Significant edges in the case of a non-stationary Gaussian noise

Isabelle Abraham, Romain Abraham, Agnes Desolneux, Sebastien Li-Thiao-Te

Significant edges in the case of a non-stationary Gaussian noise

I. Abraham [a], R. Abraham [b], A. Desolneux [c], S. Li-Thiao-Te [d]

[a] *CEA/DIF, 91680 Bruyères le Chatel, France*

[b] *Laboratoire MAPMO, Fédération Denis Poisson, Université d'Orléans, B.P. 6759, 45067 Orléans cedex 2, France*

[c] *Laboratoire MAP5, Université René Descartes, 45 rue des Saints-Pères, 75270 Paris cedex 06, France*

[d] *Laboratoire CMLA, ENS Cachan, 61 avenue du Président Wilson, 94235 Cachan cedex, France*

Abstract

In this paper, we propose an edge detection technique based on some local smoothing of the image followed by a statistical hypothesis testing on the gradient. An edge point being defined as a zero-crossing of the Laplacian, it is said to be a significant edge point if the gradient at this point is larger than a threshold $s(\varepsilon)$ defined by: if the image I is pure noise, then $\mathbb{P}(\|\nabla I\| \geq s(\varepsilon) \mid \Delta I = 0) \leq \varepsilon$. In other words, a significant edge is an edge which has a very low probability to be there because of noise. We will show that the threshold $s(\varepsilon)$ can be explicitly computed in the case of a stationary Gaussian noise. In images we are interested in, which are obtained by tomographic reconstruction from a radiograph, this method fails since the Gaussian noise is not stationary anymore. But in this case again, we will be able to give the law of the gradient conditionally on the zero-crossing of the Laplacian, and thus compute the threshold $s(\varepsilon)$. We will end this paper with some experiments and compare the results with the ones obtained with some other methods of edge detection.

Key words: Edge detection, Significant edges, Inverse problem, Statistical hypothesis testing

Email addresses: isabelle.abraham@cea.fr (I. Abraham), romain.abraham@univ-orleans.fr (R. Abraham),

1 Introduction

This work is part of some specific physical experiments which consist in studying radially symmetric objects [3]. These objects are composed of several materials and one of the interesting features is the location of the frontier between the different materials.

To describe such an object, it is enough to give the densities of the materials on a slice of the object that contains the symmetry axis. An example of studied object is given on Figure 1.

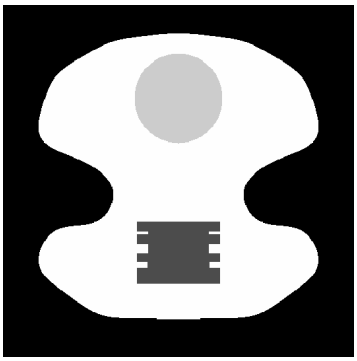
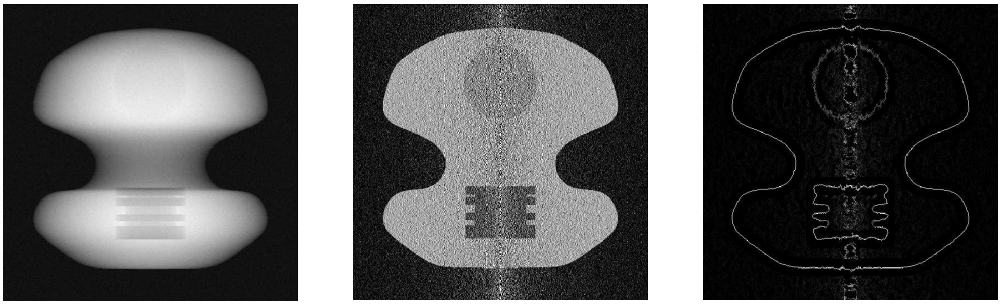


Fig. 1. Slice of a studied object.

To look at the interior of this object, a radiography is performed (see Figure 2(a)), then a tomographic reconstruction is computed (Figure 2(b)) and finally an edge detection is made (Figure 2(c)).

agnes.desolneux@math-info.univ-paris5.fr (A. Desolneux),
lithiaote@cmla.ens-cachan.fr (S. Li-Thiao-Te).



(a) (b) (c)

Fig. 2. (a) Radiograph of the object of Figure 1. (b) Tomographic reconstruction. (c) Edge detection on the tomography.

At this point, let us mention that the tomographic reconstruction is not an usual one. Indeed, the usual inverse Radon transform (and the usual reconstruction algorithms such as filtered back-projection) operates on a slice of the object that is orthogonal to the rotation axis. Here, because of the radial symmetry assumption, the reconstruction can be global [5]. This reconstruction will be detailed in Section 4.1.

As we can see on Figure 2(c), many detected edges do not correspond to real features. This is due to the high level of noise. For the time being, the selection of the edges is manually executed. The goal of this work is to perform this selection automatically. For that purpose, the edge detector will not be changed but we will compute also other significant features that will allow us to select the "true" edges.

The ideas used here mainly come from previous work of Desolneux, Moisan and Morel [4]. Informally speaking, they define the notion of significant edges by computing the probability that some edge-related events appear in an image of pure noise. When this probability is small enough, the edge is probably a feature of the image and not due to the noise. Unfortunately, their method assumes that the noise is stationary which, as easily seen on Image 2(b), is not the case in our study because of the tomographic inversion (see Section 4.5.1 for some examples of results obtained with their method). Moreover, their study is quite general and, apart from the stationarity, no assumption is made on the noise.

In our case, as we deal with specific images, the noise is well-known and some statistical models can be used. Indeed, we may suppose that the noise on the radiograph (2(a)) is a Gaussian white noise with mean zero and with a variance that can easily be estimated. Then, a tomographic inversion is performed. As this operation is linear, we still obtain a Gaussian noise but it is now correlated

and non-stationary.

The edge detector will not be modified here. It consists in estimating the Laplacian at each point, and edge points are then defined as the zero-crossings of the Laplacian. As we already said, we only add some features that will discriminate the significant edges. The easiest feature to compute is a kind of contrast measurement C based on a gradient estimate. To be more precise, we consider an image I of pure noise (that is a realization of our model of noise after tomographic reconstruction), estimate the gradient and the Laplacian of I at a point (u, v) (with an abuse of notation, we will denote by $\nabla I(u, v)$ and $\Delta I(u, v)$ these estimates and by $C(u, v)$ the contrast value) and we compute, for a fixed $\varepsilon > 0$, the smallest value $s(\varepsilon)$ for which

$$\mathbb{P}(C(u, v) \geq s(\varepsilon) \,\Big|\, \Delta I(u, v) = 0) \leq \varepsilon. \tag{1}$$

Then, we perform an edge detection on the studied image f (where we also estimate ∇f and Δf by the same method) and we keep the points (u, v) of the studied image f that satisfy

- $\Delta f(u, v) = 0$ (an edge is present at point (u, v)).
- $C(u, v) \geq s(\varepsilon)$ (this edge is significant).

