) + v(t),$$

and v is an additive zero-mean gaussian noise with variance σ^2.

Most of the proposed approaches try to emphasize original signal properties by exploiting the correlation between the corresponding coefficients in a suitable expansion basis. These correlations are more evident if a scale space decomposition is used, for example the wavelet transform. In fact, it allows the analysis of singularity points exploiting the lack of correlation of the gaussian noise [7, 12, 14].

In [1, 2] is has been provided a model for further compacting wavelet coefficients information. The model attempts to reduce the redundancy of the wavelet representation in correspondence to significant structures of the signal, while preserving correlation between adjacent coefficients. This is possible by representing the wavelet transform of a signal by means of *overlapping atoms*. The atoms correspond to basic singularity points in the original signal and they are dependent on the wavelet basis. They can be discriminated at each scale level independently by modelling their interference through scale.

Nonetheless, the wavelet transform reveals an intrinsic time-scale structure, as shown in Fig. 1.

In fact, there is a precise link between successive scales. This link has been empirically exploited by defining the *persistency* property of wavelet coefficients [4]. A step forward has been done in investigating the wavelet transform of signals having singularities of different order. As shown in Chapter 1, the decay of modulus maxima of the wavelet transform gives a measure of the kind of singularity within a given interval (*cone of influence*). 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 intersect. For that reason, 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 [12, 14]. This leads to some false alarms or the missing of some important information. In Fig. 3 there is an example. It depicts two interfering singularities. In this

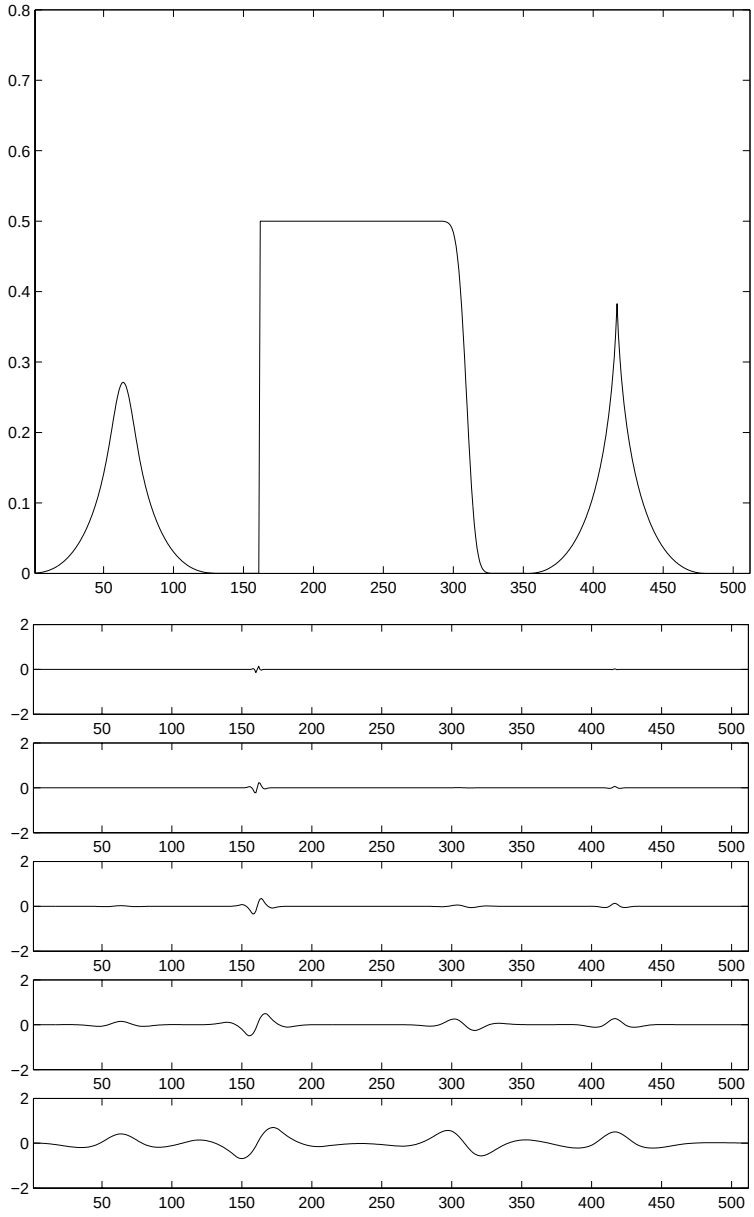


Figure 1: *Top)* Signal having different kinds of singularity. *Bottom)* Its wavelet transform at different scales.

case, since the two cones of influence intersect, only the two highest modulus maxima are wrongly selected. Moreover, once modulus maxima have been detected, we need an algorithm for reconstructing the signal from them. The maxima projection algorithm proposed in [14] is useful but it has three main drawbacks:

- it could not converge to the original signal (Mallat counterexample in [10]);
- the convergence of the algorithm requires a minimum distance between two successive maxima [14];
- maxima chains are guaranteed for each scale s only using wavelets which are derivative of a gaussian kernel (see Chapter VI of [11]).

These drawbacks can be solved by further characterize each modulus maximum. A first attempt has been done by Dragotti and Vetterli in [7], who tried to exactly model piecewise polynomial signals. Nonetheless, also in this case the distance between two adjacent singularities becomes important for distinguishing them. In other words, when two singularities interfere, *footprints* are not able to discriminate them.

The main contribution of this work consists of providing the trajectories of significant modulus maxima of the wavelet transform, $Wf(u,s)$, of a signal f in a theoretical and almost precise manner. These trajectories model the evolution law of some predefined basic atoms whose superimposition approximates $Wf(u,s)$. For each atom, the significant maximum is the one having the greatest amplitude (see Fig. 2). A significant maximum 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 (see Fig. 4). Using this representation, we can say that significant modulus maxima characterize the atomic decomposition of the wavelet transform. Notice that the latter gives only an approximation of a signal and not its perfect reconstruction, as shown in [1, 2].

The outline of the paper is the following. In Section 2 a pde for $Wf(u,s)$ is written both in the general case and for a single atom. Section 3 provides the equations of modulus maxima chains and generalizes the atom for modelling the decay of singularities of different order. Section 4 shows how these evolution laws can be exploited in denoising while some experimental results and comparative studies are presented in the last Section.

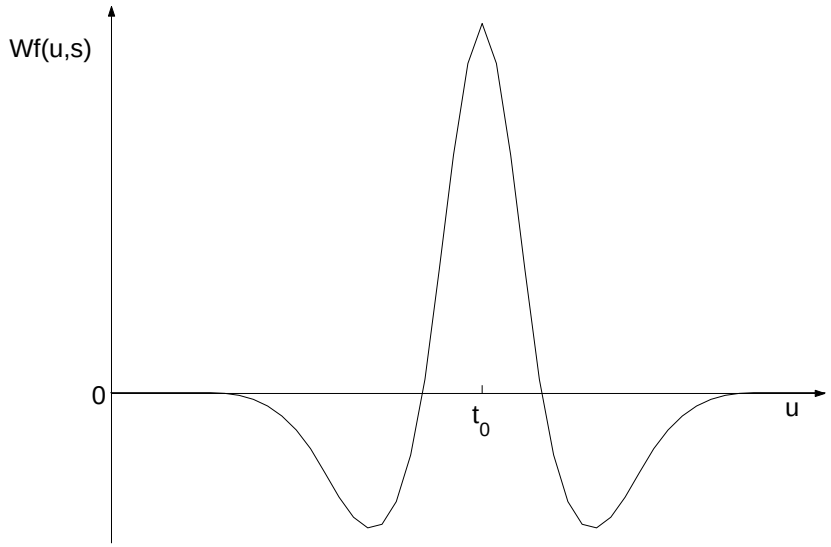


Figure 2: Basic atom computed at a fixed scale s and using the spline biorthogonal wavelet 3/9 [1, 2].

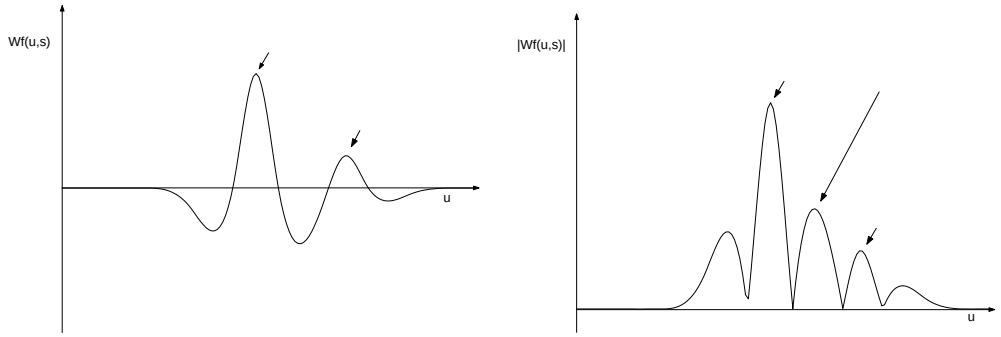


Figure 3: Two interfering atoms make a spurious maximum inside their cones of influence: a) wavelet transform composed of two atoms having the same sign, whose maxima are indicated by arrows; b) modulus of the wavelet transform: the longer arrow indicates the spurious modulus maximum that should be selected in place of the real rightmost one, following the classical procedure [12] .

