Computing 61, 215–234 (1998)

Computing

2D-Discontinuity Detection from Scattered Data

M. Rossini, Milano

Received February 6, 1998; revised April 8, 1998

Abstract

We describe a numerical approach for the detection of discontinuities of a two dimensional function distorted by noise. This problem arises in many applications as computer vision, geology, signal processing. The method we propose is based on the two-dimensional continuous wavelet transform and follows partially the ideas developed in [2], [6] and [8]. It is well-known that the wavelet transform modulus maxima locate the discontinuity points and the sharp variation points as well. Here we propose a statistical test which, for a suitable scale value, allows us to decide if a wavelet transform modulus maximum corresponds to a function value discontinuity. Then we provide an algorithm to detect the discontinuity curves from *scattered and noisy data*.

AMS Subject Classifications: 65R30, 65D15, 65R10.

Key words: Discontinuities, wavelets, edge-detection.

1. Introduction

We describe a method to detect the curves across which a two dimensional function distorted by noise is discontinuous. Before starting, we state the problem. Our aim is to detect a discontinuity curve from a set of functional data $\{z_i\}_{i=1}^n$ which has been sampled at some points $\{\mathbf{r}_i\}_{i=1}^n$ of a parameter domain $Q \subset \mathbb{R}^2$. The values $\{z_i\}_{i=1}^n$ can be thought as a realization

$$z_i = f(\mathbf{r}_i) + e(\mathbf{r}_i), \qquad i = 1, \dots, n \tag{1}$$

of a process

$$z(x, y) = f(x, y) + e(x, y), \tag{2}$$

where the trend $f(x, y)$ is discontinuous across an unknown curve ℓ of Q and smooth in any neighbourhood of Q which does not intersect ℓ; $e(x, y)$ is white noise with expected value $E\{e\} = 0$ and variance $E\{e^2\} = \sigma^2 < \infty$;

We suppose ℓ being a Q one-curve with equation $y = l(x)$ (all the things we prove, hold also for the case $x = l(y)$). Without loss of generality we will suppose Q to be the unitary square $[0,1] \times [0,1]$.

This problem arises in many applications as computer vision, geology, signal processing. In fact, the most important information of a particular phenomenon is often carried by irregular structures. The accurate detection of such discontinuity curves is then of basic importance to analyze and recover the considered phenomenon correctly. We can think, for instance, about subsoil or depth faults which represent discontinuities in geological layers caused by severe movements of the earth crust. Their localization provides useful information for geologists about, for instance, the occurrence of oil reservoirs.

The method we propose in this paper is based on the two-dimensional continuous wavelet transform and follows partially the ideas developed in [2], [6] and [8].

It is well-known (see for instance [4], [6]) that a signal can be well localized in time and frequency by the wavelet transform which is therefore well-adapted to describe transient phenomena like signal sharp variation and singularities. Moreover, it is possible to characterize the local regularity by theorems relating the Lipschitz exponent to the wavelet transform evolution across scales. It is known [6], [7] that the important information is carried by the wavelet transform modulus local maxima which locate either the strong function variation or the discontinuities.

The question is how to distinguish between maxima corresponding to discontinuities and maxima corresponding to sharp but continuous variation. The noise fluctuations, then, introduce false discontinuities and, as consequence, we could also find maxima due to the noise.

For the one-dimensional case, Mallat and Hwang provide in [6] a numerical procedure to estimate the Lipschitz exponent for some kind of singularities when the function is not distorted by noise. They also provide a denoising algorithm, for one and two dimensional signals, which removes the maxima due to the noise by studying their evolution across different scales. Moreover, for the two dimensional case, they discriminate the irregularities due to the noise also using some coherence hypotheses: the singularities belong to regular curves and vary smoothly along these curves while the noise, usually, does not produce smooth curves. In this way they detect the signal important structures. But, as already proved in [2] for the one dimensional case and how it will be shown later, this is not enough to solve our problem.

For this reason we propose a method, based on a statistical test, which, for a suitable scale value, allows us to decide if a wavelet transform modulus maximum corresponds to a function value discontinuity. In Section 2 we will introduce the two dimensional wavelet transform; according to Canny's edge detection [3], we will define the wavelet transform modulus maxima and we'll prove a theorem which characterizes the wavelet transform behaviour in a neighbourhood of a maximum corresponding to a function value discontinuity. We will consider the presence of the noise in Section 3 and, using the results obtained in Section 2, we will perform a statistical test to detect the discontinuity curve. In Section 4 we shall provide a detection algorithm from data collected on scattered points $\{\mathbf{r}_i\}_{i=1}^n$ of Q. In applied problems we often deal with this situation which presents more difficulties than the case of gridded data already studied in [8]. Finally, some numerical results are shown in Section 5.

2. Introductory Material and the Characterization Theorem

In this section we recall, following [6], some notations and definitions which will be largely used in the paper.

For any function $h(x,y)$, $h_a(x,y)$ denotes the dilation of $h(x,y)$ by the scale factor $a, a \in \mathbb{R}$, $a > 0$

$$h_a(x,y) = \frac{1}{a^2} h\left(\frac{x}{a}, \frac{y}{a}\right).$$

Let us call *smoothing* function any function $\phi(x)$ such that $\phi(x) = O(1/(1+x^2))$ and whose integral $\int_{\mathbb{R}} \phi(x)dx$ is nonzero. We call two-dimensional *smoothing* function, any function $\theta(x,y)$ whose double integral is nonzero. We define two basic wavelets $\psi_1(x,y)$, $\psi_2(x,y)$ that are, respectively, the partial derivative along x and y of a *smoothing function* $\theta(x,y)$

$$\psi_1(x,y) = \frac{\partial}{\partial x}\theta(x,y), \quad \psi_2(x,y) = \frac{\partial}{\partial y}\theta(x,y). \tag{3}$$

The two-dimensional wavelet transform of a function $f(x,y) \in L^2(\mathbb{R}^2)$ defined with respect to $\psi_{1,a}(x,y)$ and $\psi_{2,a}(x,y)$ has two components

$$W_{1,f,a}(b_1,b_2) = f * \psi_{1,a}(b_1,b_2),$$
$$W_{2,f,a}(b_1,b_2) = f * \psi_{2,a}(b_1,b_2).$$

Let us consider the vector

$$\mathbf{W}_{f,a}(b_1,b_2) = \left(W_{1,f,a}(b_1,b_2), W_{2,f,a}(b_1,b_2)\right); \tag{4}$$

we call wavelet transform modulus at the scale a the modulus of $\mathbf{W}_{f,a}$, that is the function

$$SW_{f,a}(b_1,b_2) = \sqrt{W_{1,f,a}^2(b_1,b_2) + W_{2,f,a}^2(b_1,b_2)}. \tag{5}$$

We consider the angle

$$\alpha_{\mathbf{w}}(b_1,b_2) = \arctan\left(\frac{W_{2,f,a}(b_1,b_2)}{W_{1,f,a}(b_1,b_2)}\right) \tag{6}$$

between the vector (4) and the horizontal.