¿From a statistical point of view, this consists in performing an hypothesis test. We consider a point (u, v) where an edge takes place ($\Delta(u, v) = 0$) and we test the null hypothesis "the zero-crossing of the Laplacian is due to the noise". The level of the test ε is arbitrarily chosen and related to the number of false detections allowed. It will be set to $\varepsilon = 10^{-5}$ hereafter. Let us mention that the threshold value $s(\varepsilon)$ varies slowly with respect to ε. For instance, in the case of a white noise (see Section 3), the threshold value can be computed explicitly and is proportional to $\sqrt{-\ln \varepsilon}$. When the null hypothesis is rejected, the edge is retained as it comes from a "true" feature of the image, whereas when the null hypothesis is accepted, the zero-crossing of the Laplacian may come from the noise and the edge is not meaningful.

Let us mention at this point that such statistical approaches have already been used for edge detection in [11], [10] or [8]. They usually use estimates of the gradient based on finite differences which fail in our case. Moreover, the noise is in most cases stationary. Let us also cite [2] where the authors have modified the method of [4] to take into account the non-stationarity of some images, by a local noise estimate. Their work is still general and does not make any assumption on the noise structure. As we deal with specific experiments, the noise is always the same and well-known and we can take proper advantage of this knowledge.

The paper is organized as follows: in Section 2, we present the edge detector based on the estimate of the gradient and the Laplacian. Then, in Section

3, our method is presented in the case of a Gaussian white noise. Of course, this does not correspond to our case but the computations are easier and show the performance of this method. In Section 4, we will first describe the tomographic inversion and the operators involved, and then describe the noise model we have to deal with. We will then apply the significant edges detection method in the framework of this non-stationary noise. We will end the section with some experiments and comparisons with other methods.

2 Estimating the Gradient and the Laplacian

In this section, we introduce a method for edge detection. We consider that the image is a real-valued function $(u, v) \mapsto f(u, v)$ of two continuous real parameters u and v. Then, we say that there exists an edge at point (u, v) if the Laplacian of f is zero at this point. Moreover, the computation of the contrast function C will be based on the gradient of f (see the end of this section for the choice of this function). As the images are very noisy, these derivatives cannot be estimated by usual finite differences. The method used here, sometimes known as Savitsky-Golay smoothing, consists in locally approximating the image by a polynomial. The derivatives of the polynomial are then identified with those of the image.

2.1 An optimization problem

Let (u, v) denote the point where we want to compute the first and second order derivatives of the image f. We choose 2 parameters : d which is the maximum degree of the approximating polynomial and r which is the radius of the ball on which we perform the approximation. We denote by $B_r(u, v)$ the ball of radius r centered at point (u, v). We will simply write B_r when the center of the ball is the origin $(0, 0)$ of $\mathbb{R}^2$. We are then looking for a polynomial P of degree less that d such that

$$E(P) = \int_{B_r} \Big(f(u + x, v + y) - P(x, y) \Big)^2 \, dx \, dy \tag{2}$$

is minimal among all polynomials of degree less than d. In other words, we are looking for the best approximation of f by a polynomial of degree less than d on the ball $B_r(u, v)$ in the sense of the L^2-norm.
This is an optimization problem where the unknowns are the coefficients of the polynomial. As the problem is convex, there is a unique solution (given by the orthogonal projection of f on the space of polynomials of degree less

than d) which is easily computed by solving the equations

$$\frac{\partial E}{\partial a_i} = 0$$

where the a_i's denote the coefficients of the polynomial.

Role of the ball radius. Two parameters are arbitrary chosen in this method. The first one is the ball radius r. The larger r is, the more effective the smoothing is. The influence of the noise is therefore attenuated with a large r but the location of the edge is then less precise. We must consequently make a balance between noise smoothing and edge detection accuracy. For instance, if we have a small level of noise or if the edges are very complicated (with high curvature), we must choose a small value for r.

Role of the polynomial degree. The second parameter is the polynomial degree. Here again a large value of d gives a better approximation but does not smooth the noise enough. In fact, as we are, in a first step, interested in the points where the Laplacian is zero, it appears that a second-order polynomial is enough. Of course, the estimate of the first order derivatives with a polynomial of degree 2 is not very good and highly depends on the size of the window B_r. But we will see that this drawback can be useful for the choice of a contrast function.

In what follows, the approximation is made with a polynomial of degree $d = 2$, and the first and second order derivatives of the image are identified with those of the approximating polynomial.

2.2 Computations with a second order polynomial

Let us first introduce some notations. In the following, we will set

$$\forall i, j \in \mathbb{N}, \quad b_{ij}(r) = \int_{B_r} x^i y^j dx\, dy.$$

As the ball B_r is symmetric, we have that $b_{ij}(r) = 0$ as soon as i or j is odd and that $b_{ij}(r) = b_{ji}(r)$ for all i,j. In order to have simple expressions, we will also set:

$$b(r) = b_{20}(r), \;\; \alpha(r) = -\frac{2b_{20}(r)}{b_{00}(r)} \quad \text{and} \quad \beta(r) = \frac{1}{2}\left(b_{40}(r) + b_{22}(r) - \frac{2b_{20}^2(r)}{b_{00}(r)}\right)$$

Lemma 1 *The gradient and the Laplacian of the polynomial of degree* 2 *which is the best approximation of f on the ball $B_r(u, v)$ for the L^2-norm, being respectively denoted by $\nabla_r f(u, v) = (\frac{\partial_r f}{\partial x}(u, v), \frac{\partial_r f}{\partial y}(u, v))$ and $\Delta_r f(u, v)$, are*

given by:

$$\begin{aligned}
\frac{\partial_r f}{\partial x}(u,v) &= \frac{1}{b(r)} \int_{B_r} x\, f(u+x, v+y)\, dx\, dy \\
\frac{\partial_r f}{\partial y}(u,v) &= \frac{1}{b(r)} \int_{B_r} y\, f(u+x, v+y)\, dx\, dy \\
\Delta_r f(u,v) &= \frac{1}{\beta(r)} \int_{B_r} f(u-x, v-y) \left(\alpha(r) + x^2 + y^2\right)\, dx\, dy.
\end{aligned}$$

Proof :

We consider a polynomial of degree 2 which we write

$$P(x,y) = a_{xx}x^2 + a_{yy}y^2 + a_{xy}xy + a_x x + a_y y + a_0.$$

The equations obtained by writing $\nabla E(P) = 0$, where $E(P)$ is given by Equation (2), are:

$$\begin{cases}
b_{40}(r)a_{xx} + b_{22}(r)a_{yy} + b_{20}(r)a_0 = \displaystyle\int_{B_r} x^2\, f(u+x, v+y)\, dx\, dy \\
b_{22}(r)a_{xx} + b_{40}(r)a_{yy} + b_{20}(r)a_0 = \displaystyle\int_{B_r} y^2\, f(u+x, v+y)\, dx\, dy \\
b_{22}(r)a_{xy} = \displaystyle\int_{B_r} x\, y\, f(u+x, y+v)\, dx\, dy \\
b_{20}(r)a_x = \displaystyle\int_{B_r} x\, f(u+x, y+v)\, dx\, dy \\
b_{20}(r)a_y = \displaystyle\int_{B_r} y\, f(u+x, v+y)\, dx\, dy \\
b_{20}(r)a_{xx} + b_{20}(r)a_{yy} + b_{00}(r)a_0 = \displaystyle\int_{B_r} f(u+x, v+y)\, dx\, dy
\end{cases}$$

We then obtain the following estimates for the derivatives:

$$\begin{aligned}
\frac{\partial P}{\partial x}(0,0) = a_x &= \frac{1}{b_{20}(r)} \int_{B_r} x\, f(u+x, v+y)\, dx\, dy \\
\frac{\partial P}{\partial y}(0,0) = a_y &= \frac{1}{b_{20}(r)} \int_{B_r} y\, f(u+x, v+y)\, dx\, dy \\
\Delta P(0,0) &= 2(a_{xx} + a_{yy}) \\
&= \frac{2}{b_{40} + b_{22} - \frac{2b_{20}^2}{b_{00}}} \int_{B_r} f(u+x, v+y) \left(-\frac{2b_{20}}{b_{00}} + x^2 + y^2\right)\, dx\, dy.
\end{aligned}$$

□

2.3 Choice of the contrast function

We would like to use a contrast function based on the estimates of the first and second derivatives of the image f obtained in the previous section.

