

Didier AUROUX

Laboratoire MIP - Institut de Mathématiques de Toulouse

Université Paul Sabatier Toulouse 3

France

auroux@mip.ups-tlse.fr



Image restoration and classification by topological asymptotic expansion.

Mathematics and Image Analysis 2006,

MIA'06, Paris, France

September 18-21, 2006



Talk overview



1. Introduction to topological asymptotic expansion
2. Image restoration
 - (a) Linear approach
 - (b) Nonlinear variational approach
 - (c) Topological gradient approach
3. Regularized image classification
 - (a) Topological gradient applied to the classical approach
 - (b) Restoration-based preprocessing method
4. A few more numerical results
5. Conclusion



Introduction

Image restoration problem : how to de-noise the image? In order to prevent the image becoming blurred, we must **identify its contours** in order to preserve them.

To find an optimal shape (the set of contours) is equivalent to find its characteristic function (0 – 1 optimization problem).

The goal of image classification is to find a partition of the image in subsets with a constant color level in each subset.

Main problem in both cases : **the non differentiability**.



ASYMPTOTIC ANALYSIS AND TOPOLOGICAL GRADIENT



Topological gradient



General formulation : Let Ω be a bounded **open set**, and let us consider a **PDE** on this set. We denote by u_Ω the solution of this PDE. We also consider a **cost function** $J(u_\Omega)$:

$$\Omega \longrightarrow u_\Omega \longrightarrow j(\Omega) := J(u_\Omega).$$

The topological asymptotic measures the **impact of creating a hole** around a point x in the domain Ω on the cost function :

$$j(\Omega \setminus \varepsilon B_x) - j(\Omega) = f(\varepsilon)g(x) + o(f(\varepsilon))$$

where $f(\varepsilon) > 0$ and $\lim_{\varepsilon \rightarrow 0} f(\varepsilon) = 0$.

g is called the **topological gradient**, and one should create holes where the topological gradient is negative.



Topological gradient computation



We consider the variational formulation of the PDE :

$$a_\varepsilon(u_\varepsilon, w) = l_\varepsilon(w), \quad \forall w \in V.$$

We assume that V is a Hilbert space, a_ε and l_ε are continuous functions, and that a_ε is coercive. **Hypothesis :**

$$\begin{cases} J_\varepsilon(u_\varepsilon) - J_\varepsilon(u_0) = L_\varepsilon(u_\varepsilon - u_0) + f(\varepsilon)\delta J_1 + o(f(\varepsilon)), \\ J_\varepsilon(u_0) - J_0(u_0) = f(\varepsilon)\delta J_2 + o(f(\varepsilon)), \\ (a_\varepsilon - a_0)(u_0, v_\varepsilon) = f(\varepsilon)\delta a + o(f(\varepsilon)), \\ (l_\varepsilon - l_0)(v_\varepsilon) = f(\varepsilon)\delta l + o(f(\varepsilon)), \end{cases}$$

where v_ε is solution of the adjoint equation :

$$a_\varepsilon(w, v_\varepsilon) = -L_\varepsilon(w), \quad \forall w \in V.$$

Then, the topological gradient is given by $\delta a - \delta l + \delta J_1 + \delta J_2$.



IMAGE RESTORATION

Problem : how to restore a noised image ?