One can easily prove that the two components of the wavelet transform are proportional to the gradient vector of f smoothed by θ_a. Canny [3] defines the edge points at the scale a as the points where the modulus of (4) is maximum in the direction where the vector points to. This direction is where f has the sharpest variation. Using this approach, Mallat and Hwang have defined the modulus maxima. As in [6] we call wavelet transform modulus maxima, the points $(\bar{x}, \bar{y})$ of Q where the function $SW_{f,a}$ is locally maximum along the direction given by $\alpha_w(\bar{x}, \bar{y})$. Namely a point $(\bar{x}, \bar{y})$ is called modulus maximum if the restriction of $SW_{f,a}(b_1, b_2)$ to the straight line of Q, $b_2 = \tan(\alpha_w(\bar{x}, \bar{y}))(b_1 - \bar{x}) + \bar{y}$, is locally maximum at $b_1 = \bar{x}$. Then the modulus maxima locate the sharp

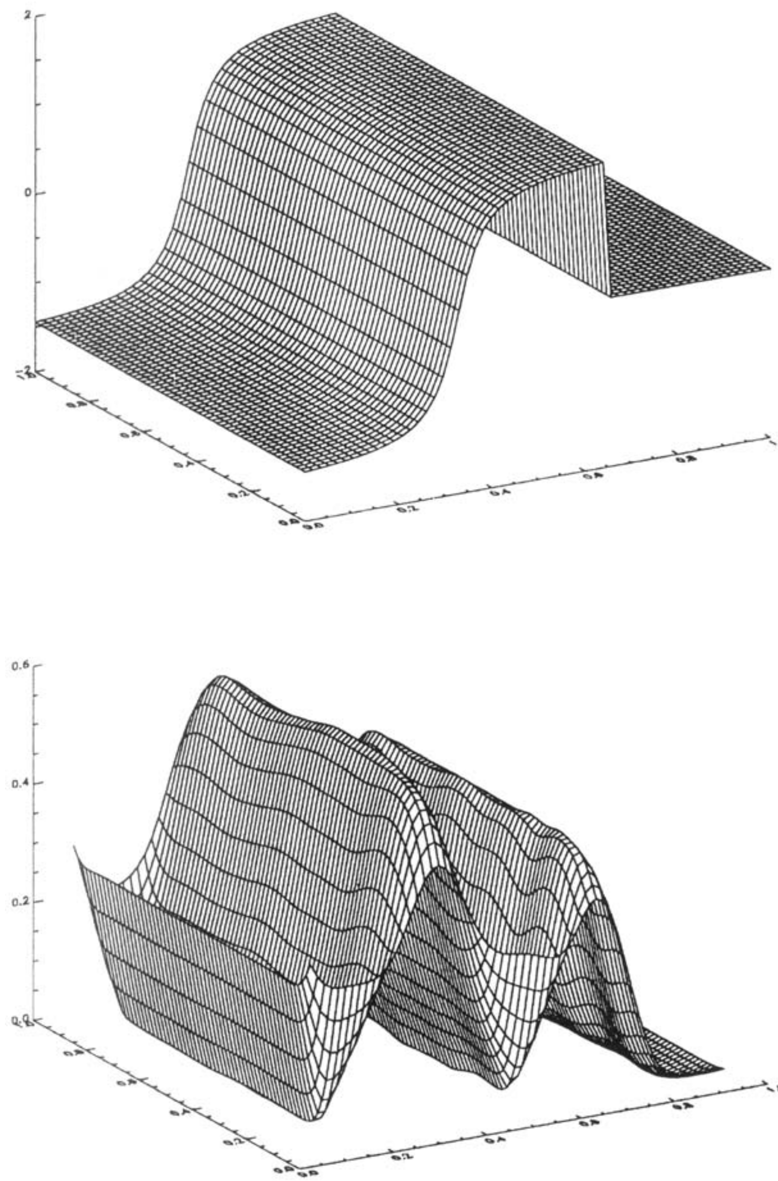


Figure 1. Top: the function, bottom: the modulus of its wavelet transform

variation of f, and the angle (6) indicates locally the direction where the signal has the sharpest variation.

From above it is clear that the modulus maxima of $SW_{f,a}$ locate either the discontinuity curve ℓ or the sharp but continuous variation points of f. The example of Fig. 1 explains what is stated above: the function f is shown in the top picture, while its wavelet transform modulus is displayed on the bottom. As we can easily see, the graphic shows two curves of modulus maxima: the straight line $x = 0.65$ across which f is discontinuous and the straight line $x = 0.325$ across which f varies sharply.

It follows that it is not possible to tell which is the discontinuity curve analysing $SW_{a,f}$ only.

It is then necessary to dispose of an instrument to distinguish between this different situation. Our approach is based on Theorem 2.1 which, for a fixed scale a, gives the wavelet transform modulus value at the points belonging to a neighbourhood of a discontinuity point (see the following relation (8)). Theorem 2.1 has been proved in the particular case a constant C exists such that the function $f(x, y)$ is equal to a constant k_1 for all $(x, y) \in Q$ with $l(x) < y < l(x) + C$, and equal to a different constant k_2 for all $(x, y) \in Q$ with $l(x) - C < y < l(x)$.

Theorem 2.1. *Let $f(x, y) \in L^2(\mathbb{R}^2)$ and suppose $f(x, y)$ be discontinuous across the planar curve ℓ of equation $y = l(x)$. Let the following conditions hold*

i) *a constant C exists such that $\forall(x, y) \in Q$ with $l(x) < y < l(x) + C$, $f(x, y)$ is equal to a constant k_1 and equal to a different constant k_2 $\forall(x, y)$ with $l(x) - C < y < l(x)$; let $\Omega = \{(x, y) \in Q$ such that $l(x) - C < y < l(x) + C\}$*

ii) *let $l(x)$ be continuous and piecewise derivable with $|l'(x)| \leq H < \infty$.*
Let us consider a family of wavelets $\psi_{1,a}$, $\psi_{2,a}$ such that

iii) *$\theta(x, y)$ is equal to $\phi(x)\phi(y)$, where ϕ is a one-dimensional smoothing function; let ψ be the first derivative of ϕ;*

iv) *the basic wavelets* (3) *have compact support* $[-A, A] \times [-A, A]$.
Set for simplicity

$$c(x, y) = \int_{-A}^{A} \phi(t)\phi\left(\frac{l(at + x) - y}{a}\right) l'(at + x)\,dt,$$

$$d(x, y) = \int_{-A}^{A} \phi(t)\phi\left(\frac{l(at + x) - y}{a}\right) dt,$$

and

$$G(x, y) = \sqrt{c^2(x, y) + d^2(x, y)}.$$

Let (x_0, y_0) be a point belonging to ℓ. If the scale parameter a is such that the support of $\psi_{1,a}(x - x_0, y - y_0)$ and $\psi_{2,a}(x - x_0, y - y_0)$ is included in Ω, that is

$v)$ $[x_0 - aA, x_0 + aA] \times [y_0 - aA, y_0 + aA] \subset \Omega$,

we have

$$SW_{f,a}(x_0, y_0) = |k_1 - k_2| G(x_0, y_0), \tag{7}$$

where $G(x_0, y_0) \neq 0$.