The simplest contrast function we can choose is the norm of the gradient:

$$C_1(u,v) = \|\nabla_r f(u,v)\| .$$

Indeed, the value of this norm tells how sharp the edge is. This contrast function is efficient and will be used when the images we deal with are piecewise constant.

However, in many cases, the objects we handle are not homogeneous and their images contain some slopes (see Figure 3). In this case, the gradient norm is not a good contrast function. Indeed, let us consider an image with a constant slope with some noise (see Figure 4). We would like to say that no edge is significant in that case. However, the value of the gradient norm (which will be close to the value of the slope) will always be greater that the threshold value s when the noise level is small.

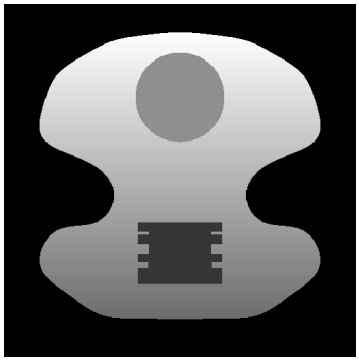


Fig. 3. Object with an inhomogeneous material

In the latter case, we take advantage of the dependence of the first order derivatives estimates with respect to the ball radius. Indeed, the estimates of the gradient in the case of the constant slope in Figure 4 will not depend on the size of the window (see Figure 4) whereas, when an edge (a discontinuity) occurs, the estimates do depend on that radius (see Figure 5). So, we can use as a contrast function the function

$$C_2(u,v) = \|\nabla_{r_1} f(u,v) - \nabla_{r_2} f(u,v)\|$$

where $r_1 < r_2$ and $\nabla_r f$ denotes the value of the gradient estimate with a ball

of radius r.

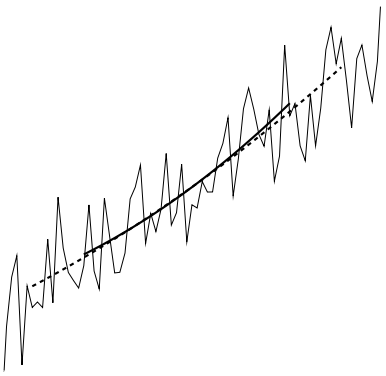


Fig. 4. A noisy constant slope: the gradient of the approximating polynomial does not depend on the radius r.

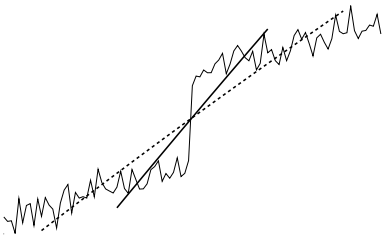


Fig. 5. An edge on a noisy slope: the approximating polynomials with two different values of the radius r.

3 Significant edges in the case of a Gaussian white noise

3.1 White noise and Wiener integral

We recall here the definition and the main properties of a white noise in a continuous setting and of the Wiener integral. We refer to [12], [6] or [7] for more on white noise and the Wiener integral.

Definition 1 *A Gaussian white noise on $\mathbb{R}^2$ of variance σ^2 is a random function W defined on the Borel sets A of $\mathbb{R}^2$ of finite Lebesgue measure (denoted by $|A|$) such that*

- *$W(A)$ is a Gaussian random variable (r.v.) with mean 0 and variance $\sigma^2|A|$,*
- *If $A_1 \cap A_2 = \emptyset$, the r.v. $W(A_1)$ and $W(A_2)$ are independent and*

$$W(A_1 \cup A_2) = W(A_1) + W(A_2).$$

Such a function W exists but is not a true measure since the two-parameters process

$$B(s,t) := W\Big((0,s] \times (0,t]\Big)$$

(usually called the Brownian sheet) is of unbounded total variation.

Nevertheless we can define the so-called Wiener integral $\int f dW$ for every function f in $L^2(\mathbb{R}_+^2)$. We can also define the derivatives of the Brownian sheet in the sense of Schwartz distributions (although the Brownian sheet is nowhere differentiable). Thus, we define

$$\dot{B}(s,t) = \frac{\partial^2 B(s,t)}{\partial s \partial t}$$

and we have

$$\int f dW = \int_{\mathbb{R}_+^2} f(u,v)\dot{B}(u,v)du\,dv \qquad a.s.$$

for every function f in the Schwartz space.

With a slight abuse of notations, we call $\dot{B}$ a Gaussian white noise and we always denote by $\int_{\mathbb{R}_+^2} f(u,v)\dot{B}(u,v)du\,dv$ the Wiener integral with respect to this white noise, for every function $f \in L^2$. The main properties of this integral are

- For every f, the r.v. $\displaystyle\int_{\mathbb{R}_+^2} f(u,v)\dot{B}(u,v)du\,dv$ is a Gaussian r.v. with mean 0 and variance $\sigma^2 \displaystyle\int_{\mathbb{R}_+^2} f(u,v)^2 du\,dv$.
- For every f, g, the random vector

$$\left(\int_{\mathbb{R}^2} f(u,v)\dot{B}(u,v)du\,dv, \int_{\mathbb{R}^2} g(u,v)\dot{B}(u,v)du\,dv\right)$$

  is Gaussian with covariance

$$\sigma^2 \int_{\mathbb{R}^2} f(u,v)g(u,v)du\,dv.$$

We will use these properties to compute the laws of ∇I and ΔI.

3.2 Laws of the gradient and of the Laplacian

We suppose here that our noise is a Gaussian white noise, of variance σ^2. As we have already said, this case is not the one we are interested in and our method is probably over-performed by other standard methods in that case. The goal of this section is to present our method in a simple case where the computations are easy to do and can be carried out in a continuous setting.

We will only focus here on the case of piecewise constant objects and therefore we will use the contrast function C_1.