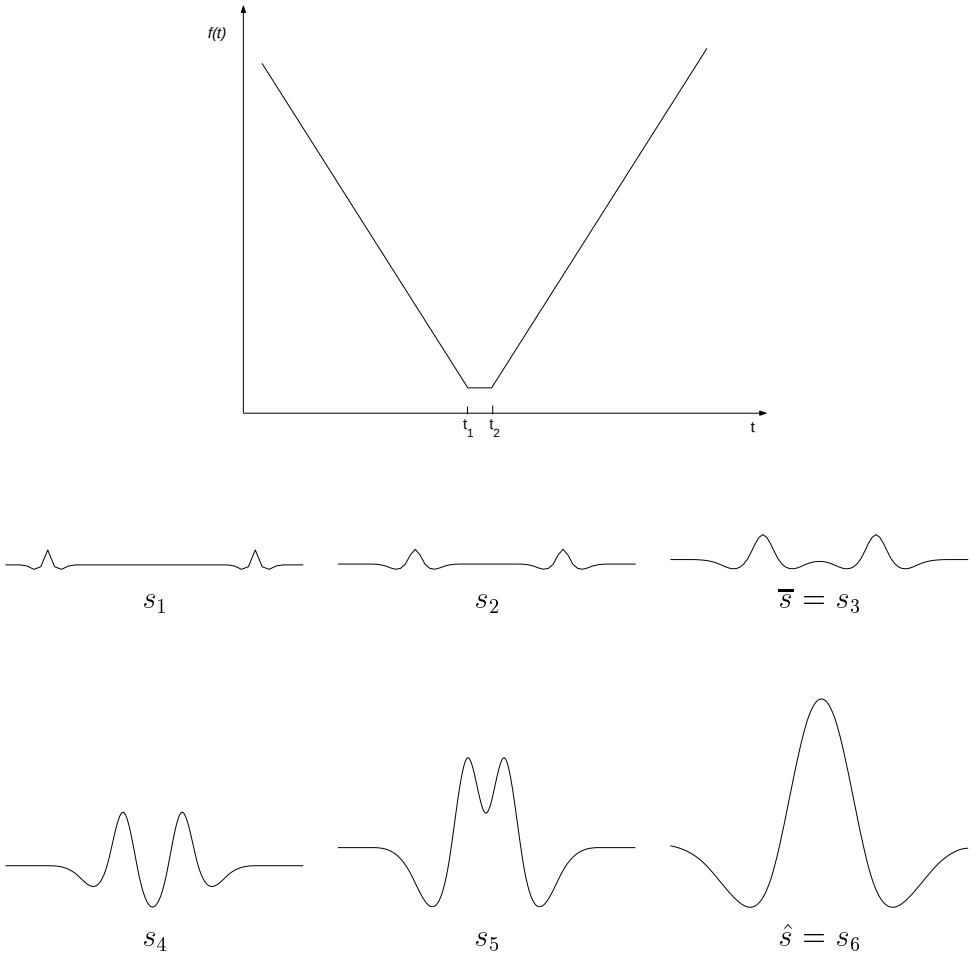


Figure 4: Topmost figure depicts a signal having two singularity points at t_1 and t_2. Below there is its corresponding wavelet transform computed at successive scales $s = \{s_1, \ldots, s_6\}$. It is composed of two atoms having the same sign which interfere as long as the scale increases.

2 Scale dependencies in the atomic representation

Let ψ be a real and continuous wavelet and let us suppose that its first derivative is defined. The function $\overline{\psi}(u,s)$ is obtained by scaling and dilating ψ by a factor s, i.e.

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

where $s \in \mathbf{R}^+$ and $u \in \mathbf{R}$. Therefore, its partial derivatives with respect to the variables s and u are:

$$\overline{\psi}_s = -\frac{1}{2s\sqrt{s}}\psi\left(-\frac{u}{s}\right) + \frac{u}{s^2\sqrt{s}}\psi'\left(-\frac{u}{s}\right)$$

$$\overline{\psi}_u = -\frac{1}{s\sqrt{s}}\psi'\left(-\frac{u}{s}\right).$$

Hence, by comparing ψ_s and ψ_u, we have

$$\overline{\psi}_s = \left(-\frac{u}{s}\overline{\psi}\right)_u + \frac{1}{2s}\overline{\psi}. \tag{2}$$

It is a first order, non homogeneous and semi-linear partial differential equation and represents the evolution law through scales for $\overline{\psi}$ [8]. Let us remind that the wavelet transform of a generic function f by using the mother wavelet ψ is

$$Wf(u,s) = f * \overline{\psi} = \frac{1}{\sqrt{s}}\int_{-\infty}^{+\infty} f(t)\psi\left(\frac{t-u}{s}\right)dt.$$

To simplify the notation, in this section $Wf(u,s)$ will be simply indicated with $w(u,s)$.

Therefore, by convolving both members of (2) with a generic function f, we achieve:

$$f * \overline{\psi}_s = f * \left(-\frac{u}{s}\overline{\psi}\right)_u + \frac{1}{2s}f * \overline{\psi}$$

and by using the property of the convolution with respect to the derivative[1] it follows:

$$w_s = \left(f * (-\frac{u}{s}\overline{\psi})\right)_u + \frac{1}{2s}w. \tag{3}$$

[1] $f * g(t) = g * f(t)$; $\frac{d}{dt}(f * g)(t) = \frac{df}{dt} * g(t) = f * \frac{dg}{dt}(t)$

Using simple computation eq. (3) can be rewritten as follows

$$w_s = -\frac{u}{s} w_u - \frac{1}{2s} w + \frac{1}{s} v_u, \tag{4}$$

where v is the wavelet transform of the function $tf(t)$. This equation shows two different effects. The first term after equal, i.e. $-\frac{u}{s} w_u$, guides a sort of transport along the scale s. The second and third terms, i.e. $-\frac{1}{2s} w + \frac{1}{s} v_u$, are source terms and guide the decay and the shape of the wavelet transform along scales.

It is worth outlining that equation (3) is a more general result than the one in [11], since we did not make any assumption about the function f or the kernel ψ. In fact, as stated in [11], if ψ is the n^{th} order derivative of a gaussian kernel $\theta(t)$, i.e. $\psi(t) = (-1)^n \theta^{(n)}(t)$, then the function $\rho = \frac{1}{\sqrt{s}} f^{(n)} * \overline{\theta}$ satisfies the heat equation:

$$\rho_s = \frac{1}{2s} \rho_{uu}. \tag{5}$$

We will now prove that assuming $\psi(t) = (-1)^n \theta^{(n)}(t)$, (5) derives from (3).

Proposition 1 *Let $\psi(t) = (-1)^n \theta^{(n)}(t)$, where θ is a gaussian kernel, then (3) implies (5).*

Proof:
From (1) we have $\overline{\psi}(t,s) = s^n \frac{\partial^n}{\partial t^n} \overline{\theta}(t,s)$ and then $w(u,s) = s^n f^{(n)} * \overline{\theta}(u)$. It is trivial to verify that if $\theta(t) = e^{-t^2}$, then

$$\theta^{(n+1)}(t) = -2t\theta^{(n)}(t) - 2n\theta^{(n-1)}(t), \quad \forall\, n \in \mathbf{N},\ n \neq 0. \tag{6}$$

Hence, for $\psi = (-1)^n \theta^{(n)}(t)$, the equation (3) becomes

$$\left(s^n f^{(n)} * \overline{\theta}(u)\right)_s = -\left(f * \frac{u}{s} s^n \overline{\theta}^{(n)}(u)\right)_u + \frac{1}{2s}(s^n f^{(n)} * \overline{\theta}(u)) \tag{7}$$

By computing the derivative with respect to s in the first member of (7) and using (6) in the first term of the second member of (7), we have

$$s^n \left(f^{(n)} * \overline{\theta}(u)\right)_s + n s^{n-1}[f^{(n)} * \overline{\theta}(u)] = \frac{s^{n-1}}{2} \left(f * \overline{\theta}^{(n+1)}(u))\right)_u +$$

$$+\frac{s^{n-1}}{2}(f * 2n\overline{\theta}^{(n-1)}(u))_u + \frac{1}{2s}(s^n f^{(n)} * \overline{\theta}(u)).$$

Using the property of convolution, it follows

$$s^n \left(f^{(n)} * \overline{\theta}(u)\right)_s + ns^{n-1}[f^{(n)} * \overline{\theta}(u)] = \frac{s^{n-1}}{2}\left(f * \overline{\theta}^{(n)}(u))\right)_{uu} +$$

$$+\frac{s^{n-1}}{2}2n(f * \overline{\theta}^{(n)}(u)) + \frac{1}{2s}(s^n f^{(n)} * \overline{\theta}(u)).$$