Moreover, let (b_1, b_2) *be such that the support of* $\psi_{1,a}(x - b_1, y - b_2)$ *and* $\psi_{2,a}(x - b_1, y - b_2)$ *is included in* Ω, *that is* $[b_1 - aA, b_1 + aA] \times [b_2 - aA, b_2 + aA] \subset \Omega$. *Then, if*

$vi)$ (b_1, b_2) *belongs to the neighbourhood of* (x_0, y_0), $I(x_0, y_0) = [x_0 - aA/2, x_0 + aA/2] \times [y_0 - aA/2, y_0 + aA/2]$ *but not to* ℓ,

we have that the value of the wavelet transform modulus in (b_1, b_2) *is given by the following relation*

$$SW_{f,a}(b_1, b_2) = \frac{G(b_1, b_2)}{G(x_0, y_0)} SW_{f,a}(x_0, y_0). \tag{8}$$

Proof: Let (x_0, y_0) be a point of ℓ and consider $W_{1,f,a}(x_0, y_0)$, $W_{2,f,a}(x_0, y_0)$. By definition and by *iv)*

$$W_{1,f,a}(x_0, y_0) = \frac{1}{a^2} \int_{y_0 - aA}^{y_0 + aA} \int_{x_0 - aA}^{x_0 + aA} f(x, y) \psi_1\left(\frac{x - x_0}{a}, \frac{y - y_0}{a}\right) dx\, dy,$$

$$W_{2,f,a}(x_0, y_0) = \frac{1}{a^2} \int_{y_0 - aA}^{y_0 + aA} \int_{x_0 - aA}^{x_0 + aA} f(x, y) \psi_2\left(\frac{x - x_0}{a}, \frac{y - y_0}{a}\right) dx\, dy.$$

Observe that the domain of integration is the union of the two sub-domains having ℓ as common boundary. Let a be such that they are normal with respect to x or to y. Let us suppose they are normal with respect to x (the other case is the same). So, if $v)$ holds, splitting the domain of integration, making the change of variables $t = (x - x_0)/a, u = (y - y_0)/a$, using *iii)* and using *i)*, we get

$$W_{1,f,a}(x_0, y_0) = k_1 \int_{-A}^{A} dt \int_{\frac{l(at + x_0) - y_0}{a}}^{A} \psi(t) \phi(u) du$$

$$+ k_2 \int_{-A}^{A} dt \int_{-A}^{\frac{l(at + x_0) - y_0}{a}} \psi(t) \phi(u) du, \tag{9}$$

$$W_{2,f,a}(x_0, y_0) = k_1 \int_{-A}^{A} dt \int_{\frac{l(at + x_0) - y_0}{a}}^{A} \phi(t) \psi(u) du$$

$$+ k_2 \int_{-A}^{A} dt \int_{-A}^{\frac{l(at + x_0) - y_0}{a}} \phi(t) \psi(u) du. \tag{10}$$

Integrating (9) and (10) we obtain

$$W_{1,f,a}(x_0,y_0)=(k_1-k_2)\int_{-A}^{A}\phi(t)\phi\left(\frac{l(at+x_0)-y_0}{a}\right)l'(at+x_0)dt, \quad (11)$$

$$W_{2,f,a}(x_0,y_0)=(k_2-k_1)\int_{-A}^{A}\phi(t)\phi\left(\frac{l(at+x_0)-y_0}{a}\right)dt. \quad (12)$$

Then, from (5), (11), and (12) we get (7).

Let us observe that $G(x_0,y_0)\neq 0$. This follows immediately from the assumptions on the smoothing function θ.

Consider a point (b_1,b_2) such that $|b_1-x_0|<aA/2$, $|b_2-y_0|<aA/2$ and such that $[b_1-aA,\ b_1+aA]\times[b_2-aA,\ b_2+aA]\subset\Omega$. Going on as in the previous part of the proof, we get

$$W_{1,f,a}(b_1,b_2)=(k_1-k_2)\int_{-A}^{A}\phi(t)\phi\left(\frac{l(at+b_1)-b_2}{a}\right)l'(at+b_1)dt, \quad (13)$$

$$W_{2,f,a}(b_1,b_2)=(k_2-k_1)\int_{-A}^{A}\phi(t)\phi\left(\frac{l(at+b_1)-b_2}{a}\right)dt. \quad (14)$$

Then from (5), (13), and (14) we get

$$SW_{f,a}(b_1,b_2)=|k_1-k_2|\sqrt{c^2(b_1,b_2)+d^2(b_1,b_2)}. \quad (15)$$

Combining (7) with (15) we get (8).

Remark 1. *It is not difficult to show that when* $(b_1,b_2)\to(x_0,y_0)$, $G(b_1,b_2)\to G(x_0,y_0)$.

The theorem has been proved in a particular case. Usually a function which is discontinuous across a curve does not satisfy to *i)*. For the general case we have the following result.

Corollary 2.2. *Let the hypotheses of Theorem* 2.1 *hold and replace* i) *with*
I) $f(x,y)\in C^1\forall(x,y)$ *with* $y<l(x)$ *and* $\forall(x,y)$ *with* $y>l(x)$.

Let (x_0,y_0) *be a point of* ℓ. *Then for all points* (b_1,b_2) *satisfying to condition vi) of Theorem* 2.1, *we have*

$$SW_{f,a}^2(b_1,b_2)=\frac{G^2(b_1,b_2)}{G^2(x_0,y_0)}\,SW_{f,a}^2(x_0,y_0)+\varepsilon(b_1,b_2,a), \quad (16)$$

where for all fixed a, $\varepsilon(b_1,b_2,a)\to 0$ *when* $(b_1,b_2)\to(x_0,y_0)$.

Proof: Using the mean value theorem and taking into account Remark 1, we get the proof.

Then if (b_1, b_2) belongs to a suitable neighbourhood of a discontinuity point (x_0, y_0), $G^2(b_1, b_2)/G^2(x_0, y_0)SW^2_{f,a}(x_0, y_0)$ will approximate the value $SW^2_{f,a}(b_1, b_2)$.

Finally we want to show how relation (8) of Theorem 2.1 is modified when we choose a non compactly supported wavelet.

Corollary 2.3. *Let the assumptions of Theorem* 2.1 *hold except for iv), and consider as smoothing function*

$$\theta(x,y) = \exp\left(\frac{-\beta x^2}{2}\right)\exp\left(\frac{-\beta y^2}{2}\right),$$

with $\beta \geq 1$. *Let* $\bar{A}$ *be a real number such that* $[x_0 - a\bar{A}, x_0 + a\bar{A}] \times [y_0 - a\bar{A}, y_0 + a\bar{A}] \subset \Omega$. *Then if* (b_1, b_2) *belongs to* $I(x_0, y_0) = [x_0 - a\bar{A}/2, x_0 + a\bar{A}/2] \times [y_0 - a\bar{A}/2, y_0 + a\bar{A}/2]$ *but not to* ℓ *and if* $[b_1 - a\bar{A}, b_1 + a\bar{A}] \times [b_2 - a\bar{A}, b_2 + a\bar{A}] \subset \Omega$, *relation* (8) *becomes*

$$SW^2_{f,a}(b_1, b_2) = \frac{G^2(b_1, b_2)}{G^2(x_0, y_0)} SW^2_{f,a}(x_0, y_0) + O\left(\exp\left(\frac{-\beta\bar{A}^2}{2}\right)\right), \tag{17}$$

where in the expression of G, A *has been changed with* $\bar{A}$.

Proof: The proof follows immediately from the exponential decay of the wavelets we considered.

Thus, if we choose β and $\bar{A}$ properly, $SW^2_{f,a}(b_1, b_2)$ can be approximated by $(G^2(B_1, B_2)/G^2(x_0, y_0))SW^2_{f,a}(x_0, y_0)$.

The results of Theorem 2.1, Corollary 2.2 and Corollary 2.3, suggest how to detect the function value discontinuities. Let us remember that the modulus maxima of $SW_{f,a}$ locate the discontinuities and the sharp variation as well, but only the maxima corresponding to discontinuities will satisfy (8), (16) or (17).