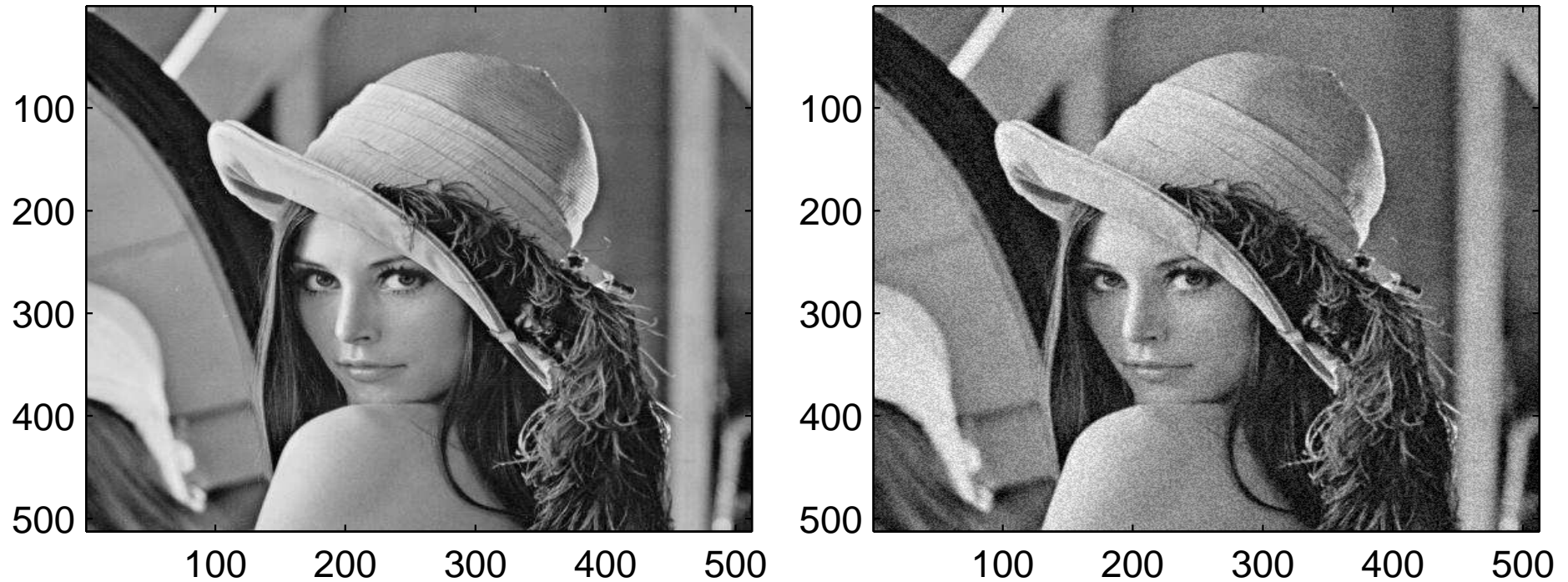


FIG. 1 – Original image (left) and noised image (right).



Linear approach



We consider the following operator

$$\begin{aligned} K : H^1(\Omega) &\rightarrow L^2(\Omega), \\ u &\mapsto u \end{aligned}$$

and we have to solve the following problem

$$Ku = v.$$

A necessary optimality condition is

$$K^*Ku = K^*v,$$

which is an ill-posed problem. The regularization of Tikhonov gives

$$K^*Ku + cu = K^*v.$$

The weak formulation is

$$\langle K^*Ku + cu, w \rangle_{H^1(\Omega)} = \langle K^*v, w \rangle_{H^1(\Omega)}, \quad \forall w \in H^1(\Omega).$$



Linear approach



By definition of K , the previous equation is equivalent to the following one :

$$\langle u, w \rangle_{L^2(\Omega)} + \langle cu, w \rangle_{H^1(\Omega)} = \langle v, Kw \rangle_{L^2(\Omega)}, \quad \forall w \in H^1(\Omega).$$

The **classical variational formulation** of image restoration is then the following :

$\Omega \subset \mathbb{R}^2$, $v \in L^2(\Omega)$ is the noised image, and we have to find the **solution** $u \in H^1(\Omega)$ of

$$\begin{cases} -\operatorname{div}(c\nabla u) + u = v & \text{in } \Omega, \\ \partial_n u = 0 & \text{in } \Gamma = \partial\Omega. \end{cases}$$



Nonlinear approach



$$E(u) = \frac{1}{2} \int_{\Omega} |v - u|^2 dx + \lambda \int_{\Omega} \psi(|\nabla u|) dx.$$

By choosing $\psi(|\nabla u|) = |\nabla u|^2$, we obtain the linear approach. If $E(u)$ has a minimum u , u must satisfy the Euler-Lagrange equation :

$$-\lambda \operatorname{div} \left(\psi'(|\nabla u|) \frac{\nabla u}{|\nabla u|} \right) + u = v.$$

Hypothesis on ψ :

- $\psi'(0) = 0$ and $\lim_{t \rightarrow 0^+} \frac{\psi'(t)}{t} = \psi''(0) > 0$: isotropic regularization where the gradient is weak.
- $\lim_{t \rightarrow +\infty} \frac{\psi'(t)}{t} = \lim_{t \rightarrow +\infty} \psi''(t) = 0$ and $\lim_{t \rightarrow +\infty} \frac{\psi''(t)}{\psi'(t)/t} = 0$: anisotropic regularization where the gradient is strong.
- $\lim_{t \rightarrow +\infty} \psi(t) = +\infty$: well posed problem.