Lemma 2 *If the image I is a Gaussian white noise of variance σ^2, then*

$$\left(\frac{\partial_r I}{\partial x}, \frac{\partial_r I}{\partial y}, \Delta_r I\right)$$

is a Gaussian vector with mean zero and covariance matrix

$$\begin{pmatrix} \frac{\sigma^2}{b(r)} & 0 & 0 \\ 0 & \frac{\sigma^2}{b(r)} & 0 \\ 0 & 0 & V(r,\sigma) \end{pmatrix}, \quad \textit{where} \quad V(r,\sigma) = \frac{\sigma^2}{\beta^2(r)} \int_{B_r} \left(\alpha(r) + x^2 + y^2\right)^2 dx\, dy.$$

Proof : We compute the laws of the approximate derivatives of I when $I = \dot{B}$. We recall that these derivatives are given by

$$\begin{aligned}
\frac{\partial_r I}{\partial x}(u,v) &= \frac{1}{b(r)} \int_{B_r} x\, \dot{B}(u+x, v+y)\, dx\, dy \\
\frac{\partial_r I}{\partial y}(u,v) &= \frac{1}{b(r)} \int_{B_r} y\, \dot{B}(u+x, v+y)\, dx\, dy \\
\Delta_r I(u,v) &= \frac{1}{\beta(r)} \int_{B_r} \dot{B}(u+x, v+y) \left(\alpha(r) + x^2 + y^2\right) dx\, dy.
\end{aligned}$$

Because of the stationarity of $\dot{B}$, they have the same law as

$$\begin{aligned}
\frac{\partial_r I}{\partial x}(0,0) &= \frac{1}{b(r)} \int_{B_r} x\, \dot{B}(x,y)\, dx\, dy \\
\frac{\partial_r I}{\partial y}(0,0) &= \frac{1}{b(r)} \int_{B_r} y\, \dot{B}(x,y)\, dx\, dy \\
\Delta_r I(0,0) &= \frac{1}{\beta(r)} \int_{B_r} \dot{B}(x,y) \left(\alpha(r) + x^2 + y^2\right) dx\, dy.
\end{aligned}$$

As we deal with Wiener integrals, we deduce that the vector

$$\left(\frac{\partial_r I}{\partial x}, \frac{\partial_r I}{\partial y}, \Delta_r I\right)$$

is a Gaussian vector with mean zero.

To compute its covariance matrix, let us recall that, if X and Y are random variables defined by

$$X = \int_{B_r} h_1(x,y)\dot{B}(x,y)\,dx\,dy$$
$$Y = \int_{B_r} h_2(x,y)\dot{B}(x,y)\,dx\,dy$$

then we have

$$\mathrm{Cov}(X,Y) = \sigma^2 \int_{B_r} h_1(x,y)h_2(x,y)\,dx\,dy.$$

Consequently, we have for instance:

$$\mathrm{Cov}\left(\frac{\partial_r I}{\partial x}, \frac{\partial_r I}{\partial y}\right) = \frac{\sigma^2}{b^2(r)} \int_{B_r} x\,y\,dx\,dy = 0.$$

By some analogous calculations, we finally get the following covariance matrix for our Gaussian vector:

$$\begin{pmatrix} \frac{\sigma^2}{b(r)} & 0 & 0 \\ 0 & \frac{\sigma^2}{b(r)} & 0 \\ 0 & 0 & V(r,\sigma) \end{pmatrix}$$

where

$$V(r,\sigma) = \frac{\sigma^2}{\beta^2(r)} \int_{B_r} \left(\alpha(r) + x^2 + y^2\right)^2 dx\,dy.$$

□

Thanks to this lemma, we immediately have the following properties:

- The random variable $\|\nabla_r I\|^2$ is the sum of two squared independent Gaussian random variables which have the same variance. It is therefore distributed as a χ^2-law. More precisely, its law is

$$\frac{\sigma^2}{b(r)}\chi^2(2)$$

  where $\chi^2(2)$ denotes a χ^2-law with two degrees of freedom.
- The random variable $\Delta_r I$ is a Gaussian random variable with mean zero and variance $V(r,\sigma)$.
- The random variables $\|\nabla_r I\|^2$ and $\Delta_r I$ are independent.

3.3 Computation of the threshold

Proposition 1 *Let I be a Gaussian white noise and let $s(\varepsilon)$ be the threshold value such that*

$$\mathbb{P}\Big(\|\nabla_r I\| \geq s(\varepsilon) \;\Big|\; \Delta_r I = 0\Big) \leq \varepsilon.$$

Then

$$s(\varepsilon) = \sqrt{-\frac{2\sigma^2}{b(r)} \ln \varepsilon}.$$

Proof : To begin with, as the random variables $\|\nabla_r I\|^2$ and $\Delta_r I$ are independent, we can forget the conditioning and only compute

$$\mathbb{P}\Big(\|\nabla_r I\| \geq s(\varepsilon)\Big) = \mathbb{P}\Big(\|\nabla_r I\|^2 \geq s(\varepsilon)^2\Big).$$

As a consequence of Lemma 2, we have that the law of $\|\nabla_r I\|^2$ is $\frac{\sigma^2}{b(r)}\chi^2(2)$. Now, since the density of a $\chi^2(2)$ law is the one of a $\Gamma\left(\frac{1}{2}, 1\right)$ law, we have that the law of $\|\nabla_r I\|^2$ is given by

$$\mathbb{P}\Big(\|\nabla_r I\|^2 \geq s^2\Big) = \int_{\frac{b(r)}{\sigma^2}s^2}^{+\infty} \frac{1}{2} e^{-\frac{t}{2}} dt = \exp\left(-\frac{b(r)s^2}{2\sigma^2}\right).$$

This finally leads to the announced threshold value $s(\varepsilon)$. □

3.4 Experiments

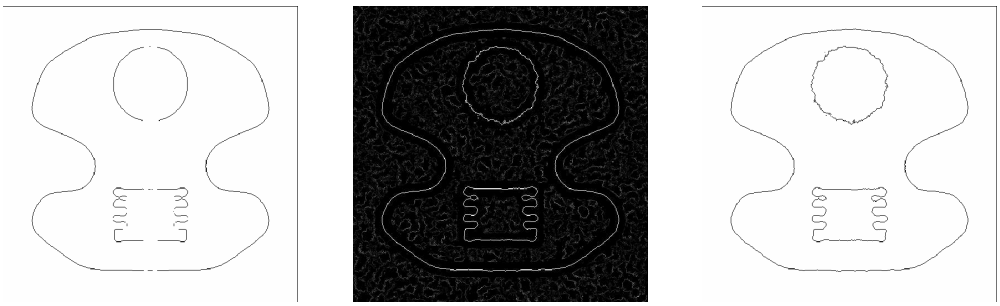
We consider the piecewise constant object of Figure 1 with some additive Gaussian white noise. The densities of the different materials of this object are:

- 1 for the outer material,
- 0.8 for the material inside the circle,
- 0.3 for the other inner material.