Then, if there is a wavelet transform modulus maximum in $(\bar{x}, \bar{y})$, we can match $G^2(b_1, b_2)/G^2(\bar{x}, \bar{y})SW^2_{f,a}(\bar{x}, \bar{y})$ with $SW^2_{f,a}(b_1, b_2)$, for $(b_1, b_2) \in I(\bar{x}, \bar{y})$. A successful matching will indicate that the wavelet maximum $(\bar{x}, \bar{y})$ belongs to the discontinuity curve ℓ.

3. Detection of ℓ from the Process z

The previous section results provide a method to tell if a maximum belongs to the discontinuity curve ℓ. Now we consider the presence of the noise. Let us consider the process (2), its two-dimensional continuous wavelet transform

$$\mathbf{W}_{z,a}(b_1, b_2) = \mathbf{W}_{f,a}(b_1, b_2) + \mathbf{W}_{e,a}(b_1, b_2), \tag{18}$$

and its wavelet modulus maxima.

Usually abrupt changes due to the noise may introduce false discontinuities. It is known that when the scale a increases, the noise fluctuations decrease; in fact the signal–to–noise ratio, as proved in [2] for the one dimensional case and in the following Proposition 3.1 for the present situation, is proportional to a. Moreover (see [2] and [6]) the expected number of maxima due to the noise is proportional to $1/a$. By consequence, when a increases, the signal dominates the noise and the maxima due to the noise are expected to disappear.

It is important to point out that, on one hand, the noise cannot be completely removed and, on the other increasing a means to loose information on the underlying function (see, for instance Example 3 of Section 5). Hence the choice of a is very important. If we choose a "too small" the noise dominates the signal and it may destroy all the knowledge given by the wavelet transform. If we choose a "large", the signal prevails on the noise but we can loose significant information. It is then important to find the right balance between reducing the spurious responses and preserving the information on the trend.

Taking into account the above considerations, we have studied a method based on a statistical test which, for a suitable scale a, allows us to discriminate the maxima corresponding to discontinuity points either from the sharp variation points or from the maxima due to the noise.

To this purpose we associate to each maximum $(\bar{x}, \bar{y})$ of the wavelet transform, the random variable

$$R^2(b_1, b_2) = \frac{SW^2_{z,a}(b_1, b_2)}{SW^2_{f,a}(\bar{x}, \bar{y}) \dfrac{G^2(b_1, b_2)}{G^2(\bar{x}, \bar{y})}} \tag{19}$$

where (b_1, b_2) belongs to a neighbourhood of $(\bar{x}, \bar{y})$.

In Section 3.1 we will study the properties of (19) and we will show that if a maximum belongs to ℓ and if we choose a properly, then the realizations of (19) are between 0.5 and 1.5 with probability close to one. These results are used in Section 3.2 to perform a test which detects the maxima belonging to ℓ.

Proposition 3.1. *The inverse of the signal-to-noise ratio ρ of the process $SW^2_{z,a}$ at a generic point (b_1, b_2) is*

$$\frac{1}{\rho} = \frac{1}{a} \frac{8\sigma^2\gamma}{SW^2_{f,a}(b_1, b_2)} \left(1 + \frac{1}{a} \frac{\sigma^2\gamma}{SW^2_{f,a}(b_1, b_2)}\right), \tag{20}$$

where $\gamma = \|\psi_1\|_2^2 = \|\psi_2\|_2^2$ and $\sigma^2 = E\{e^2\}$.

Proof: Relation (20) is obtained from the definition of the signal-to-noise ratio ρ [1], and from the statistic of the processes $W_{1,e,a}$, $W_{2,e,a}$, $SW^2_{e,a}$.

3.1. The Properties of $R^2(b_1, b_2)$

In the following propositions are stated the properties of (19) associated to the maxima belonging to the discontinuity curve. Namely, in Proposition 3.2 we shall prove that the expected value and variance of R^2 are respectively $1 + O(1/a)$ and $O(1/a)$. These results will then be used in Proposition 3.3 to prove that the realizations of R^2 are between 0.5 and 1.5 with probability close to one.

Proposition 3.2. *Let (x_0, y_0) be a modulus maximum belonging to ℓ. Let the assumptions of Theorem* 2.1 *hold, and set for simplicity*

$$\gamma = \|\psi_1\|_2^2 = \|\psi_2\|_2^2, \qquad K = \frac{G^2(b_1, b_2)}{G^2(x_0, y_0)} SW_{f,a}^2(x_0, y_0).$$

Then $\forall (b_1, b_2)$ satisfying to condition vi) of Theorem 2.1, *the expected value and the variance of* (19) *are*

$$E\big(R^2(b_1, b_2)\big) = 1 + \frac{1}{a}\frac{2\sigma^2\gamma}{K} \tag{21}$$

$$Var\big(R^2(b_1, b_2)\big) = \frac{1}{a}\frac{4\sigma^2\gamma}{K}\left(1 + \frac{1}{a}\frac{\sigma^2\gamma}{K}\right), \tag{22}$$

where $\sigma^2 = E\{e^2\}$

Proof: Consider a point $(b_1, b_2) \in I(x_0, y_0)$ defined as in *vi)* of Theorem 2.1. From (5) and (8) we get

$$\begin{aligned} SW_{z,a}^2(b_1, b_2) &= K + SW_{e,a}^2(b_1, b_2) \\ &\quad + 2W_{1,f,a}(b_1, b_2) W_{1,e,a}(b_1, b_2) \\ &\quad + 2W_{2,f,a}(b_1, b_2) W_{2,e,a}(b_1, b_2). \end{aligned}$$

Then

$$\begin{aligned} R^2(b_1, b_2) &= 1 + \frac{2W_{1,f,a}(b_1, b_2) W_{1,e,a}(b_1, b_2)}{K} + \frac{SW_{e,a}^2(b_1, b_2)}{K} \\ &\quad + \frac{2W_{2,f,a}(b_1, b_2) W_{2,e,a}(b_1, b_2)}{K}. \end{aligned}$$

Using the statistic of the processes $W_{1,e,a}$, $W_{2,e,a}$, $SW_{e,a}^2$, relations (21), (22) are achieved by simple calculations.

Proposition 3.3. *Let (x_0, y_0) be a modulus maximum belonging to ℓ. If $a \gg \bar{a}$, then $\forall (b_1, b_2)$ satisfying to condition vi) of Theorem* 2.1, *the realizations of $R^2(b_1, b_2)$ fall between* 0.5 *and* 1.5 *with probability P almost one.*

Proof: From the well-known Chebyshev inequality we have that

$$\begin{aligned} P\{.5 < R^2(b_1, b_2) < 1.5\} &= P\{-.5 < R^2(b_1, b_2) - 1 < .5\} \\ &\geq 1 - \frac{E\big((R^2(b_1, b_2) - 1)^2\big)}{(.5)^2}. \end{aligned}$$

Using (21), and (22), and if $a \gg \bar{a}$ we have that

$$E\left(\left(R^2(b_1,b_2) - 1\right)^2\right) \ll (.5)^2.$$

Then

$$P\left\{.5 < R^2(b_1,b_2) < 1.5\right\} \approx 1.$$

By simple algebra, it is not difficult to prove that $\bar{a} = \dfrac{4(\sqrt{6}+2)\sigma^2}{\|f\|_2^2}$.