Let us now divide both members by $s^{n-1}\sqrt{s}$, hence

$$\frac{s}{\sqrt{s}}\left(f^{(n)} * \overline{\theta}(u)\right)_s = \frac{1}{2\sqrt{s}}\left(f^{(n)} * \overline{\theta}(u)\right)_{uu} + \frac{s}{2s\sqrt{s}}(f^{(n)} * \overline{\theta}(u)). \qquad (8)$$

Since $\left(\frac{1}{\sqrt{s}}f^{(n)} * \overline{\theta}(u)\right)_s = -\frac{1}{2s\sqrt{s}}(f^{(n)} * \overline{\theta}(u)) + \frac{1}{\sqrt{s}}\left(f^{(n)} * \overline{\theta}(u)\right)_s$,
eq. (8) becomes

$$s\left(\frac{1}{\sqrt{s}}f^{(n)} * \overline{\theta}(u)\right)_s + \frac{s}{2s\sqrt{s}}(f^{(n)} * \overline{\theta}(u)) = \frac{1}{2\sqrt{s}}\left(f^{(n)} * \overline{\theta}(u)\right)_{uu} + \frac{s}{2s\sqrt{s}}(f^{(n)} * \overline{\theta}(u))$$

and then

$$\left(\frac{1}{\sqrt{s}}f^{(n)} * \overline{\theta}(u)\right)_s = \frac{1}{2s}\left(\frac{1}{\sqrt{s}}f^{(n)} * \overline{\theta}\right)_{uu},$$

that is

$$\rho_s = \frac{1}{2s}\rho_{uu}.$$

$\oslash$

Although the generality of the result, the equation (3) is still unuseful in this form since we have not a priori information about the function f. In the following section, it will be shown how (3) changes by using the atomic representation proposed in [1, 2].

3 Evolution laws for wavelet atoms

Let us now to simplify the problem by using the atomic decomposition for the wavelet transform of f. We start with the case of an infinite ramp function and compute the source term v_u in (4).

Let $f(t)$ be

$$f(t) = \begin{cases} \beta & t < t_0 \\ \alpha_1(t - t_0) + \beta & t \geq t_0 \end{cases} \tag{9}$$

and $k(t) = tf(t)$, i.e.

$$k(t) = \begin{cases} \beta t & t < t_0 \\ \alpha_1(t^2 - t_0 t) + \beta t & t \geq t_0. \end{cases} \tag{10}$$

Then,

$$w(u, s) = \alpha_1 F(t_0, u, s),$$

where

$$F(t_0, u, s) = s\sqrt{s}\left(\int_{\left(\frac{t_0-u}{s}\right)}^{b} t\psi(t)dt - \left(\frac{t_0 - u}{s}\right)\int_{\left(\frac{t_0-u}{s}\right)}^{b}\psi(t)dt\right) \tag{11}$$

is the basic atom centered at t_0, and

$$v(u, s) = \frac{1}{\sqrt{s}}\int_{u+sa}^{u+sb} tf(t)\psi^*\left(\frac{t-u}{s}\right)dt =$$

$$= \frac{1}{\sqrt{s}}\int_{u+sa}^{t_0} \beta t\psi\left(\frac{t-u}{s}\right)dt + \frac{1}{\sqrt{s}}\int_{t_0}^{u+sb}(\alpha_1(t^2 - t_0 t) + \beta t)\psi\left(\frac{t-u}{s}\right)dt =$$

$$= \beta s\sqrt{s}\int_a^b y\psi(y)dy - \alpha_1 u\sqrt{s}(t_0 - u)\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy +$$

$$+\alpha_1 s^2\sqrt{s}\int_{\frac{t_0-u}{s}}^{b} y^2\psi(y)dy - \alpha_1 s\sqrt{s}(t_0 - 2u)\int_{\frac{t_0-u}{s}}^{b} y\psi(y)dy,$$

where the change of variable $y = \frac{t-u}{s}$ and the zero mean property of ψ have been used.

Hence, the partial derivative of $v(u, s)$ with respect to u is

$$v_u = -\alpha_1\sqrt{s}(t_0-2u)\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy - \alpha_1\sqrt{s}(t_0u-u^2)\frac{1}{s}\psi\left(\frac{t_0-u}{s}\right) +$$

$$+\alpha_1 s^2\sqrt{s}\frac{1}{s}\left(\frac{t_0-u}{s}\right)^2\psi\left(\frac{t_0-u}{s}\right) + 2\alpha_1 s\sqrt{s}\int_{\frac{t_0-u}{s}}^{b} y\psi(y)dy +$$

$$-\alpha_1 s\sqrt{s}(t_0-2u)\frac{1}{s}\left(\frac{t_0-u}{s}\right)\psi\left(\frac{t_0-u}{s}\right).$$

By summing equal terms we get

$$v_u = -\alpha_1\sqrt{s}t_0\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy + \alpha_1\sqrt{s}2u\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy + 2\alpha_1 s\sqrt{s}\int_{\frac{t_0-u}{s}}^{b} y\psi(y)dy.$$

By adding and subtracting the quantity $\alpha_1\sqrt{s}t_0\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy$ and comparing the result with (11), it holds

$$v_u = 2\alpha_1 s\sqrt{s}\left\{\int_{\frac{t_0-u}{s}}^{b} y\psi(y)dy - \left(\frac{t_0-u}{s}\right)\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy\right\} + \alpha_1 t_0\sqrt{s}\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy =$$

$$= 2\alpha_1 F(t_0,u,s) + \alpha_1 t_0\sqrt{s}\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy = 2w + \alpha_1 t_0\sqrt{s}\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy.$$

Since $w_u = \alpha_1\sqrt{s}\int_{\frac{t_0-u}{s}}^{b}\psi(y)dy$, then

$$v_u = 2w + t_0 w_u. \tag{12}$$

Putting (12) in (4), we achieve the evolution law for the isolated atom $\alpha_1 F(t_0,u,s)$:

$$w_s = \frac{t_0-u}{s}w_u + \frac{3}{2s}w. \tag{13}$$

This equation is more tractable than the general one (4). In fact, it is semi-linear and the inhomogeneous term is affine with respect to w [8]. Hence, we can study its characteristic curves.

We want to find the projection of the solution of (13) onto the parametric curves

$$\begin{cases} s = s(\tau) \\ u = u(\tau) \\ w = w(s(\tau),u(\tau)) \end{cases} \tag{14}$$

where $\tau \in \mathbf{R}$. Since $\frac{d}{d\tau}w = w_s \frac{d}{d\tau}s + w_u \frac{d}{d\tau}u$, from (13) we derive $w(s(\tau), u(\tau))$ as the solution of the following ordinary differential equation:

$$\frac{d}{d\tau}w = \frac{3}{2s}w \tag{15}$$

if and only if

$$\begin{cases} \frac{d}{d\tau}s = 1 \\ \frac{d}{d\tau}u = -\frac{t_0 - u}{s} \end{cases}$$

By solving both equations using the separation of variables, it follows:

$$\begin{cases} s = \tau \\ \log(t_0 - u) = \log(C_0 \tau). \end{cases}$$

Hence

$$u(s) = t_0 - C_0 s,$$

where C_0 is the initial condition at scale $s = 1$, i.e $C_0 = t_0 - u(1)$.

Summing up, the characteristic curves of (13) are straight lines, as shown in Fig. 5, whose equations are:

$$u(s) = t_0 - (t_0 - u(1))s. \tag{16}$$

They are continuous curves and provide the trajectory of all the wavelet coefficients composing the atom in the scale-space domain.

It means that a single atom travels through scales without moving from its original position. The two only effects that modify its shape are tied to the terms on the right side of (13). The first one is *transport*: the atom becomes wider along scales. The second one is a *source* term. The atom amplitude is higher along scales ($O(s^{3/2})$) in agreement with the theorem of Jaffard (Theorem 6.4 in [11]). These two effects are shown in Fig. 6.

In Section 2.4, we stated that each atom can be characterized by its modulus maxima. In particular, we saw that each atom is centered at location t_k. Choosing a symmetric wavelet, t_k corresponds to the atom center of mass that for the shape in Fig. 2, coincides with the global maximum location, i.e. its mid point. Whenever singularities can be detected by means of modulus maxima of their corresponding atom, it is possible to provide the trajectories of those maxima along scales. This is a crucial result since it gives a precise rule for constructing significant maxima chains in the wavelet domain.