The standard deviation of the Gaussian noise is $\sigma = 0.2$ in the experiments of Figure 6 and is $\sigma = 0.4$ in the experiments of Figure 7. Both images have the same size 512×512 pixels. The experiments have been carried out with a ball of radius $r = 12$ pixels.
The different images of Figures 6 and 7 are respectively:

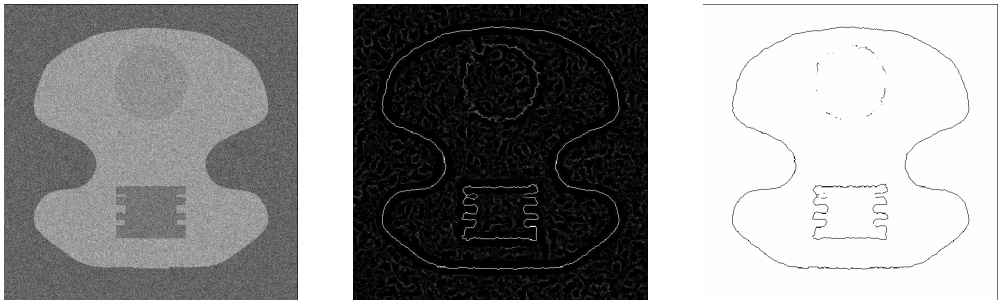
- (a) The noisy image.

- (b) The zero-crossings of the Laplacian with the contrast function C_1 visualized in grey-level (the white color corresponds to high value for the contrast function C_1).
- (c) The extracted significant edges ($\varepsilon = 10^{-5}$).



(a) (b) (c)

Fig. 6. (a) The noisy image ($\sigma = 0.2$). (b) The zero-crossings of the Laplacian with the contrast function C_1 visualized in grey-level. (c) The extracted significant edges ($\varepsilon = 10^{-5}$).



(a) (b) (c)

Fig. 7. (a) The noisy image ($\sigma = 0.4$). (b) The zero-crossings of the Laplacian with the contrast function C_1 visualized in grey-level. (c) The extracted significant edges ($\varepsilon = 10^{-5}$).

In the case of a signal-to-noise ratio large enough (Figure 6), all the edges are well detected and the "false" edges are removed. Let us nevertheless mention that, with our method, the edges which have a high curvature are smoothed. This drawback is even more important when the ball radius r is large (the influence of the value of this radius will be studied in the experiments of the next section).

When the noise level is rather large (Figure 7), some edges of the image cannot be extracted from the noise (it happens when the contrast associated to this edge is close to the noise level).

4 Significant edges in the case of a Gaussian white noise on the radiograph

4.1 Tomography

Let us turn now to the more realistic case we are interested in. As we mentioned it in the introduction, we first make a radiography of an object. Tomography is the inverse problem associated with reconstructing the initial object from its radiograph. This is now a well-known problem as it is the key tool in medical scanner imagery (or other medical imaging systems).

To begin with, let us describe what a radiography is from a mathematical point of view. The studied object is exposed to X-rays that go through it. Some of the X-photons are absorbed. As an output, we observe the quantity of X-photons that have not been absorbed by the material, and we thus measure in some sense the "mass" of material the ray went through. More precisely, if the object is described by its density μ (which is a function of the space coordinates), what can be measured at some point of the receptor is

$$\int_{ray} \mu \, d\ell$$

where "ray" means the straight line that goes from the source to the studied point of the receptor (we suppose that the X-rays source is just a point, which implies that the previous line is unique).

We also assume that the X-rays source is far away from the object so that the rays are assumed to be parallel. Then, to reconstruct any object from its radiographs, we must turn around the object and make a radiography for every angle $\theta \in [0, \pi)$. This leads to the so-called Radon transform of the object, which is known to be invertible. This is the principle of the medical scanner.

In our case, as the object is radially symmetric, if we turn around the object with for rotation axis the symmetry axis of the object, all the radiographs are exactly the same. Consequently, a single radiograph of such an object is enough to perform the tomographic reconstruction. Indeed, if $f(x, y)$ denotes the density along a slice that contains the symmetry axis (see Figures 1 and

8), then a radiograph of this object is given by

$$g(u,v) = 2\int_{|u|}^{+\infty} f(x,v)\frac{x}{\sqrt{x^2-u^2}}dx.$$

This is a linear transform and we will denote it hereafter by

$$g = Hf.$$

As we already said, this linear operator H is invertible and we in fact know explicitly its inverse on the space of continuously differentiable functions g:

$$f(x,y) = (H^{-1}g)(x,y) = -\frac{1}{\pi}\int_x^{+\infty} \frac{1}{\sqrt{u^2-x^2}}\frac{\partial g}{\partial u}(u,y)du.$$

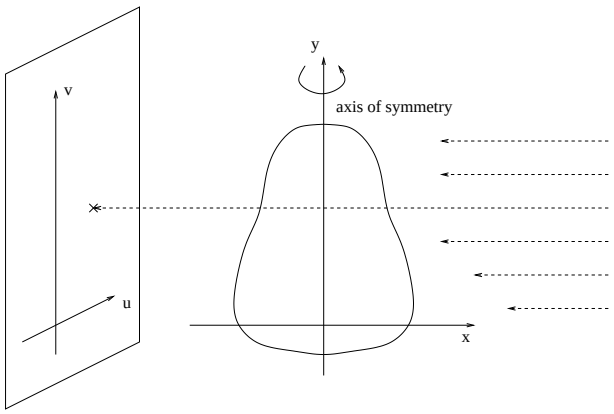


Fig. 8. Radiography of a radially symmetric object.

Our assumption on the noise is that it is an additive Gaussian white noise on the radiograph (i.e. on g). But what we want is to study the object given by f. So we must transform the white noise by the operator H^{-1}. Unfortunately, because of the singularity of the integral at $x = 0$, we cannot apply the operator H^{-1} to a white noise $\dot{B}$, even in a L^2-sense. Therefore, we will work in a discrete framework: the images f and g are naturally discretized (as they are numerical images). This leads to a discretization of the operator H, which we will still denote by H and which now may be viewed as a matrix. The discretization is made in such a way that the symmetry axis ($x = 0$) is settled between two pixels so that the previous singularity does not appear. This matrix is then invertible and we denote by H^{-1} its inverse which we can make now operate on a discrete Gaussian white noise.

4.2 Law of the noise on the tomographic reconstruction

Let us consequently consider a field $\eta = (\eta_{i,j})_{1\leq i\leq p, 1\leq j\leq n}$ of i.i.d. random Gaussian variables with mean 0 and variance σ^2. Let us define $I = (I_{i,j})$ the random field obtained after tomographic reconstruction i.e. after making H^{-1} operate on $\eta = (\eta_{i,j})$. In fact, as the X-rays are supposed to be parallel, the reconstruction can be made line by line independently and therefore, if we consider the row vectors

$$\vec{\eta_i} = (\eta_{i,1}, \ldots, \eta_{i,n}) \quad \text{and} \quad \vec{I_i} = (I_{i,1}, \ldots, I_{i,n})$$

then, there exists an invertible matrix M (independent of i, and of size $n \times n$) such that

$$\vec{I_i} = \vec{\eta_i} M.$$

Consequently, the law of I is characterized by the following properties:

- $I = (I_{i,j})$ is a Gaussian random field.
- For $i \neq k$, $\vec{I_i}$ and $\vec{I_k}$ are independent.
- For each i, the vector $\vec{I_i}$ is a Gaussian vector of mean 0 and covariance matrix
$$\Gamma = \sigma^2 M^t M,$$
where M^t denotes the transpose of M.