Therefore, if we select a in $(\bar{a}, +\infty)$ and if (x_0, y_0) is a maximum corresponding to a discontinuity, then

- from Proposition 3.1 we have that the signal prevails on the noise and we can suppose that

$$SW_{f,a}^2(x_0,y_0) \approx SW_{z,a}^2(x_0,y_0);$$

- from Propositions 3.2 and 3.3 we have that $\forall(b_1,b_2) \in I(x_0,y_0)$, the realizations of $R^2(b_1,b_2)$ are between 0.5 and 1.5 with probability almost one

3.2. The Detection Test

The results of the previous sections suggest a detection scheme. We can perform an hypothesis test to decide if a wavelet modulus maximum $(\bar{x},\bar{y})$ is a discontinuity point. The test is based on the fact that if $(\bar{x},\bar{y}) \in \ell$ and if we choose a as stated above, then $\forall(b_1,b_2)$ belonging to the neighbourhood $I(\bar{x},\bar{y})$ defined as in Theorem 2.1, the realizations of $R^2(b_1,b_2)$ are between 0.5 and 1.5 with probability close to one. Hence the test steps are:

- fix a significance level δ;
- extract a random sample $R^2(b_{1,j},b_{2,j})$ $j=1,\ldots,N$, $(b_{1,j},b_{2,j}) \in I(\bar{x},\bar{y})$;
- evaluate the relative frequence π of the event

$$S = \left\{.5 < R^2(b_1,b_2) < 1.5, \quad (b_1,b_2) \in I(\bar{x},\bar{y})\right\}.$$

Since π estimates the probability that $(\bar{x},\bar{y})$ belongs to ℓ, if $\pi \geq 1-\delta$ we accept it as discontinuity point and reject it otherwise.

4. Detection from Scattered and Noisy Data: an Algorithm

In several applied problems, we are given a set of noisy data collected on scattered points of the domain Q. Let us consider n scattered points $\mathbf{r}_i \in Q = [0,1] \times [0,1]$, and a realization (1) of the process (2).

The method we have studied, has essentially two phases: the first is to find the wavelet modulus maxima and the second is to discriminate, by the detection test, the maxima corresponding to discontinuities.

In the first phase of the method, we need to approximate the continuous wavelet transform $\mathbf{W}_{z,a}$ of the process (2) by a function obtained from the sample. In order to do this, we consider a partition of the domain Q in n rectangles Q_i, $i = 1, \ldots, n$ such that in each rectangle Q_i, only one point $\mathbf{r}_i$ falls, and

$$\bigcap_{i=1}^{n} Q_i = \emptyset,$$

$$\bigcup_{i=1}^{n} Q_i = Q.$$

Since the trend f is discontinuous across the curve ℓ, we approximate the process (2) with the function

$$g(x,y) = z_i \quad \forall (x,y) \in Q_i.$$

Then we approximate $\mathbf{W}_{z,a}$ by the wavelet transform $\mathbf{W}_{g,a}$ of $g(x,y)$. It is not difficult the prove that $\mathbf{W}_{z,a} = \mathbf{W}_{g,a} + \mathbf{O}(1/n)$.

Regarding the phase two of the method, let us observe that when we perform the test described in Section 3.2, we need to compute $G(x,y)$ which depends on the unknown ℓ equation $y = l(x)$. But we may observe (see [6]) that if $(\bar{x}, \bar{y})$ belongs to ℓ, the direction along which the modulus of the wavelet transform is maximum is approximately orthogonal to the tangent to ℓ in $(\bar{x}, \bar{y})$. So, in the neighbourhood of $(\bar{x}, \bar{y})$, defined in Theorem 2.1, we can approximate the equation $y = l(x)$ with the tangent

$$y = -\frac{W_{1,g,a}(\bar{x},\bar{y})}{W_{2,g,a}(\bar{x},\bar{y})}(x - \bar{x}) + \bar{y}.$$

We can now sketch the algorithm. Chosen a suitable value of a and a significance level δ, the main steps are

1. To compute $\mathbf{W}_{g,a}$ and $SW_{g,a}$.
2. To compute the modulus maxima $(\bar{x}_\iota, \bar{y}_\iota)$ of the wavelet transform. Let s be their number.
3. For $\iota = 1, \ldots, s$, $m = 0$
   extract a random sample $(B_{1,j}, B_{2,j}) \in I(\bar{x}_\iota, \bar{y}_\iota)$ $j = 1, \ldots, N$.
   For $j = 1, \ldots, N$ compute $R^2(B_{1,j}, B_{2,j})$;
   if $0.5 < R^2(B_{1,j}, B_{2,j}) < 1.5$ then $m = m + 1$, else next j; next j
   $\pi = \frac{m}{N}$, if $\pi \geq 1 - \delta$ then accept $(\bar{x}_\iota, \bar{y}_\iota)$ else reject it; next ι
4. Using the selected points, $y = l(x)$ is reconstructed by some smoothing method.

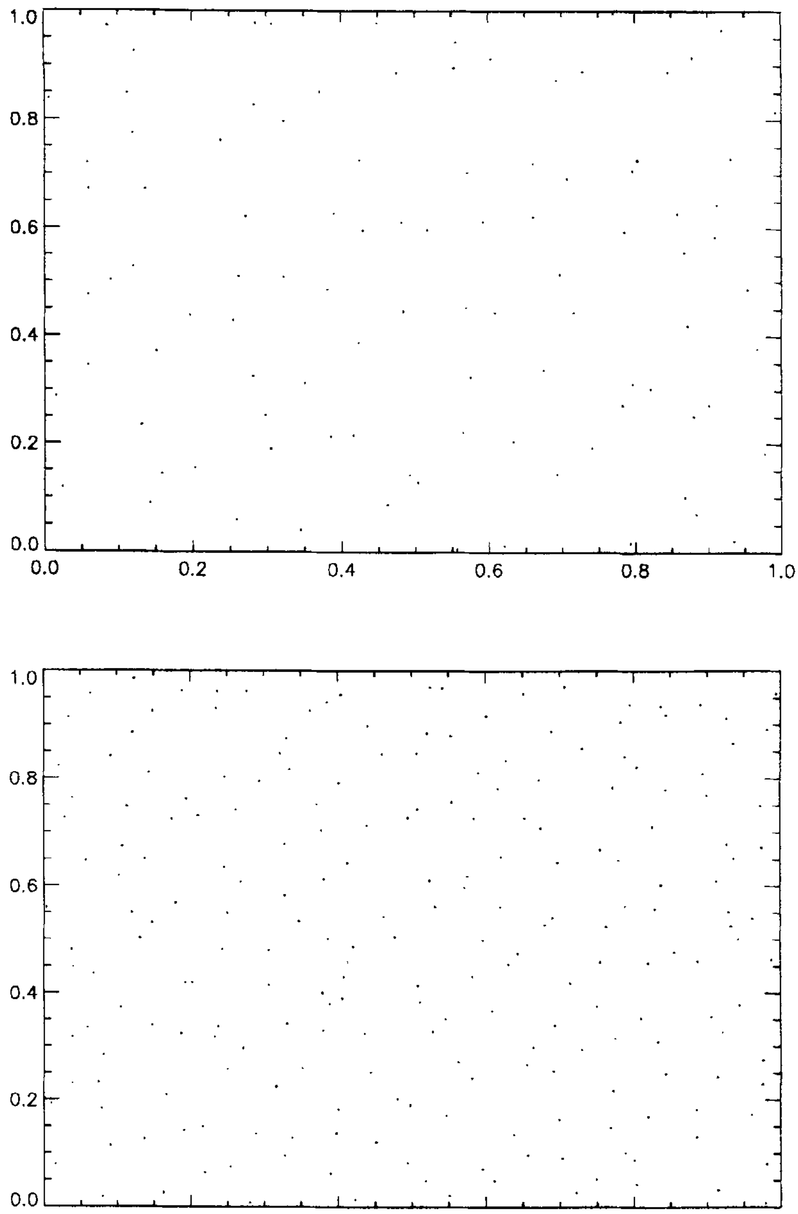


Figure 2. The scattered data: top $n = 100$, bottom $n = 225$

5. Numerical Results

In the numerical experiments we have considered n scattered points $\mathbf{r}_i$ $i = 1,\dots,n$ in the unitary square of $\mathbb{R}^2$ and the corresponding functional values distorted by noise