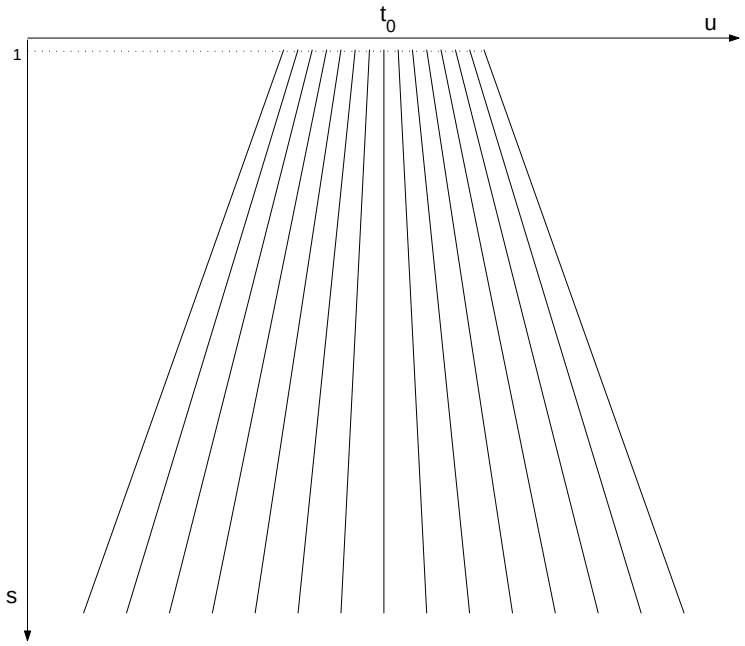


Figure 5: Characteristic curves of (13).

Let us compute the derivative with respect to u of both members of eq. (13) at critical points, i.e. $w_u(s, u(s)) = 0$. Hence

$$w_{su} = \left(\frac{t_0 - u}{s}\right) w_{uu}. \tag{17}$$

On the other hand, along atom extrema chains $\overline{u} = u(s)$ we have:

$$\left((w(s, u(s))_u\right)_s = 0$$

i.e.,

$$w_{su} + w_{uu}\dot{u} = 0, \tag{18}$$

where with $\dot{u}$ we indicate the derivative of $u(s)$ with respect to s. By comparing (17) and (18) we have:

$$-w_{uu}\dot{u} = \frac{t_0 - u}{s} w_{uu}.$$

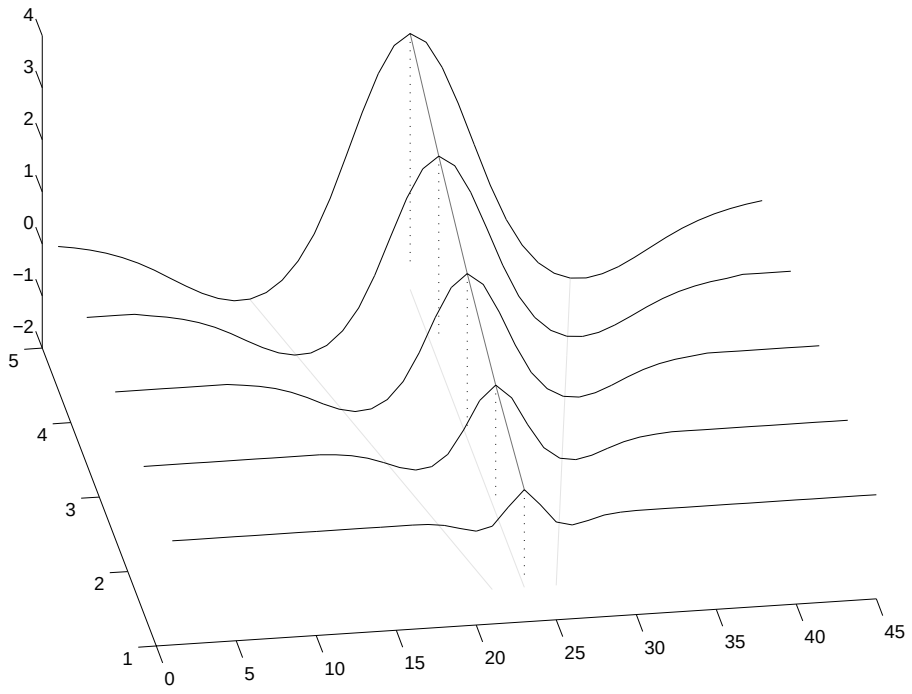


Figure 6: Black solid lines represent an isolated wavelet atom at increasing scales. Its behavior accounts for both the diffusive and the source effect. The amplitudes of global absolute maxima along scales belong to the curve $s^{3/2}$ (dark gray solid line). Light gray lines depict atoms absolute maxima trajectories.

Since $w_{uu}(s, u(s)) \neq 0$, it follows

$$\dot{u} = -\frac{t_0 - u}{s}. \tag{19}$$

Hence, (19) gives the trajectories of all atom extrema of (11). Since we are interested in following the global modulus maximum, we have to consider the following initial condition $u(1) = t_0$. Therefore the trajectory of the global absolute maximum of (11) is

$$u(s) = t_0, \quad \forall s \ \in \mathbf{R}^+.$$

It follows that the global maximum of (11) does not move from its initial location.

3.1 Trajectories for interfering atoms

Let us now try to generalize to a function $f(t)$ having two singularity points, like the one depicted in Fig. 4. In this case, (13) becomes:

$$w_s = \frac{\overline{t} - u}{s} w_u + \frac{3}{2s} w + \frac{d}{2\sqrt{s}} \left\{ \alpha_2 \int_{z_2}^{+\infty} \psi(y) dy - \alpha_1 \int_{-\infty}^{z_1} \psi(y) dy \right\} \tag{20}$$

with $\overline{t} = \frac{t_1 + t_2}{2}$, $\{z_k = \frac{t_k - u}{s}, \quad k = 1, 2\}$ and $d = t_2 - t_1$.

The corresponding maxima trajectories are:

$$\dot{u} = -\frac{\overline{t} - u}{s} - \frac{d}{2s} \left(\frac{\alpha_2 \psi(z_2) + \alpha_1 \psi(z_1)}{\alpha_2 \psi(z_2) - \alpha_1 \psi(z_1)} \right). \tag{21}$$

It can be seen that (21) differs from (19) in its last term. This latter makes (21) non linear and describes how maxima generated by the two atoms move along scales. While there is no interference, i.e. the two cones of influence do not overlap, the behavior of their trajectories is the one depicted in Fig. 7 for $s < \overline{s}$. The two atoms are independent of each other. They are only constrained to the transport and source effects, but they do not move from their initial position. This fact can be easily seen if we put $z_2 \notin supp\{\psi(\frac{t_1 - u}{s})\}$. $\psi(z_2) = 0$ and then (21) trivially coincides with (19).

Once the two cones of influence begin to overlap ($s \geq \overline{s}$ in Fig. 7), each atom begins to be influenced by the other one. In particular, the atom with the greater amplitude has a stronger attractive force as it happens for

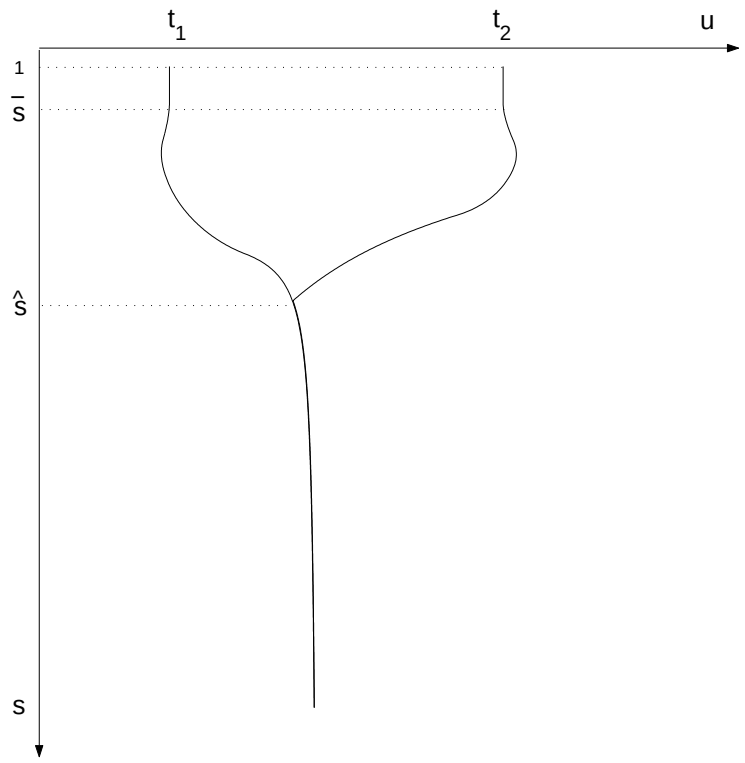


Figure 7: Modulus maxima trajectories of two interfering atoms respectively centered at t_1 an t_2. They have the same sign but different weights α_1 and α_2. In particular $|\alpha_1| > |\alpha_2|$.