4.3 Laws of the gradient and of the Laplacian

The expressions obtained in Section 2 for the gradient and for the Laplacian of an image in a continuous setting are easily translated in the discrete framework we now deal with. Indeed, we have

$$\begin{aligned}
\frac{\partial_r I}{\partial x}(u,v) &= \frac{1}{b(r)} \sum_{(i,j)\in B_r} j I_{u+i,v+j} \\
\frac{\partial_r I}{\partial y}(u,v) &= \frac{1}{b(r)} \sum_{(i,j)\in B_r} i I_{u+i,v+j} \\
\Delta_r I(u,v) &= \frac{1}{\beta(r)} \sum_{(i,j)\in B_r} \left(\alpha(r) + i^2 + j^2\right) I_{u+i,v+j}
\end{aligned}$$

where B_r now denotes the discrete ball of radius r i.e.

$$B_r = \{(i,j),\ i^2 + j^2 \leq r^2\}$$

and where the constants $\alpha(r)$, $\beta(r)$, $b(r)$, ... are the discrete analogous of the constants of Section 2.

With these estimates, the contrast functions C_1 and C_2 are easily comptuted. They are both of the form

$$C(u,v) = \sqrt{C_x^2(u,v) + C_y^2(u,v)}$$

with

$$C_x(u,v) = \sum_{i,j} j c_{ij} I_{u+i,v+j} \quad \text{and} \quad C_y(u,v) = \sum_{i,j} i c_{ij} I_{u+i,v+j},$$

where the coefficients c_{ij} are given by:

(1) In the case of the contrast function C_1,

$$c_{ij} = \frac{1}{b(r)} \mathbb{I}_{(i,j)\in B_r}.$$

(2) In the case of the contrast function C_2 with two balls of radius $r_1 < r_2$,

$$c_{ij} = \frac{1}{b(r_1)} \mathbb{I}_{(i,j)\in B_{r_1}} - \frac{1}{b(r_2)} \mathbb{I}_{(i,j)\in B_{r_2}}.$$

Therefore, the computations of the laws will be similar and they will be treated simultanously using the coefficients c_{ij}.

When the contrast function C_2 is used with two radii $r_1 < r_2$, we then compute the Laplacian $\Delta_r I$ with the larger ball radius, that is with $r = r_2$.

Lemma 3 *For both contrast functions C_1 and C_2, the vector*

$$(C_x(u,v), C_y(u,v), \Delta_r I(u,v))$$

is a Gaussian vector with mean 0 and covariance matrix of the form:

$$\begin{pmatrix} \sigma_x^2 & 0 & \sigma_{x,\Delta} \\ 0 & \sigma_y^2 & 0 \\ \sigma_{x,\Delta} & 0 & \sigma_\Delta^2 \end{pmatrix}$$

In particular, we have that C_y is independent of $(C_x, \Delta_r I)$.

Proof : The lemma is a consequence of the two following remarks. The first one is that, in both cases for the contrast function, the coefficients c_{ij} are symmetric: $c_{i,j} = c_{-i,j}$ and $c_{i,j} = c_{i,-j}$. Thus they satisfy that whenever k or l is odd then

$$\sum_{(i,j)\in B_r} i^k j^l c_{ij} = 0 \tag{3}$$

The second remark is that the vectors $\vec{I_i}$ and $\vec{I_k}$ are independent if $i \neq k$. And we thus have

$$\mathbb{E}\left[I_{i,j}I_{k,l}\right] = \begin{cases} 0 & \text{if } i \neq k, \\ \Gamma(j,l) & \text{if } i = k. \end{cases}$$

We can now compute the covariance matrix. For instance, let us start with:

$$\begin{aligned} \mathbb{E}\left[C_x C_y\right] &= \sum_{(i,j,k,l)} jkc_{ij}c_{kl}\mathbb{E}\left[I_{u+i,v+j}I_{u+k,v+l}\right] \\ &= \sum_{(i,j,l)} jic_{ij}c_{il}\Gamma(v+j,v+l) \\ &= \sum_{(j,l)} j\Gamma(v+j,v+l)\sum_{i} ic_{ij}c_{il} = 0. \end{aligned}$$

Similar computations give $\mathbb{E}\left[C_y\Delta_r I\right] = 0$ and

$$\begin{aligned} \sigma_x^2 :=& \mathbb{E}\left[C_x^2\right] = \sum_{(i,j,l)} jlc_{ij}c_{il}\Gamma(v+j,v+l); \\ \sigma_y^2 :=& \mathbb{E}\left[C_y^2\right] = \sum_{(i,j,l)} i^2c_{ij}c_{il}\Gamma(v+j,v+l); \\ \sigma_\Delta^2 :=& \mathbb{E}\left[(\Delta_r I)^2\right] = \frac{1}{\beta^2(r)}\sum_{(i,j,l)\in\Omega_r} (\alpha(r)+i^2+j^2)(\alpha(r)+i^2+l^2)\Gamma(v+j,v+l); \\ \sigma_{x,\Delta} :=& \mathbb{E}\left[C_x\Delta_r I\right] = \frac{1}{\beta(r_2)}\sum_{(i,j,l)\in\Omega_{r_2}} jc_{ij}(\alpha(r)+i^2+l^2)\Gamma(v+j,v+l), \end{aligned}$$

where we have set $\Omega_r = \{(i,j,l) \text{ such that } (i,j) \in B_r \text{ and } (i,l) \in B_r\}$.

□

4.4 Computation of the threshold

Now, as we have no more independence between the first and the second order derivatives we must compute the conditional law of the contrast function knowing that $\Delta_r I = 0$.

Proposition 2 *Let C be one of the two contrast functions. Then, the random variable $\|C\|^2$ is distributed, conditionally on $\{\Delta_r I = 0\}$, as the sum of the square of two independent Gaussian random variables, with mean zero and respective variance*

$$\sigma_y^2 \quad \textit{and} \quad \sigma_{x|\Delta=0}^2 = \frac{\sigma_x^2\sigma_\Delta^2 - \sigma_{x,\Delta}^2}{\sigma_\Delta^2},$$

that is a Gamma law with parameters $\frac{1}{2}$ *and* $\frac{1}{2}(\sigma_y^2 + \sigma_{x|\Delta=0}^2)$.

The threshold value $s(\varepsilon)$ defined by

$$\mathbb{P}\Big(\|C\| \geq s(\varepsilon) \;\Big|\; \Delta_r I = 0\Big) \leq \varepsilon$$

can no longer be computed explicitly but a numerical approximation is easy to get as the Gamma density is well-known.

Proof : C_y is independent of the pair $(C_x, \Delta_r I)$. Thus, conditionally on $\{\Delta_r I = 0\}$, the random variables C_y and C_x are still independent and the conditional law of C_y is the Gaussian distribution with mean 0 and variance σ_y^2.