$$z_i = f(\mathbf{r}_i) + e(\mathbf{r}_i), \qquad i = 1, \ldots, n.$$

The scattered points $\mathbf{r}_i$, are displayed in Fig. 2.

The algorithm has been successfully tested on several functions with different

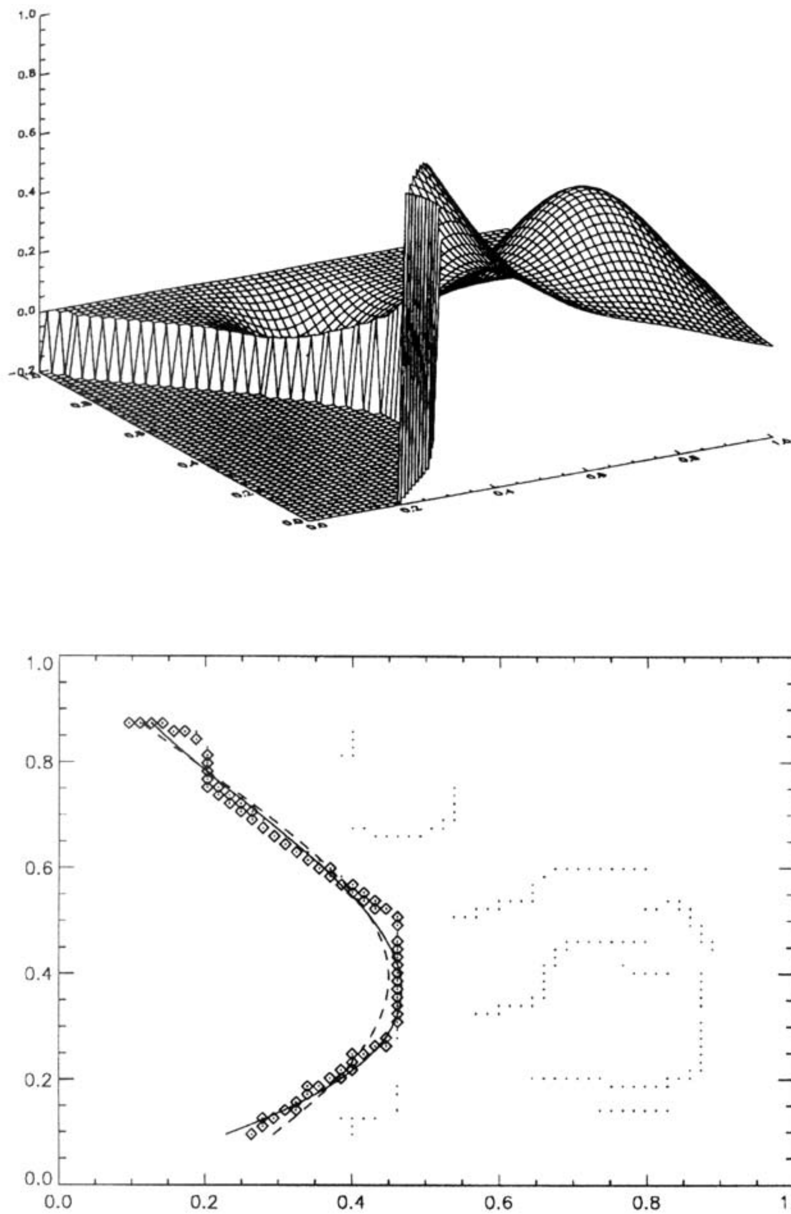


Figure 3. Example 1. Top: the exact function, bottom: the algorithm results for $a = 0.03$

kinds of discontinuity curves, different sample dimension n and different standard deviation σ.

For all the examples presented here, we have fixed the significance level δ equal to 0.1 and $N = 100$. We have chosen $\theta(x, y) = \exp(-x^2/2) \exp(-y^2/2)$ and, accordingly with Corollary 2.3, $\bar{A} = 4$.

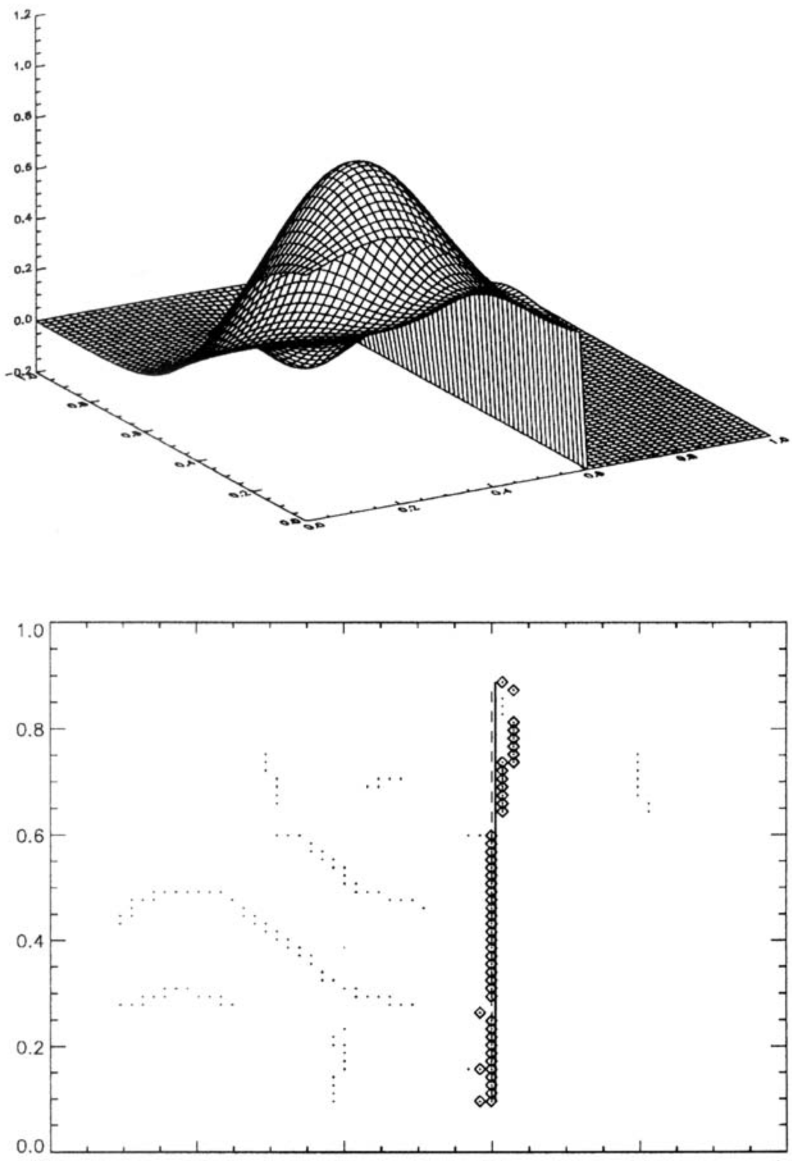


Figure 4. Example 2. Top: the exact function, bottom: the algorithm results for $a = 0.04$

The results achieved are shown in Figs. 3, 4, 6 and 7. On the top is shown the exact function, on the bottom the algorithm results: the dots are the wavelet transform modulus maxima, the diamonds are the points detected as discontinuities.