Topological gradient



We consider the classical restoration PDE :

$$\begin{cases} -\operatorname{div}(c\nabla u) + u = v & \text{in } \Omega, \\ \partial_n u = 0 & \text{in } \Gamma = \partial\Omega. \end{cases}$$

Let $x_0 \in \Omega$ and $\varepsilon > 0$ *small*, we denote by $\Omega_\varepsilon = \Omega \setminus \bar{\sigma}_\varepsilon$ the **perturbed domain** by the insertion of a **crack** $\sigma_\varepsilon = x_0 + \varepsilon\sigma(n)$, where n is a unit vector normal to the crack.

The solution $u_\varepsilon \in H^1(\Omega_\varepsilon)$ of the perturbed problem satisfies

$$\begin{cases} -\operatorname{div}(c\nabla u_\varepsilon) + u_\varepsilon = v & \text{in } \Omega_\varepsilon, \\ \partial_n u_\varepsilon = 0 & \text{in } \Gamma_\varepsilon = \partial\Omega_\varepsilon. \end{cases}$$



Topological gradient



Variational formulation :

$$a_\varepsilon(u_\varepsilon, w) = l_\varepsilon(w), \quad \forall w \in H^1(\Omega_\varepsilon)$$

where $a_\varepsilon(u, w) = \int_{\Omega_\varepsilon} (c \nabla u \nabla w + uw) dx$ and $l_\varepsilon(w) = \int_{\Omega_\varepsilon} vw dx$.

To find the contours of the image is equivalent to find a subset of Ω in which the energy is weak \Rightarrow **minimize the energy out of the contours** :

$$j(\varepsilon) = J_\varepsilon(u_\varepsilon) = \int_{\Omega_\varepsilon} \|\nabla u_\varepsilon\|^2.$$



Topological gradient



We have then the following asymptotic expansion :

$$j(\varepsilon) - j(0) = \varepsilon^2 G(x_0, n) + o(\varepsilon^2)$$

$$G(x_0, n) = -\pi(\nabla u_0(x_0) \cdot n)(\nabla v_0(x_0) \cdot n) - \pi|\nabla u_0(x_0) \cdot n|^2$$

where v_0 is solution of the adjoint problem

$$\begin{cases} -\operatorname{div}(c\nabla v_0) + v_0 = -\partial_u J(u) & \text{in } \Omega, \\ \partial_n v_0 = 0 & \text{in } \Gamma = \partial\Omega. \end{cases}$$



Topological gradient



The topological gradient can be written as

$$G(x, n) = \langle M(x)n, n \rangle$$

where $M(x)$ is the symmetric matrix defined by

$$M(x) = -\pi \frac{\nabla u_0(x) \nabla v_0(x)^T + \nabla v_0(x) \nabla u_0(x)^T}{2} - \pi \nabla u_0(x) \nabla u_0(x)^T.$$

For a given x , $G(x, n)$ takes its minimal value when n is the eigenvector associated to the lowest eigenvalue λ_{min} of M . This value will be considered as the **topological gradient** associated to the optimal orientation of the crack $\sigma_\varepsilon(n)$.



Algorithm :

- Initialization : $c = c_0$.
- Computation of u_0 and v_0 , solutions of the direct and adjoint problems.
- Computation of the 2×2 matrix M and its lowest eigenvalue λ_{min} at each point of the domain.
- $c_1 = \begin{cases} \varepsilon_c & \text{if } x \in \Omega \text{ such that } \lambda_{min} < \alpha < 0, \quad \varepsilon_c > 0 \\ c_0 & \text{elsewhere.} \end{cases}$
- Calculation of u_1 solution to the perturbed problem with $c = c_1$.