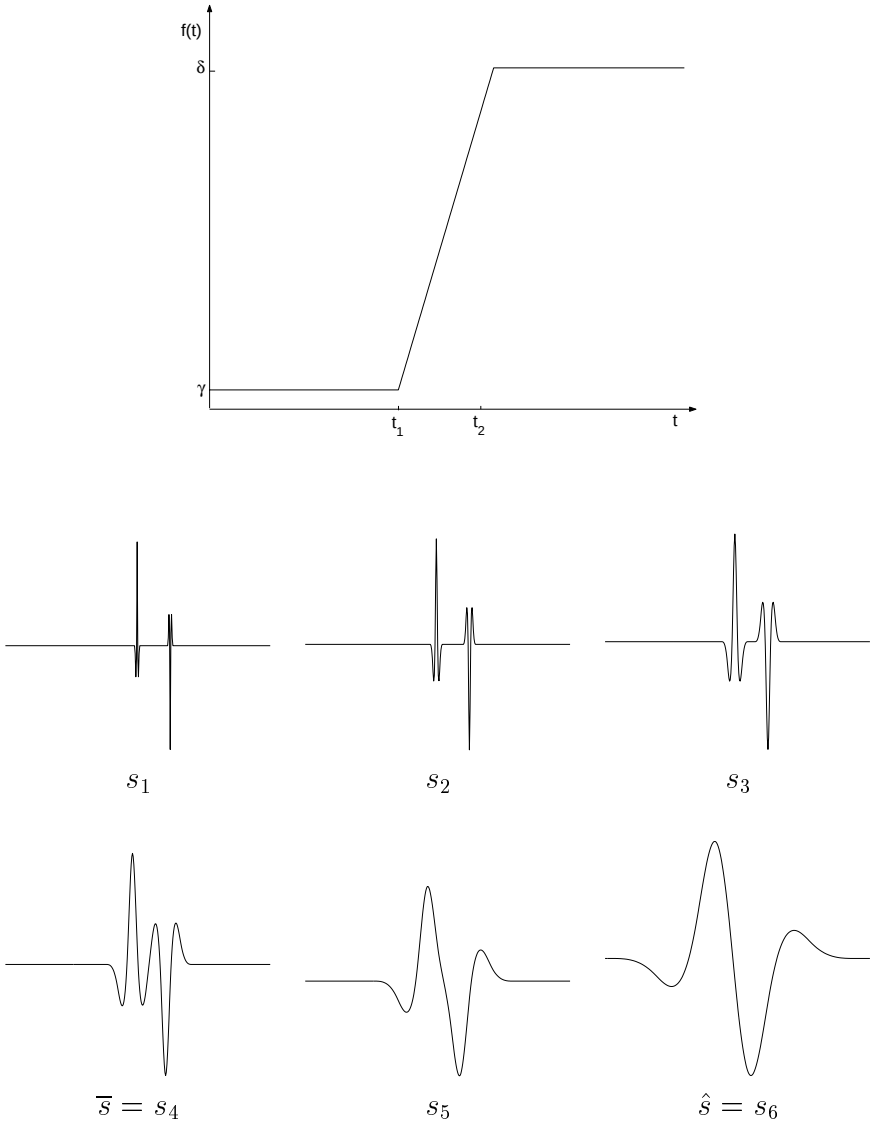


Figure 8: Topmost figure depicts a signal having two singularity points at t_1 and t_2. Below there is its corresponding wavelet transform computed at successive scales $s = \{s_1, \ldots, s_6\}$. It is composed of two atoms having opposite sign which interfere as long as the scale increases.

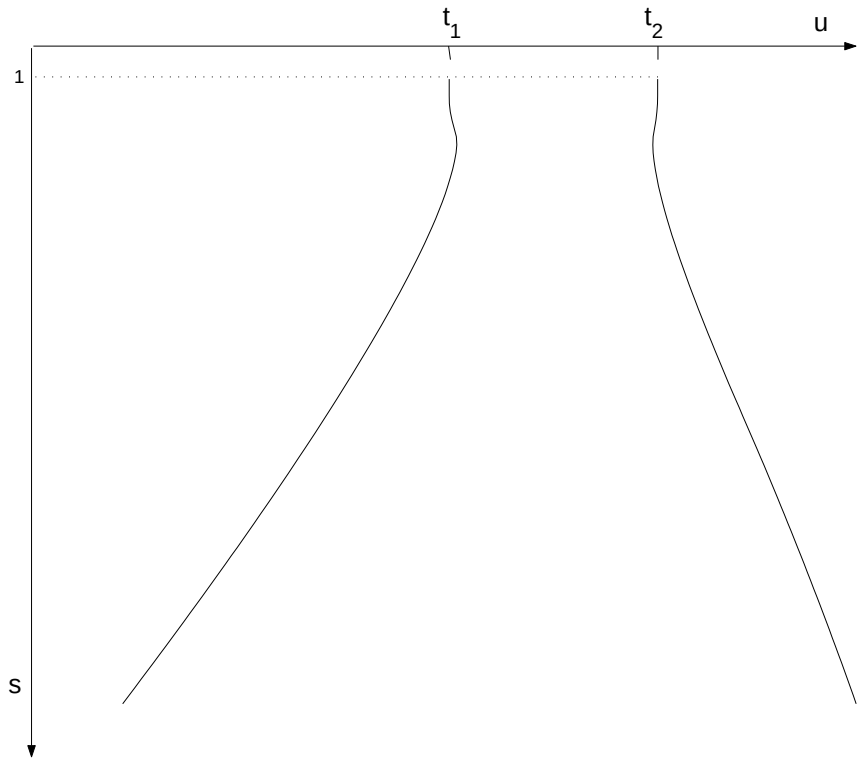


Figure 9: Modulus maxima trajectories of two interfering atoms respectively centered at t_1 an t_2. They have opposite sign and weights $|\alpha_2| > |\alpha_1|$.

gravitation. The effect is shown in Fig. 7, where it can be seen that the smaller atom has a greater deviation in its trajectory.

As stated in [1, 2], when the two atoms interfere there are two possibilities:

- The two atoms have the same sign (see Figs. 4 and 7). Maxima trajectories apparently show first repulsion and then attraction. The repulsion effect is due to dilation of the shape caused by the diffusive effect along with the fact that the atom is not a positive function over its whole support. When the interference is complete, there is an only one global maximum which keeps on travel as an isolated one till it meets other atoms.

- The two atoms have opposite signs (see Figs. 8 and 9). The final shape is depicted in Fig. 8. Maxima trajectories now apparently show first attraction and then repulsion. Also in this case, this effect is due to dilation of atoms caused by the diffusion and the non constant sign of the atom in the corresponding domain. When interference is complete, each atom independently travels along its characteristic curves till future meetings.

The aforementioned examples show that the atoms evolution law is not reversible. In fact, after the complete interference ($s \geq \hat{s}$ in Figs. 7 and 9), it is not possible to reconstruct atoms' history without additional information. In other words, it is not possible to predict the state at scale $\hat{s} - \epsilon$ from the state at $\hat{s}$. If we have an atom at a given scale s, we can never say if it corresponds to an isolated singularity or it is the composition of two or more atoms.

Eq. (20) can be easily generalized to the case of a function with N singularities, i.e. $w(u,s) = \sum_{k=1}^{N} \alpha_k F(t_k, u, s)$. Hence,

$$w_s = \frac{\overline{t} - u}{s} w_u + \frac{3}{2s} w + \frac{1}{Ns} \sum_{k=1}^{N} \left[w_u^{(k)} \left(\sum_{j=1}^{N} d_{kj} \right) \right], \tag{22}$$

with $\overline{t} = \frac{\sum_{k=1}^{N} t_k}{N}$, $d_{kj} = t_k - t_j$ and $w_u^{(k)}$ indicates the derivative with respect to u of $k-$th atom, i.e. $w_u^{(k)} = -\sqrt{s}\alpha_k \mathrm{sign}(\alpha_k)\mathrm{sign}(w^{(k)}(t_k)) \int_{-\infty}^{z_k} \psi(y)dy$.

It is trivial to show that (21) becomes:

$$\dot{u} = -\frac{\overline{t} - u}{s} - \frac{1}{Ns} \frac{\sum_{k=1}^{N} w_{uu}^{(k)} \left(\sum_{j=1}^{N} d_{kj} \right)}{w_{uu}}, \tag{23}$$

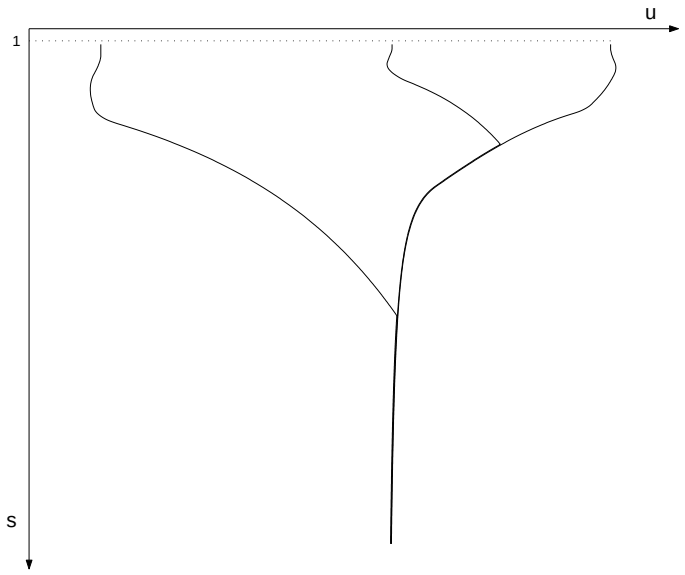


Figure 10: Modulus maxima trajectories of a signal having three singularity points. Its wavelet transform is composed of three atoms having the same sign located at $t_1 = 60$, $t_2 = 64$, $t_3 = 67$ with slopes $\alpha_1 = -8$, $\alpha_2 = -8$, $\alpha_3 = -10$.

where $w_{uu}^{(k)} = -\frac{1}{\sqrt{s}}\alpha_k \text{sign}(\alpha_k)\text{sign}(w^{(k)}(t_k))\psi(z_k)$ and $w_{uu} = \sum_{k=1}^{N} w_{uu}^{(k)}$.