In these examples the curve ℓ (dashed line) has been reconstructed using the least squares method (continuous line). Generally, the chosen method for the reconstruction depends on the informations we have on the curve itself.

Example 1.

$$f_1(x,y) = \begin{cases} f_r(x,y) & \text{if } x \geq 1/4\sin(4y) + 0.2 \\ -0.2 & \text{otherwise} \end{cases}$$

where $f_r(x,y)$ is the well-known Franke's function [5]. We have considered $n = 225$ and $\sigma = 0.01$. The trend is discontinuous across the curve of equation $x = 1/4\sin(4y) + 0.2$. On the bottom of Fig. 3, the discontinuity points detected by the algorithm for $a = 0.03$ are shown together with the approximation of the curve.

Example 2.

$$f_x(x,y) = \begin{cases} f_r(x,y) & \text{if } x \leq 0.6 \\ -0.2 & \text{otherwise} \end{cases}$$

We have considered $n = 100$ and $\sigma = 0.05$. The trend is discontinuous across $x = 0.6$ which, as shown on the bottom of Fig. 4, is recovered by the algorithm having chosen $a = 0.04$.

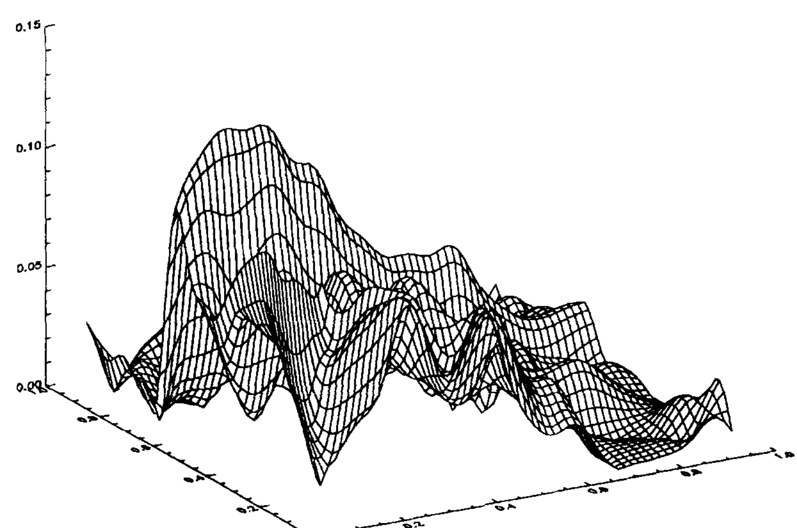


Figure 5. $SW_{g,a}$ of Example 2

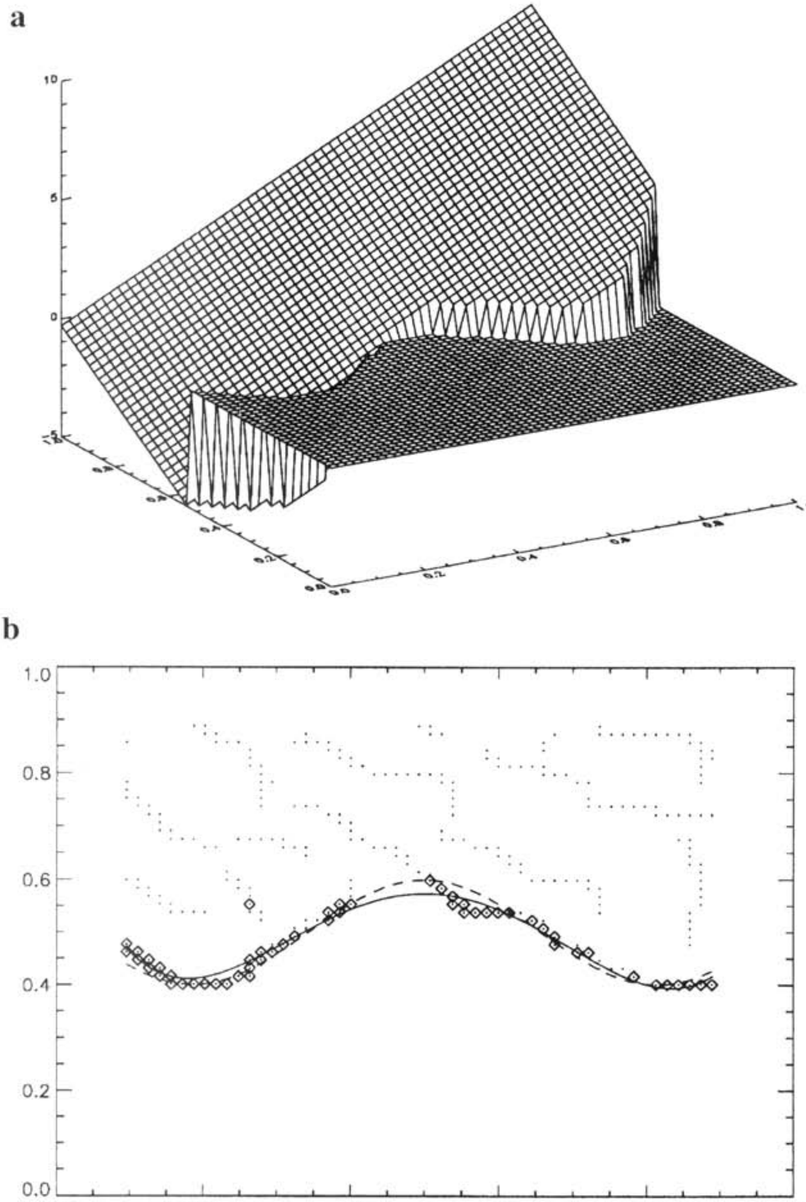
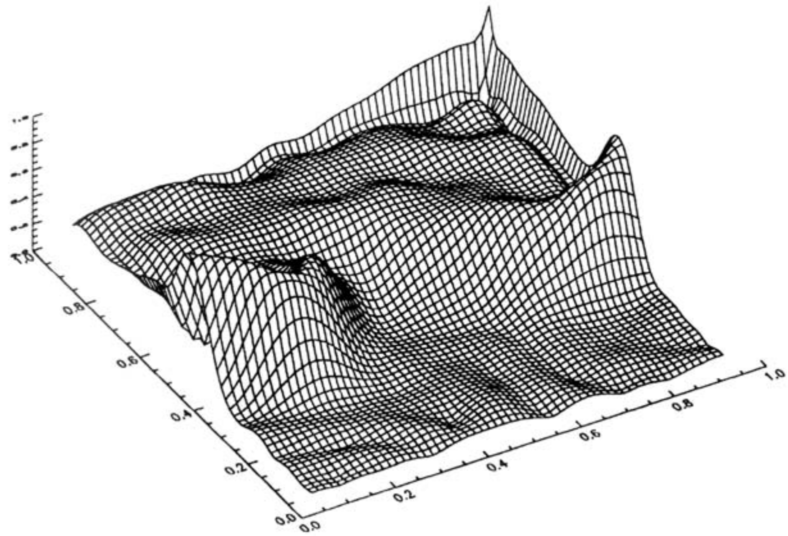