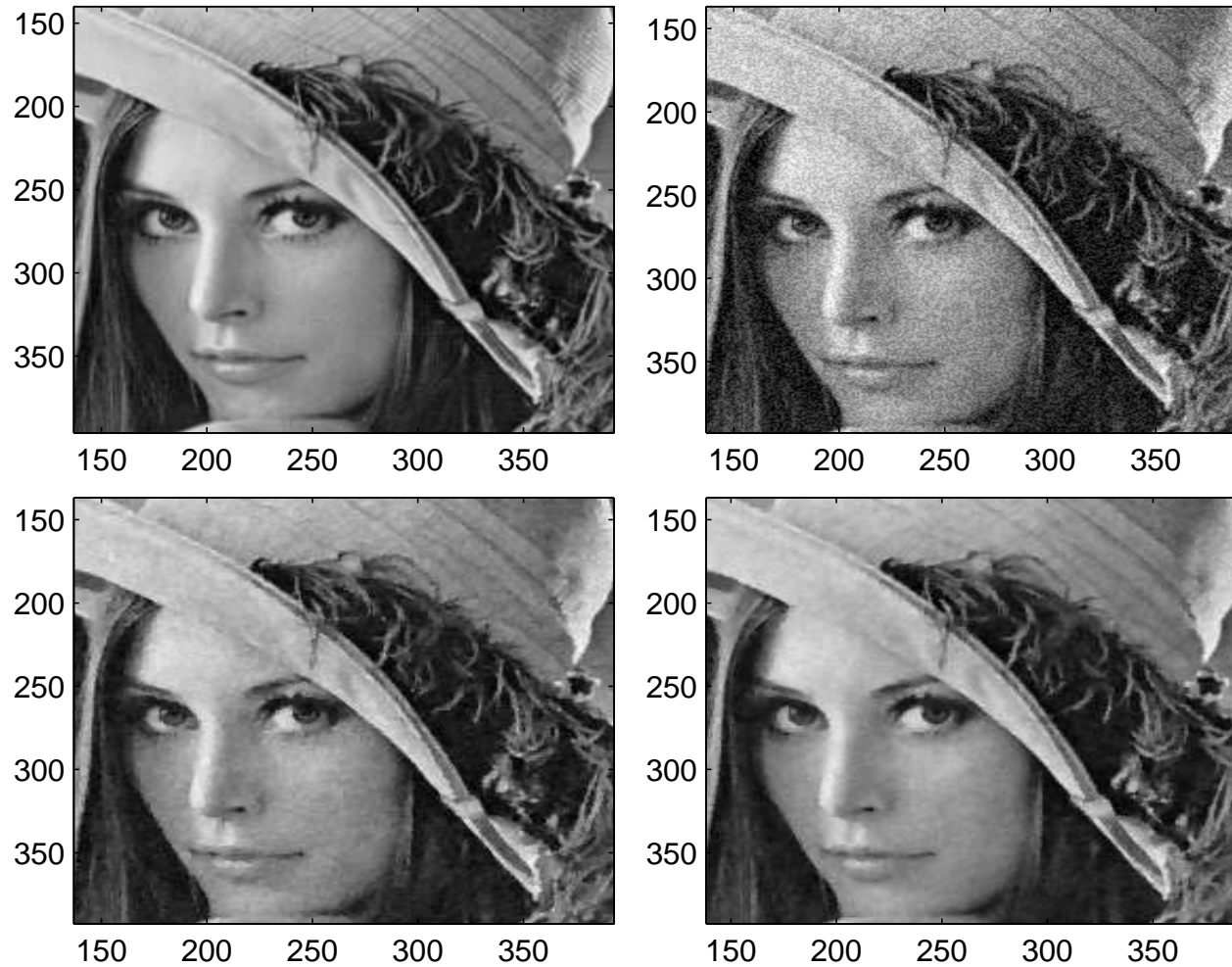


FIG. 2 – Top : Original image (left) and noised image (SNR=17) (right); Bottom : restored image by nonlinear diffusion (SNR=27) (left) and restored image by topological gradient (SNR=29) (right).