Now, if $D^2 := \sigma_x^2\sigma_\Delta^2 - \sigma_{x,\Delta}^2 \neq 0$, then the law of the pair $(C_x, \Delta_r I)$ has a density which is given by

$$f_{x,\Delta}(t_1, t_2) = \frac{1}{2\pi D} e^{-\frac{1}{2}(t_1,t_2)\Lambda(t_1,t_2)^t}$$

where Λ is the inverse of the covariance matrix, i.e.

$$\Lambda = \frac{1}{D^2}\begin{pmatrix} \sigma_\Delta^2 & -\sigma_{x,\Delta} \\ -\sigma_{x,\Delta} & \sigma_x^2 \end{pmatrix}.$$

Let us recall that, if f_Δ denotes the Gaussian density of $\Delta_r I$, then the law of C_x conditionally on $\Delta_r I = 0$ has a density given by

$$\frac{f_{x,\Delta}(t_1, 0)}{f_\Delta(0)}$$

and so is Gaussian with mean zero and variance

$$\sigma_{x|\Delta=0}^2 = \frac{D^2}{\sigma_\Delta^2}.$$

This result is still valid when $D = 0$ since it implies that C_x and $\Delta_r I$ are proportional and thus the law of C_x conditionally on $\Delta_r I = 0$ is Gaussian with mean 0 and variance 0 (it is not random anymore). □

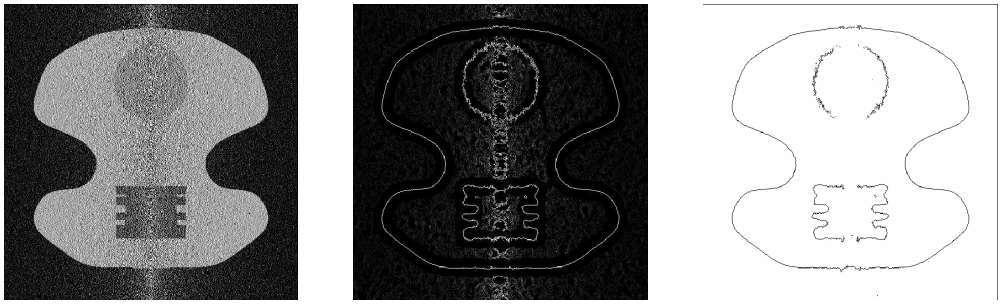
4.5 Experiments

4.5.1 Case of a piecewise constant object

To begin with, we still study the piecewise constant object of Figure 1 described in Section 3.4. Let us recall that this image represents a slice of the

object that contains the symmetry axis. The 3-dimensional object is obtained by rotation around the vertical axis that goes through the middle of the image.

In that case, we will use the contrast function C_1, which is simply the norm of the gradient. The experiments of Figure 9 correspond to a ball radius $r = 12$ pixels.
We start with the image of the radiograph obtained after the application of matrix H to our initial image. Then a Gaussian white noise is added to this radiograph. Then tomographic inversion (application of the matrix H^{-1}) is performed. This gives the image of Figure 9(a). As we already mentioned it, the noise is not stationary, it is now correlated and its variance depends on the distance from the symmetry axis. For instance, if the standard deviation of the Gaussian white noise on the radiograph is $\sigma = 4$, the variance of the noise on the tomography is about $2\sigma^2 = 32$ near the axis, $0.02\sigma^2 = 0.32$ at a distance of 65 pixels from the axis and $8.10^{-3}\sigma^2 = 0.128$ at the edge of the image located on the right at 200 pixels from the axis.



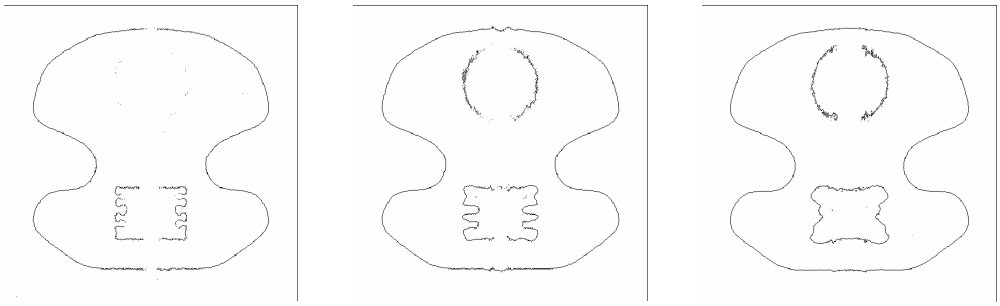
(a) (b) (c)

Fig. 9. (a) Reconstructed object from a single noisy radiograph, (b) Contrast value at the zero-crossings of the Laplacian for the contrast function C_1, (c) Significant edges.

We notice that the edges are not significant near the symmetry axis; the noise is too important here in order to extract the true edges from the noise. Let us add that the smaller the difference of the densities of the material is, the larger the region where the edges are not significant around the axis is. Even when the edges are significant, the noise and the method used to detect them can lead to noisy edges. Moreover, some details are lost because of the smoothing due to the size of the ball.

Let us compare the results obtained with different ball radii (see Figure 10). When the ball radius is small, the edges are more accurate but some are not significant: the smoothing of the noise is not enough to get rid of it. On the contrary, when the radius is large, most of the edges are detected but small

details are lost because of this smoothing.



$r = 6$ $r = 12$ $r = 20$

Fig. 10. Significant edges obtained with different ball radius: from left to right: $r = 6$, $r = 12$ and $r = 20$

Since the edges separate two materials, one included in another, they must be closed curves. Usually, an operator has to close them manually. Our method gives open edges. It does not mean that there is no edge between the materials: it simply means that the noise level is too high to give an accurate position of the edge. Therefore, we can then close the curves manually, or by usual curve completion methods, but this will not tell which closure is better (i.e. the closest to the real shape).

Comparison with other methods. We will give here the results obtain with two other methods which have both the advantage of directly providing closed curves.

• The first method is the one introduced in [4]. One keeps only the meaningful level lines of the image, which are defined by: the minimum of the norm of the gradient along the level line is larger than a threshold $T(\varepsilon)$. This threshold is computed from the gradient histogram of the image. The meaning of this definition is that such curves have a probability less than ε to appear in a pure noise image (with same gradient histogram as the original image). The results obtained with this method are shown on Figure 11. On the first row: we smooth the image of Figure 9(a) by convolution with a Gaussian kernel with respective standard deviation 2 and 4 pixels. And then, on the second row, we have the respective obtained meaningul level lines. This experiment clearly shows that, since the noise model is not adapted to the image (in particular, the non-stationarity is not taken into account), many false contours are detected.

• The second method is the famous Mumford-Shah segmentation for piecewise constant images [9]. Given an observed image g_0 defined on a domain D, one looks for the piecewise constant approximation g of g_0 that minimizes the functional

$$E_\lambda(g) = \int_D |g - g_0|^2 + \lambda \,\mathrm{Length}\Big(K(g)\Big),$$

where Length$\Big(K(g)\Big)$ is the one-dimensional measure of the discontinuity set of g (which is a set of curves denoted by $K(g)$) and λ is a parameter which weights the second term of the functional. The results obtained with this method are shown on Figure 12 for three different values of λ. The main drawbacks of this method are: (a) there is no verification that the obtained contours are not due to the noise; (b) the parameter λ has to be fixed, and the results are very dependent on its value.