Figure 6. (*Continued*)

Remarks. We may observe that in both the examples the surface has two extrema near to the discontinuity curves. As shown by Fig. 5, this closeness makes the behaviour of the modulus of the wavelet transform unclear. By consequence it is difficult to detect ℓ by analysing only the wavelet transform. Moreover the smooth variations due to the extrema are detected by the modulus c



d

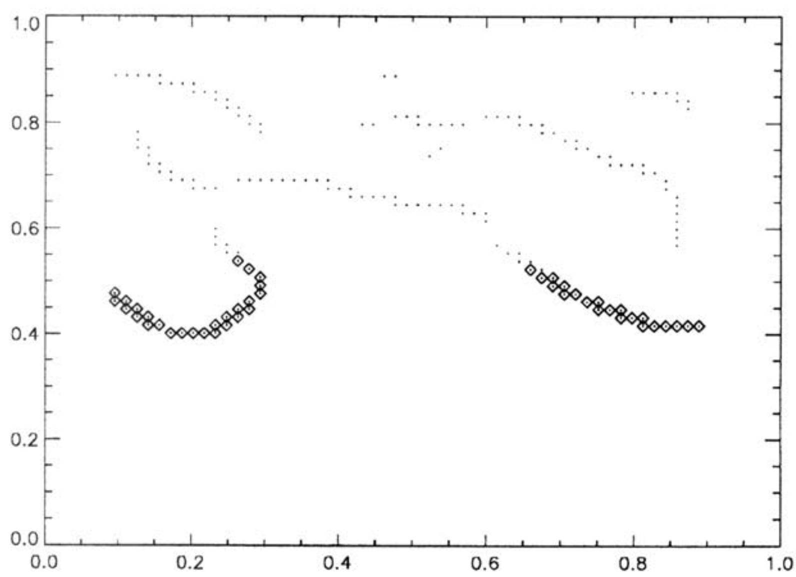


Figure 6. Example 3. **a** The exact function, **b** the algorithm results for $a = 0.03$, **c** $SW_{g,a}$ with $a = 0.05$, **d** the modulus maxima and the detected points for $a = 0.05$

maxima (see Figs. 3 and 4: bottom) and, without the detection test, they could be confused with irregular behaviours.

Example 3.

$$f_3(x, y) = \begin{cases} 10(x + y) - 10.3 & \text{if } y \geq 0.1\cos(10x - 5) + 1/2 \\ 0 & \text{otherwise} \end{cases}$$

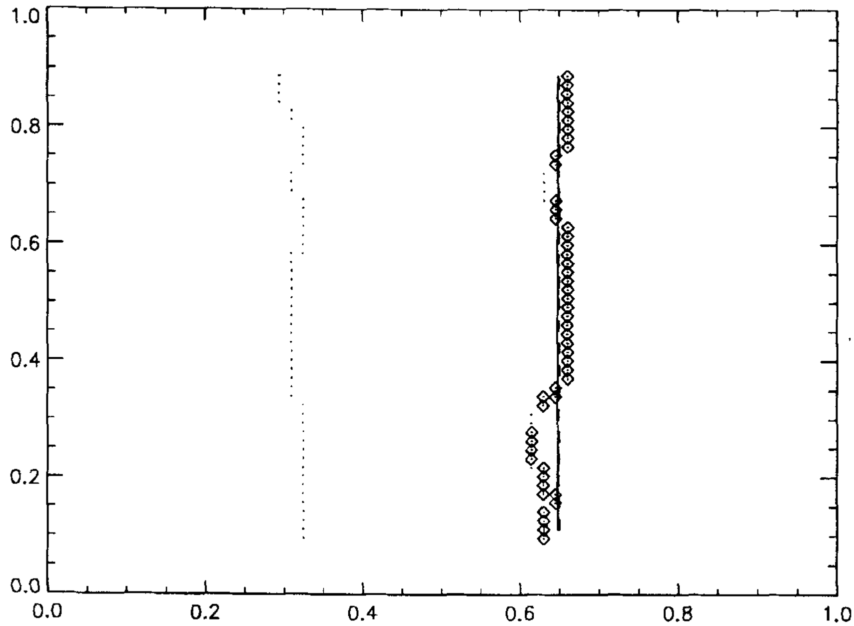


Figure 7. Example 4. The algorithm results for $a = 0.05$

We have considered $n = 225$ and $\sigma = 0.2$. Picture b of Fig. 6 shows the results achieved for $a = 0.03$.

Remarks. Across the curve ℓ: $y = 0.1\cos(10x - 5) + 1/2$, the jumps of discontinuity go to zero. In this case, the points corresponding to small discontinuity jumps could be confused with noise fluctuations. But, as shown by Picture b of Fig. 6, our method gives good results and detects these critical points too. Moreover when we increase the scale a, the number of maxima due to the noise decreases but at the same time we loose important information on the underlying function. In fact (see Figs. 6 c,d) the wavelet transform modulus and its modulus maxima do not locate anymore the discontinuity curve and we may also observe how the maxima, due to noise, seem to vary smoothly along some lines of Q.

Example 4. Finally let us consider the example discussed in Section 2.

$$f_4(x,y) = \begin{cases} \arctan(8(3x-1)+0.2) & \text{if } x \geq 0.65 \\ -0.1 & \text{otherwise} \end{cases}$$

We have considered $n = 225$ and $\sigma = 0.05$. As shown in Fig. 7, the algorithm discriminates, for $a = 0.05$ the discontinuity curve from the curve across which f varies sharply.

References

[1] Bendat, J. S.: Principles and applications of random noise theory. New York: Wiley 1958.

[2] Bozzini, M., De Tisi, F., Rossini, M.: Irregularity detection from noisy data with wavelets. In: Wavelets, images, and surface fitting (Laurent, P. J., Le Méhauté, Schumaker, L. L., eds.), pp. 59–66. Wellesley: A.K. Peters 1994.
[3] Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Machine Intell. *8*, 679–698 (1986).
[4] Daubechies, I.: The wavelet transform, time-frequency localization and signal analysis. IEEE Trans. Inf. Theory *36*, 961–1005 (1990).
[5] Franke, R., Nielson, G. M.: Smooth interpolation of large set of scattered data. Int. J. Numer. Methods Eng. *38*, 1691–1704 (1980).
[6] Mallat, S., Hwang, W. L.: Singularity detection and processing with wavelets. IEEE Trans. Inf. Theory *38*, 617–643 (1992).
[7] Mallat, S., Zhong, S.: Characterization of signals from multiscale edges. IEEE Trans. Pattern Anal. Machine Int. PAMI *14*, 710—732 (1992).
[8] Rossini, M.: Irregularity detection from noisy data in one and two dimension. Numer. Alg. *16*, 283–301 (1997).

M. Rossini
Dipartimento di Matematica
Università di Milano
via Saldini 50
I-20133 Milano
Italy
e-mail: rossini@vmimat.mat.unimi.it