This gives the equation of maxima chains along scales for the following initial conditions:

$$u_k(1) = t_k, \quad k = 1, \ldots, N,$$

where $u_k(s)$ is the trajectory of the k^{th} atom (see two examples in Figs. 10 and 11).

We can now come back again on the example of Fig. 3. Keeping in mind the atoms trajectory along scales, the spurious maximum will be never selected. In other words, the evolution law allows us to precisely follow the wavelet information, within the error produced by the atomic approximation.

3.2 About Lipschitz order of singularities

So far, an only atom shape has been used for representing every kind of singularity. It is symmetric and corresponds to a piecewise linear function. Atom maximum trajectory is exactly described by (19), while its amplitude

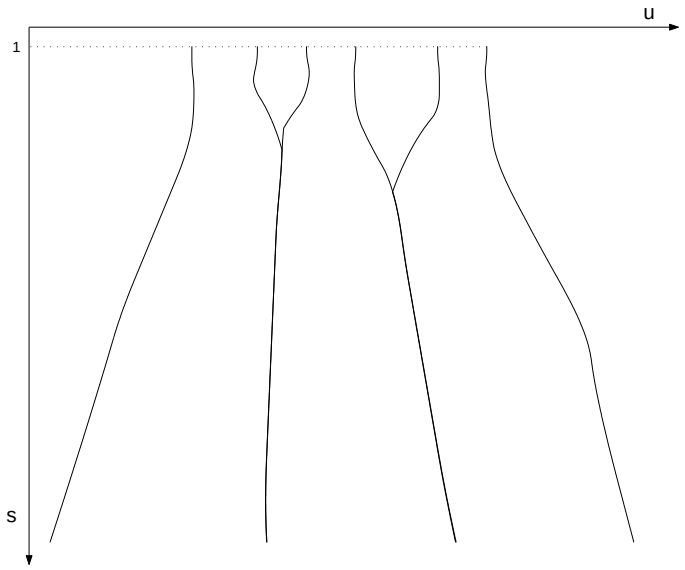


Figure 11: Modulus maxima trajectories of a signal having six singularity points respectively located at $t_1 = 60$, $t_2 = 64$, $t_3 = 67$, $t_4 = 70$, $t_5 = 75$ and $t_6 = 78$ with slopes $\alpha_1 = -8$, $\alpha_2 = 8$, $\alpha_3 = 10$, $\alpha_4 = -10$, $\alpha_5 = -6$ and $\alpha_6 = 6$.

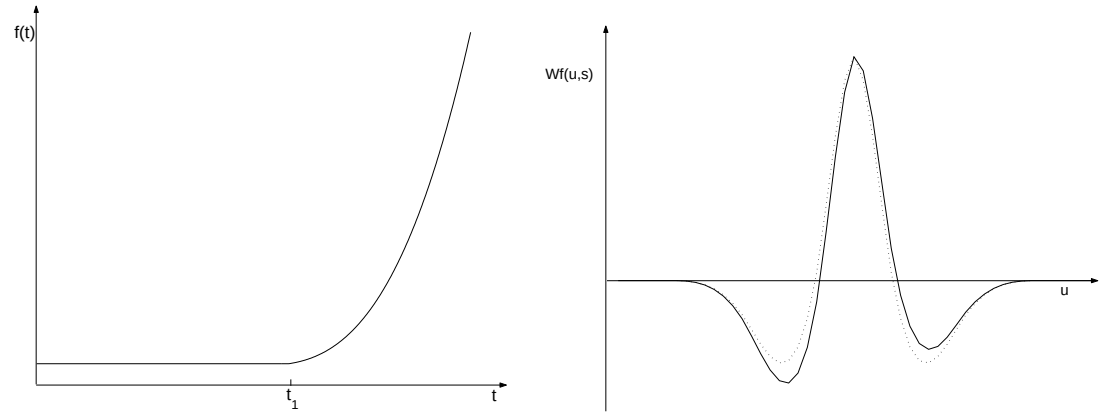
decay along scales is $O(s^{3/2})$ (see Chapter 6 in [11]). If we now consider a polynomial infinite ramp of order γ as in Fig. 12, we can see that it generates an asymmetric atom. When it is isolated, i.e. there is no interference with other atoms, its trajectory is described by (19). However, when it meets another atom we have to expect that its trajectory deviates from that in (21) since the asymmetry. Furthermore, if

$$f(t) = \begin{cases} \beta & t > t_0 \\ \alpha_1 (t - t_0)^\gamma + \beta & t \geq t_0 \end{cases}$$

the evolution law becomes:

$$w_s = \frac{t-u}{s} w_u + \frac{3}{2s} w + \frac{2(\gamma - 1)}{2s} w.$$

It has an additional term which depends on the exponent γ, which characterizes the singularity. It is obvious that this term also regulates the maxima trajectories in case of interfering atoms, as shown in Fig. 13. It is well-known that the amplitude of a wavelet modulus maximum is tied to the Lipschitz



a) b)

Figure 12: a) Third order polynomial ramp; b) the corresponding wavelet transform (solid line) and its approximation via a symmetric shape (dotted line).

order of the corresponding singularity [11]. The amplitude decay of its atom maximum is $O(s^{\gamma+1/2})$. This drawback entails the introduction of a larger vocabulary of atoms shapes, leading to the use of a *matching pursuit* [13] algorithm for selecting the best atom.

The problem can be overcome by modulating the decay along scales of an only symmetric atom. More precisely, we define a general atom $G(u,s)$ as follows:

$$G(u,s) = s^{\gamma-1}F(t_0,u,s), \tag{24}$$

where $F(t_0,u,s)$ is defined in (11). This way, the atom amplitude is modulated in agreement with the decay of the analyzed singularity. Its evolution law will be:

$$G_s = \frac{t_0-u}{s}G_u + \frac{\gamma+1/2}{s}G, \tag{25}$$

while its maximum trajectory is

$$\dot{u} = -\frac{t_0-u}{s}.$$

If the initial condition is $u(1) = t_0$, then we have the same result for the single symmetric atom.

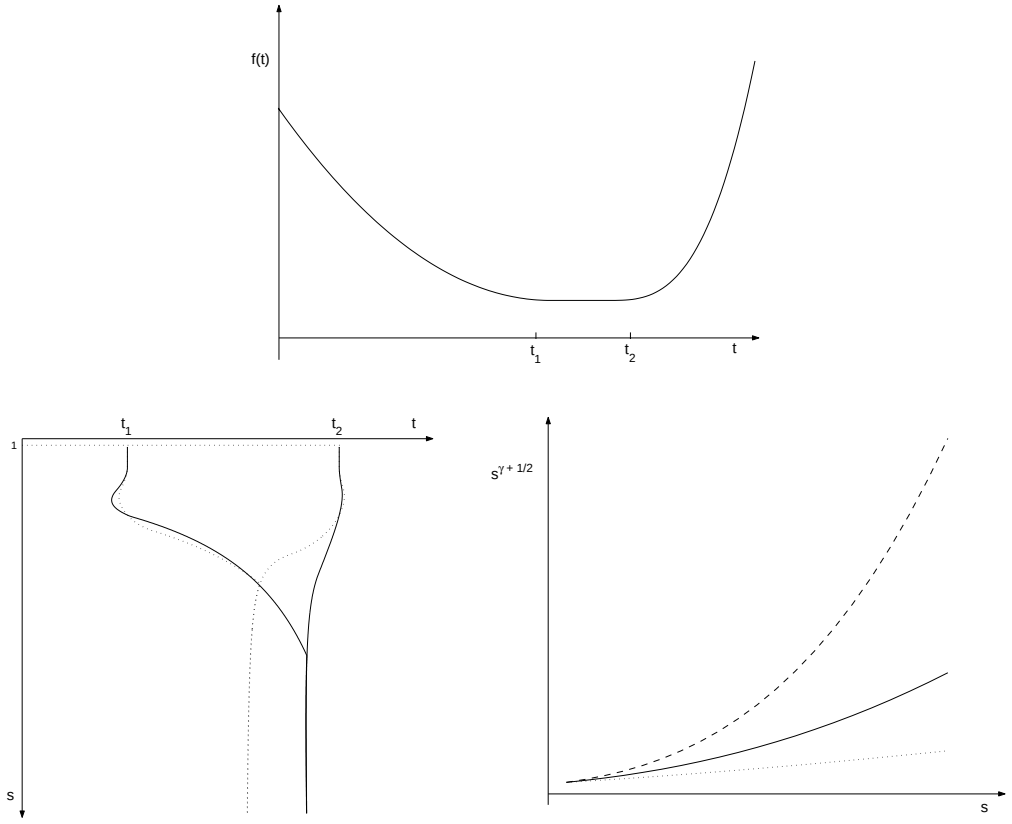


Figure 13: **Top:** Piecewise polynomial signal composed of three parts (with two singularities) whose orders are respectively $2^{nd}, 1^{st}$ and 3^{rd}. **Bottom:** Its modulus maxima trajectories (solid lines) and trajectories of modulus maxima of symmetric atoms used for their approximation (dashed) (*left*). Decays of the first (solid) and second (dashed) original singularity of the signal in a) and the one using atoms approximation (dotted) (*right*). The problem of the decay deviation can be solved using the generalized atom in (24).