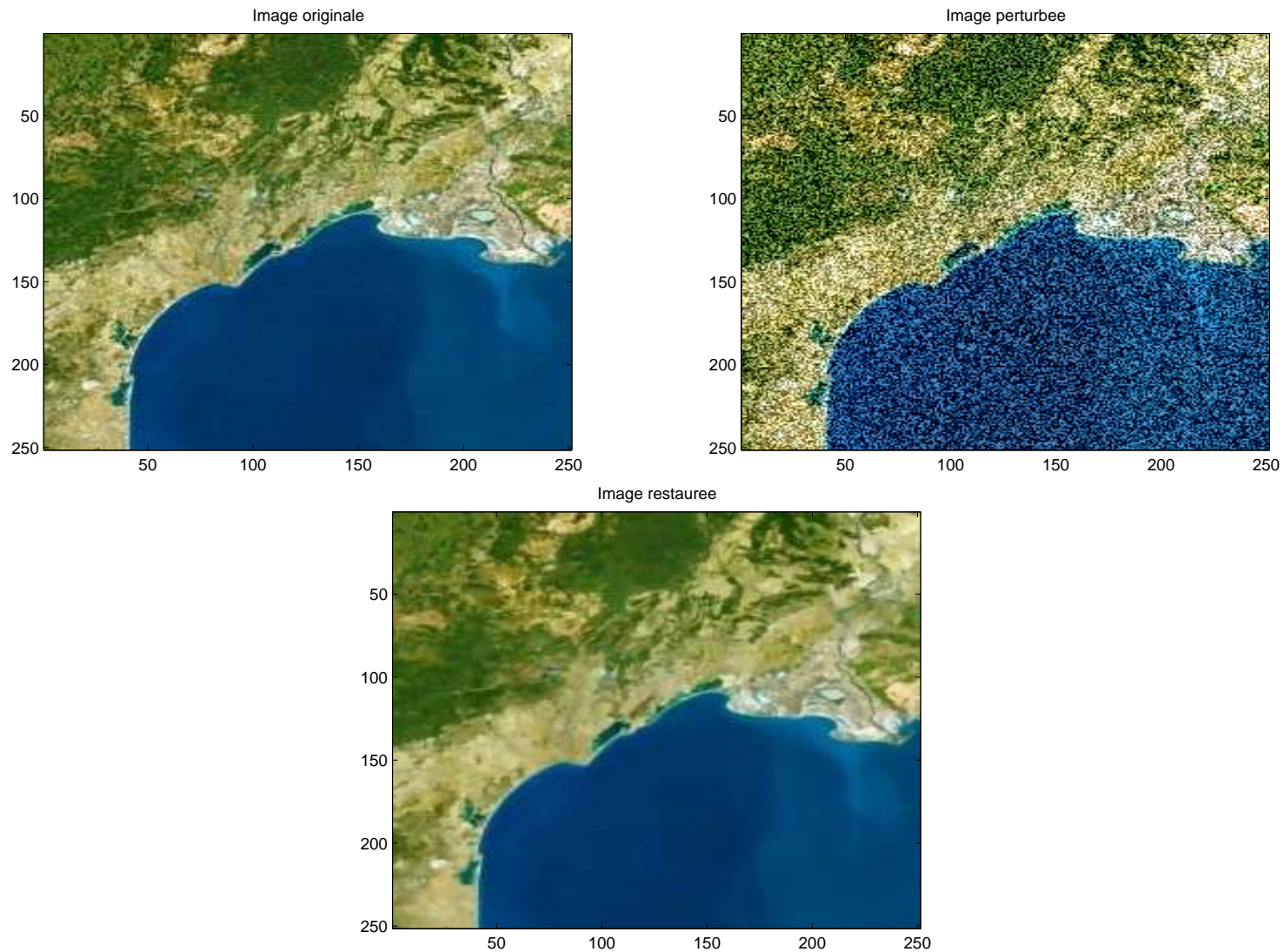


FIG. 3 – Top : Original image (left) and noised image (SNR=10) (right); Bottom : restored image by topological gradient (SNR=23).