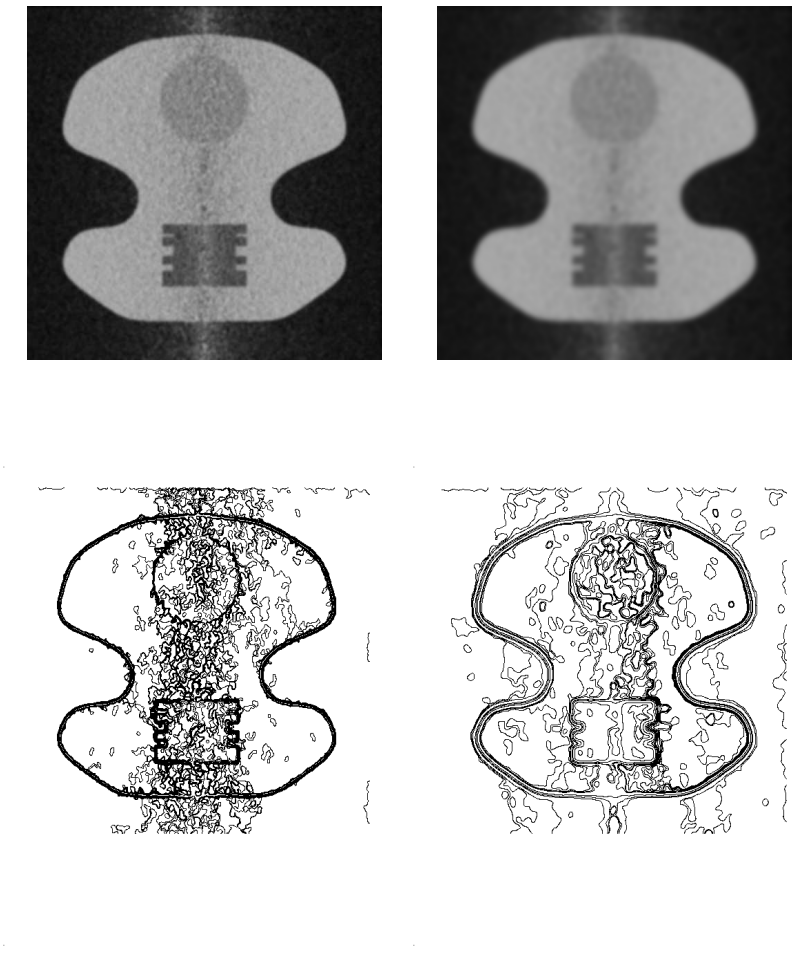


Fig. 11. First row: the image of Figure 9(a) is smoothed by convolution with a Gaussian kernel with respective standard deviation 2 and 4 pixels. Second row: the meaningul level lines of each image.

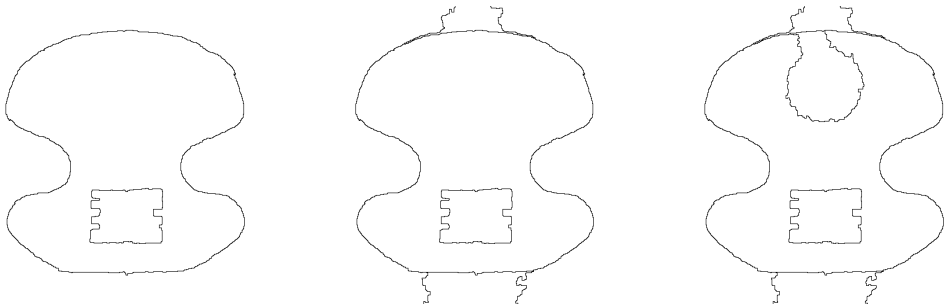


Fig. 12. Results obtained with the Mumford-Shah segmentation for piecewise constant images, for three different values of λ. From left to right, the number of regions in the segmented image is respectively 3, 6 and 7.

4.5.2 Case of an inhomogeneous material

Let us turn now to a more realistic case: the materials are not homogeneous and consequently the object is no more piecewise constant (see Figure 13). As already said, the use of the contrast function C_1 fails in that case. This is illustrated by Figure 14. In this image, one can notice that there are many false detections especially in the parts of the image where it is not constant. Figure 15 gives the significant edges obtained with the contrast function C_2 with two ball radii $r_1 = 6$ and $r_2 = 12$. With this contrast function, we eventually get only the "true" edges.

(a) (b) (c)

Fig. 13. (a) Inhomogeneous object, (b) Its noisy radiograph, (c) Tomographic reconstruction

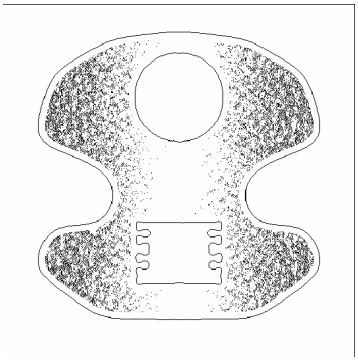


Fig. 14. Significant edges with the C_1 contrast function: there are many false detections.

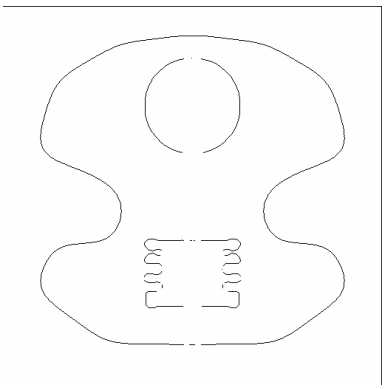


Fig. 15. Significant edges with the C_2 contrast function: only the "true" edges are obtained.

References

[1] J. Canny, A computational approach to edge detection, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 8, pp. 679-698, 1986.

[2] F. Cao, P. Musé and F. Sur, Extracting Meaningful Curves from Images, *Journal of Mathematical Imaging and Vision* 22, pp. 159-181, 2005.

[3] I. Abraham, R. Abraham, J.-M. Lagrange and F. Lavallou, Méthodes inverses pour la reconstruction tomographique X monovue, *Revue Chocs* 31 (chocs@cea.fr), 2005.

[4] A. Desolneux, L. Moisan and J.-M. Morel, Edge Detection by Helmholtz Principle, *Journal of Mathematical Imaging and Vision* 14, pp. 271-284, 2001.

[5] J.M. Dinten, Tomographie à partir d'un nombre limité de projections : Régularisation par des champs markoviens, PhD Thesis, Université Paris Sud, 1990.

[6] T. Hida, *Brownian Motion*, Applications of Mathematics 11, Springer-Verlag, 1980.

[7] T. Hida, H.H. Kuo, J. Potthoff and L. Streit, *White Noise. An infinite Dimensional Calculus.* Mathematics and its Applications 253, Kluwer Academic Publishers Group, Dordrecht, 1993.

[8] D. Marimont and Y. Rubner, A probabilistic framework for edge detection and scale selection, *6th Int. Conference on Computer Vision*, 1998.

[9] D. Mumford and J. Shah, Boundary detection by minimizing functionals, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, 1985.

[10] P. Qiu and S. Bhandarkar, An edge detection technique using local smoothing and statistical hypothesis testing, *Pattern Recognition Letters* 17, pp. 849-872, 1996.

[11] R. Touzi, A. Lopes and P. Bousquet, A statistical and geometrical edge detector for SAR images, *IEEE Transactions on Geoscience and Remote Sensing* 26, pp. 764-773, 1988.

[12] J.B. Walsh, An introduction to stochastic partial differential equations, *Ecole d'été de Probabilités de Saint-Flour XIV 1984*, Lecture Notes in Math. 1180, *Springer, Berlin* 1986.