Also in this case, we can consider the interaction between two atoms, centered at t_1 and t_2, and having Lipschitz coefficients γ_1 and γ_2. Hence,

$$w_s = \frac{\overline{t}-u}{s}w_u + \frac{w}{2s} + \frac{\gamma_1+\gamma_2}{2s}w + \frac{t_2-t_1}{2s}(w_u^{(2)} - w_u^{(1)}) + \frac{\gamma_2-\gamma_1}{2s}(w^{(2)} - w^{(1)}),$$

and

$$\dot{u} = -\frac{\overline{t}-u}{s} - \left(\frac{t_2-t_1}{2s}\right)\frac{w_{uu}^{(2)} - w_{uu}^{(1)}}{w_{uu}} - \left(\frac{\gamma_2-\gamma_1}{2s}\right)\frac{w_u^{(2)} - w_u^{(1)}}{w_{uu}},$$

where $w_u^{(k)}$ and $w_{uu}^{(k)}, \quad k = 1, 2$, are first and second order partial derivatives with respect to u of $k-$th atom, i.e.

$$w_u^{(k)} = \alpha_k s^{\gamma_k - \frac{1}{2}} \int_{z_k}^{+\infty} \psi(y)dy,$$

$$w_{uu}^{(k)} = \alpha_k s^{\gamma_k - \frac{3}{2}} \psi(z_k).$$

It is easy to verify that when $\gamma_1 = \gamma_2$, we find again (20) and (21).

For N atoms it trivially follows

$$w_s = \frac{\overline{t}-u}{s}w_u + \frac{1+2\sum_{k=1}^{N}\gamma_k}{2s}w + \frac{1}{s}\sum_{k=1}^{N} d_k w_u^{(k)} + \frac{1}{s}\sum_{k=1}^{N}\gamma_k w^{(k)}, \tag{26}$$

where $d_k = t_k - \overline{t}, \quad k = 1, \ldots, N$, and

$$\dot{u} = -\frac{\overline{t}-u}{s} - \frac{1}{s}\frac{\sum_{k=1}^{N} d_k w_{uu}^{(k)}}{w_{uu}} - \frac{1}{s}\frac{\sum_{k=1}^{N}\gamma_k w_u^{(k)}}{w_{uu}}, \tag{27}$$

The solution of the (27) is then 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 Lipschitz exponents of atoms at $s = 1$. This is an important result in terms of compaction of information.

3.3 Estimation of parameters

In the following, we briefly give the algorithm for estimating both slopes α_k and growing exponents γ_k of the atoms. At scale $s = 1$, it is possible to estimate just α_k, since they are not affected by the Lipschitz order γ_k. On the contrary, the algorithm for getting the atomic representation in [1, 2] at $s \neq 1$

gives weights $\alpha_k(s)$ which also account for the decay, i.e. $\alpha_k(s) = \alpha_k(1)s^{\gamma_k - 1}$. Hence, we can estimate the decays γ_k by solving (25) 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 (25) can be solved for each of them, yielding

$$G_s^{(k)} = \frac{\gamma_k + 1/2}{s} G^{(k)}$$

with the following initial condition $G^{(k)}(u(1), 1) = \alpha_k F(t_0, t_0, 1)$.
Hence,

$$G^{(k)} = C_0 s^{\gamma_k + 1/2},$$

that is

$$\alpha_k(s) s \sqrt{s} = \alpha_k(1) s^{\gamma_k + 1/2},$$

and then

$$\gamma_k = \log_s(\alpha_k(s)/\alpha_k(1)) + 1. \tag{28}$$

4 Denoising

The atoms based approximation improves the sparsity of the wavelet representation, since each atom models a set of coefficients. This property has been successfully exploited in de-noising [1] and allows to overcome some of the limits of standard approaches, like thresholding and attenuation. In fact, the former retains those coefficients whose value over-exceeds a given threshold, while put to zero the remaining ones. Even if this operation eliminates most of the noise, it also suppresses significant clean information. On the other hand, attenuation based approaches require a good prediction of the original information for tuning filter coefficients.

The peculiarity of the atomic representation is the modelling of all the coefficients in the atom domain while attenuating the noisy ones. In fact, atoms weights are estimated in the least squares by imposing a function model (see [1]). This operation obviously regularizes noisy data.

Even if WISDOW algorithm provides satisfying results in de-noising, it reveals some limits. The interference between atoms significantly affect the estimation of each single slope and the corresponding location. In fact, it is difficult to exactly separate each single contribution when two atoms are very close. Moreover, annoying spikes can affect the recovered signal since high amplitude noisy coefficients can be confused with real atoms.

The evolution laws, proposed in the previous section, allow to improve these results. In fact, they enable to separate atoms contribution at a fixed scale, yielding a better estimation of their weights. Therefore, for each domain of estimation we can predict how many contributions have to be considered along with their locations. The building of maxima chains enables to use many scale levels in the decomposition without loosing important information. In fact, at coarsest scales the noise flattens since its negative Lipschitz order γ [12] while the estimation domains become wider, since the dilation property of the wavelet transform. It turns out that the least squares are more precise. It is worth to further outlining that atoms decay allows to discard those coefficients dominated by noise, since their negative γ.

4.1 The Algorithm for noisy signals

In this section the algorithm for restoring a noisy signal g is described in detail.

Let us consider an overcomplete wavelet decomposition [11] of g. The overcomplete representation is employed to avoid problems due to decimation. In fact, this operation unavoidably causes distortions of the shape in Fig. 2. In particular, the phase of the decimation (even or odd) differently affects the symmetry of the atom and then it becomes impossible to draw modulus maxima trajectories.

Hence

1. Perform the over-complete wavelet decomposition 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 the algorithm for atoms estimation in [1] for estimating α_k and t_k at scale $s = 1$ and (28) for getting the corresponding decays.

4. Eliminates atoms having $\gamma_k < 0$.

5. Compute atoms trajectories by solving (27) using a 4^{th} order Runge Kutta method and extract the solution at dyadic scales $s = 2^j, \quad j = 1, \ldots, J$.

6. For each maximum, sorted in decreasing order with respect to their amplitude, at scale $s = 2^j, \quad j = 1, \ldots, J$, apply the algorithm for α_k estimation [1, 2]. 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.

7. Invert the overcomplete wavelet decomposition.

It is worth outlining that the estimation of γ_k depends on the estimation of the corresponding slope α_k. The better α_k estimation the more faithful γ_k value. To this aim we can iterate the algorithm used for slopes estimation at step 3 of the algorithm. More precisely, if $E(u, s)$ is the error for the atomic approximation at a fixed scale s, we can iterate the decomposition algorithm on it and combine the results with the ones found at the previous step. This iteration improves the final result.

5 Some experimental results

The denoising algorithm has been tested on several signals an images. As regards, images they are split into independent mono-dimensional signals. In fact, the proposed evolution laws are defined just for the 1D case. Nonetheless, even using this approximation results are very satisfying and approach the most effective denoising techniques.

A biorthogonal wavelet 3/9 associated to an over-complete decomposition has been adopted in all tests. Four scale levels are used for the decomposition and the integration step $h = 0.05$ has been selected for solving the ode in the step 5 of the denoising algorithm.

The modified *WISDOW* has been compared with some classical denoising algorithms, such as soft [5] and hard thresholding [6] and Wiener filtering [11], as shown in Fig. 14 and in Table 4.1. The piecewise polynomial signal, as in chap. 10 of [11], has been chosen for comparing both theoretical and

noisy signal	hard thresholding	soft thresholding	Wiener filter	Modified WISDOW
20.39	26.98	24.25	24.06	32.38
25.08	30.45	27.34	29.48	34.39
28.03	33.44	29.70	32.65	35.49

Table 1: Piecewise polynomial test signal (Fig. 14): a comparison between the proposed model (modified *WISDOW*) and some classical approaches for denoising in terms of SNR (*Signal to Noise Ratio*) values. Wiener filter has been implemented using the Matlab function *wiener2*.

experimental results. In fact, it is composed of jump discontinuities and therefore, it is suitable for testing the interference of singularities, since each discontinuity point is modelled as superimposition of two or more elementary atoms.

It is worth outlining that the atomic representation allows us to also predicts coefficients whose value is under the adopted threshold. That is why ripples around discontinuity points are strongly reduced. This holds even in severe noisy conditions.