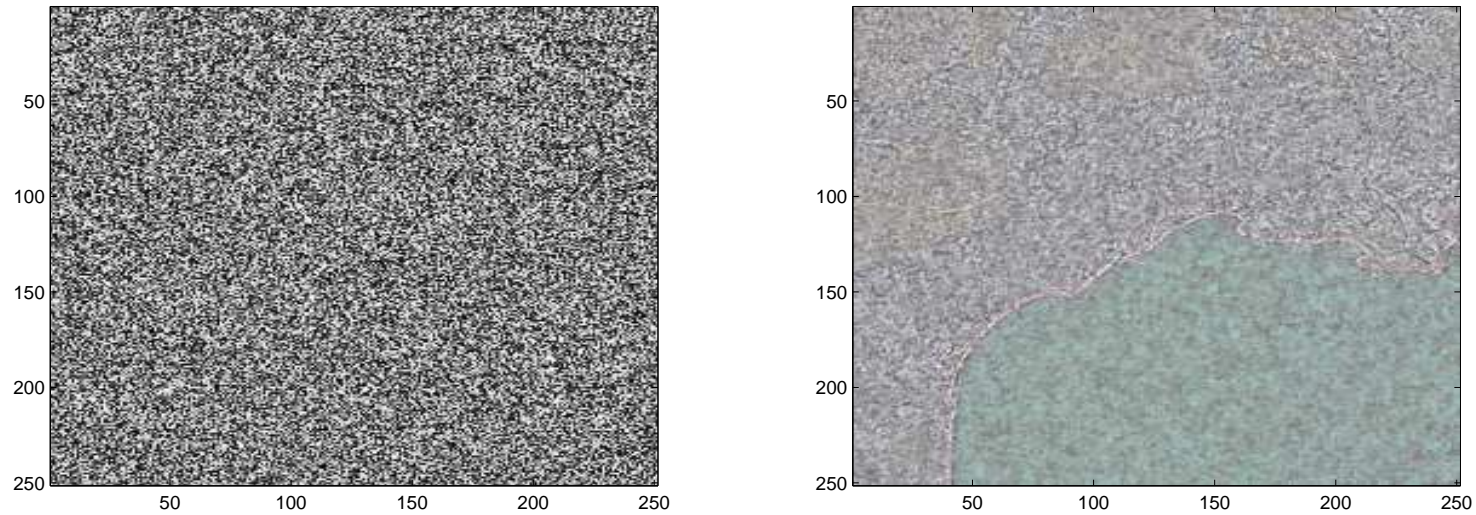


FIG. 4 – Difference between the original and noised images (left); Error on the restored image by topological gradient (right).



DCT algorithm - c constant



Let us consider the Fourier basis

$$\phi_{m,n} = \delta_{m,n} \cos(m\pi x) \cos(n\pi y)$$

where $\delta_{m,n}$ are normalisation coefficients.

If c is constant,

$$-c \Delta u + u = v$$

is equivalent to

$$\sum_{m,n} (1 + c(m\pi)^2 + c(n\pi)^2) u_{m,n} \phi_{m,n} = \sum_{m,n} v_{m,n} \phi_{m,n}.$$



DCT algorithm



Discrete Cosine Transform :

- calculate $v_{m,n}$ the DCT of v ,
- the DCT of u is then

$$u_{m,n} = \frac{v_{m,n}}{1 + c(m\pi)^2 + c(n\pi)^2}$$

- $u = \sum_{m,n} u_{m,n} \phi_{m,n}$ (inverse DCT).

Complexity : $\mathcal{O}(n \log(n))$ operations.



DCT preconditioner



When c is close to a constant :
we have to solve the linear system

$$A(c)u = B.$$

We use the Preconditioned Conjugate Gradient method :