WISDOW has also been compared with some of the most recent and effective approaches for image de-noising. $512 \times 512 \times 8$ bits Lena image has been chosen as test image (Fig. 15a) and results have been compared in terms of PSNR (*Peak Signal to Noise Ratio*). Similar results have been achieved on the several test images used in the experiments.

Experimental results show that the proposed denoising algorithm outperforms the most effective wavelet based denoising approaches. In particular, Fig. 16 compares the proposed model with the gaussian mixture estimator presented in [15], the adaptive bayesian thresholding using context modelling contained in [3] and the local Wiener filtering using elliptic directional windows for different subband in [16]. In fact, evolution law establishes a precise link between corresponding coefficients at different scales and allows us to well manage the interference between singularities even at coarser scales. This entails an almost faithful reconstruction of the original signal, avoiding constraints on the minimum distance between them, as in [7].

For a visual evaluation of the results, in Fig. 15c the denoised Lena image is shown (noise standard deviation is $\sigma = 20$).

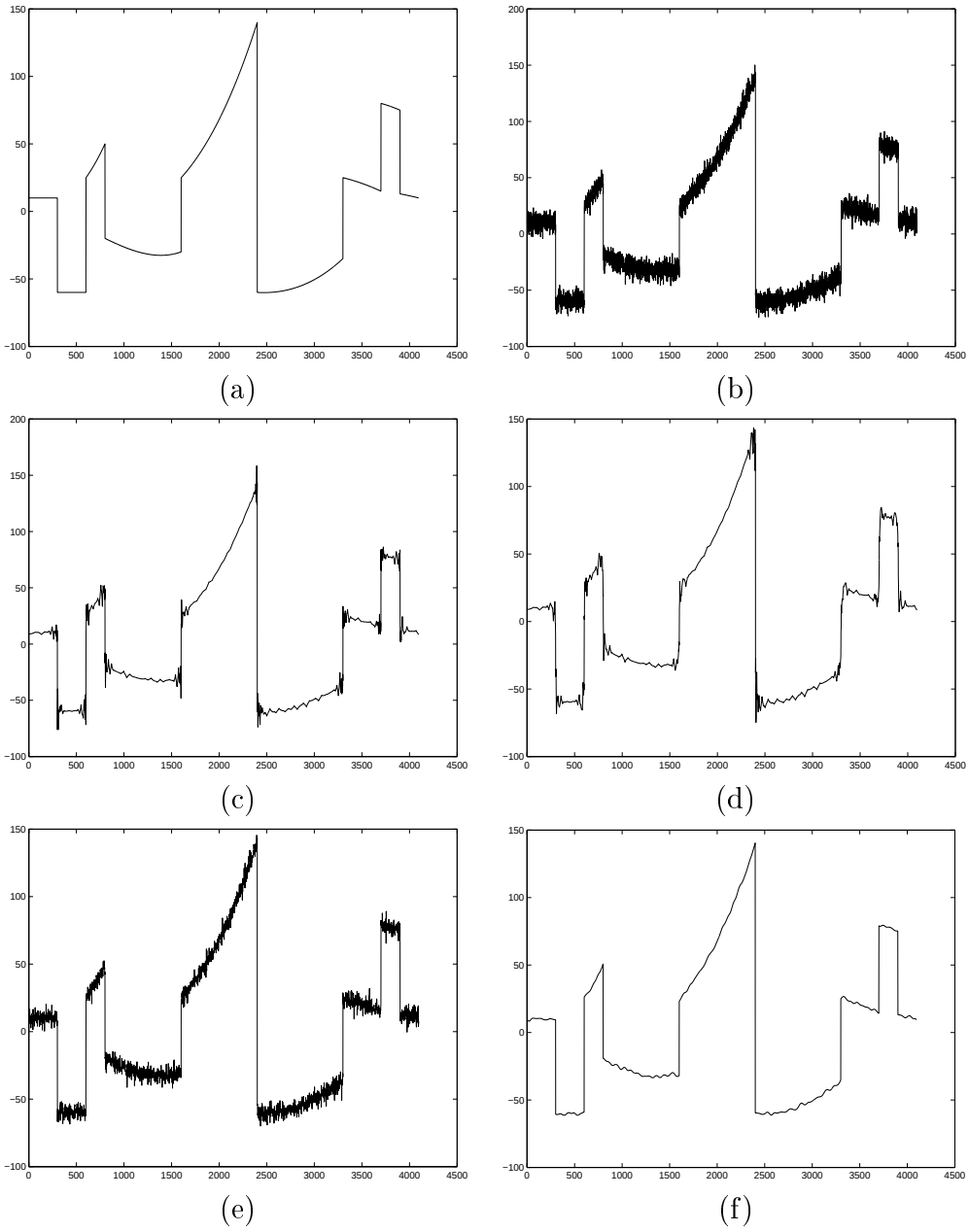


Figure 14: Piecewise polynomial signal: (a) Original signal (b) Noisy Signal (SNR = 20.37 db). (c) Hard-thresholded signal (SNR = 26.88 db). (d) Soft-thresholded signal (SNR = 23.93 db). (e) Wiener filter estimation (SNR = 24.18 db). (f) modified *WISDOW* estimation (SNR = 32.38 db).

(a)

(b)

(c)

Figure 15: (a) Original Lena image ($512 \times 512 \times 8bits$); (b) Noisy image (PSNR = 22.06 db). (c) Denoised Lena image using the proposed model (PSNR = 32.75 db).

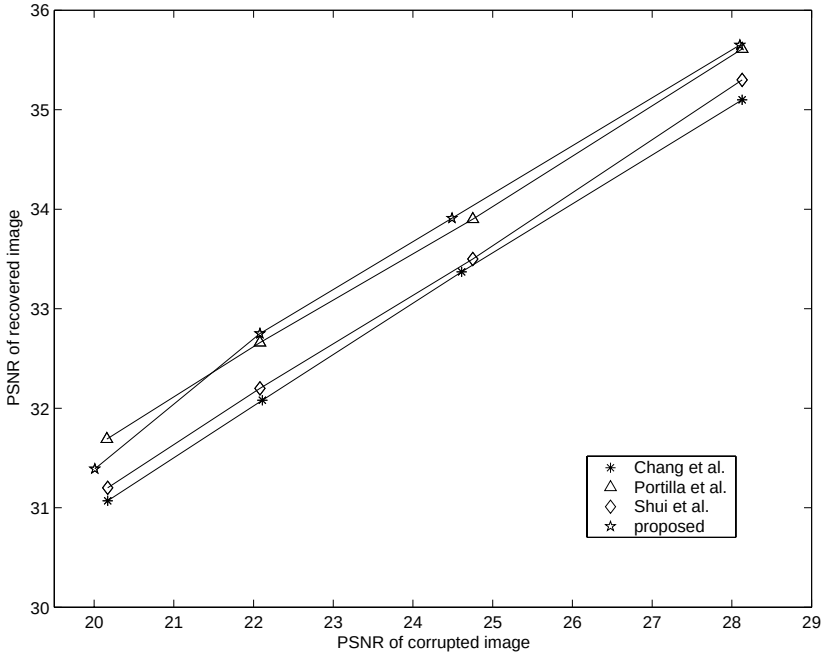


Figure 16: PSNR values versus noise variance for [15], [3], [16] and modified *WISDOW*.

6 Conclusions

In this paper a model for de-noising signals and images has been proposed. It is based on the definition of the atomic representation of the wavelet transform of a generic function. Each atom is associated to the wavelet response to a singularity point in the original signal.

It has been proved that this kind of representation is able to exploit both inter- and intra-scale dependencies of wavelet coefficients by reducing the redundancy of information. In particular, the evolution law through scales for the atoms is provided and their interactions have been investigated. All these properties make the atomic representation a powerful tool for de-noising even under severe noise conditions. The quality of the restored signals is greatly improved, since atoms strongly reduce annoying ripples around discontinuity points, according to the model. Several comparative studies show that it is able to approach and outperform the more recent wavelet based techniques for de-noising.

References

[1] V. Bruni and D. Vitulano, "Wavelet based Signal De-noising via Simple Singularities Approximation, " to appear in *Signal Processing Journal* Elsevier Science.

[2] V. Bruni and D. Vitulano, "Image and Signal Denoising in a Fixed Wavelet Basis", IAC Report, CNR, No. 32, January 2004.

[3] S.G. Chang, Bin Yu and M. Vetterli, *Spatially Adaptive Thresholding with Context Modeling for Image Denoising*, IEEE Transactions on Image Processing, Vol. 9, No. 9, pp. 1522-1530, September 2000.

[4] 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.

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

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

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

[8] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, Vol. 19, American Mathematical Society, 1999.

[9] R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, 2^{nd} ed., Prentice Hall Inc., 2002.

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

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

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

[13] S. Mallat and Z. Zhang, *Matching Pursuits with Time Frequency Dictionaries*, IEEE Transactions on Signal Processing, Vol. 41, No. 12, pp. 3397-3415, December 1993.

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

[15] J. Portilla, V. Strela, M. Wainwright and E. Simoncelli, *Image Denoising using Scale Mixtures of Gaussians in the Wavelet Domain*, IEEE Transactions on Image

[16] 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.