$$A(c_0)^{-1}A(c)u = A(c_0)^{-1}B.$$

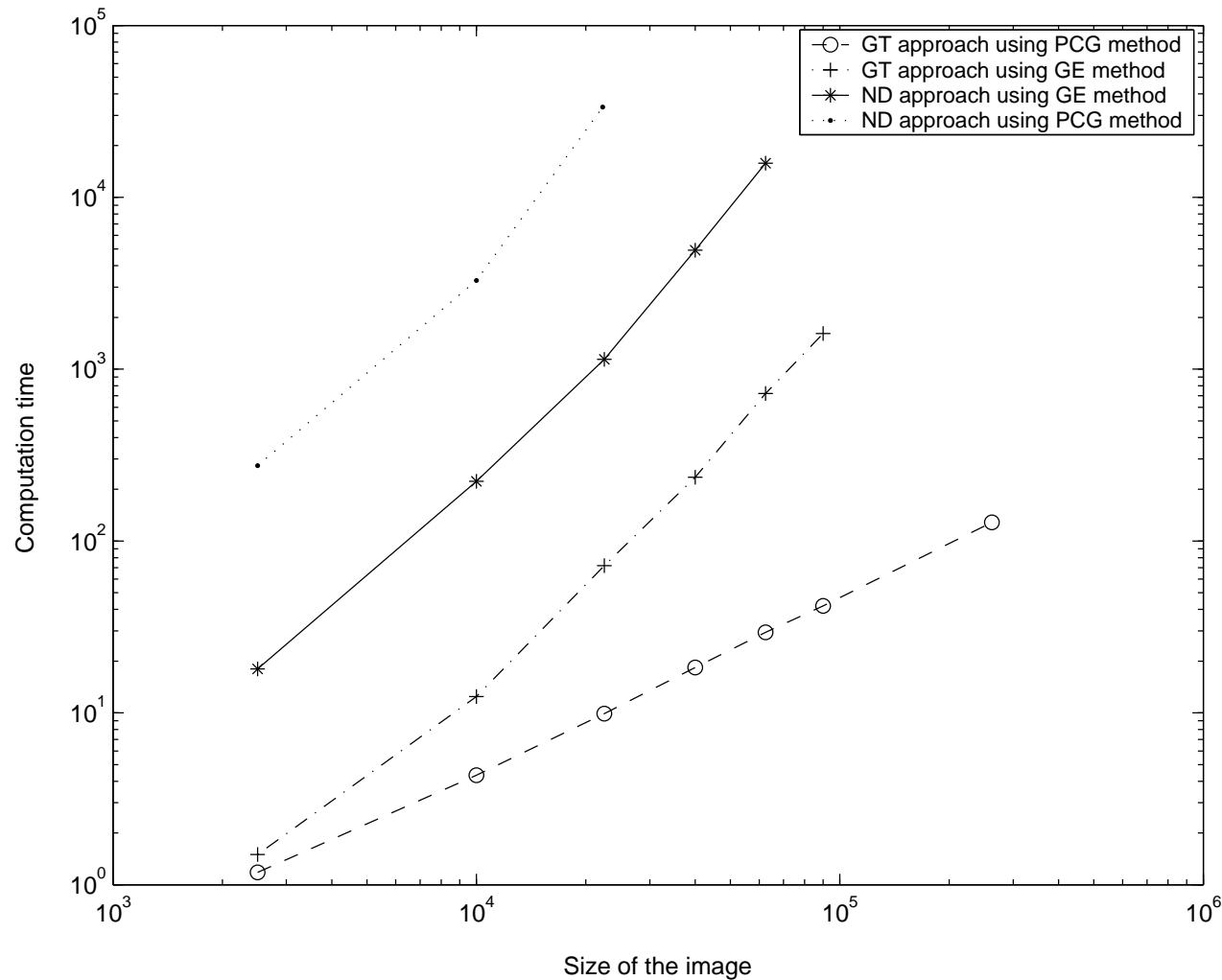


FIG. 5 – Computation time versus the size of the image for the topological gradient (GT) and nonlinear diffusion (ND) approaches. GE : Gauss elimination method, PCG : preconditioned conjugate gradient using a discrete cosine transform.



IMAGE CLASSIFICATION



Image classification



Data : $\Omega =]0, 1[\times]0, 1[$, w an image, $(\mu_i)_{i=1..K}$ color (or grey level) classes.

Problem : find a partition of Ω in subsets $(\Omega_i)_{i=1..K}$ such that :

- w is close to μ_i in Ω_i ,
- the length of interfaces between the different subsets Ω_i is minimum.

We first assume that the **number** and **values** of classes are known.

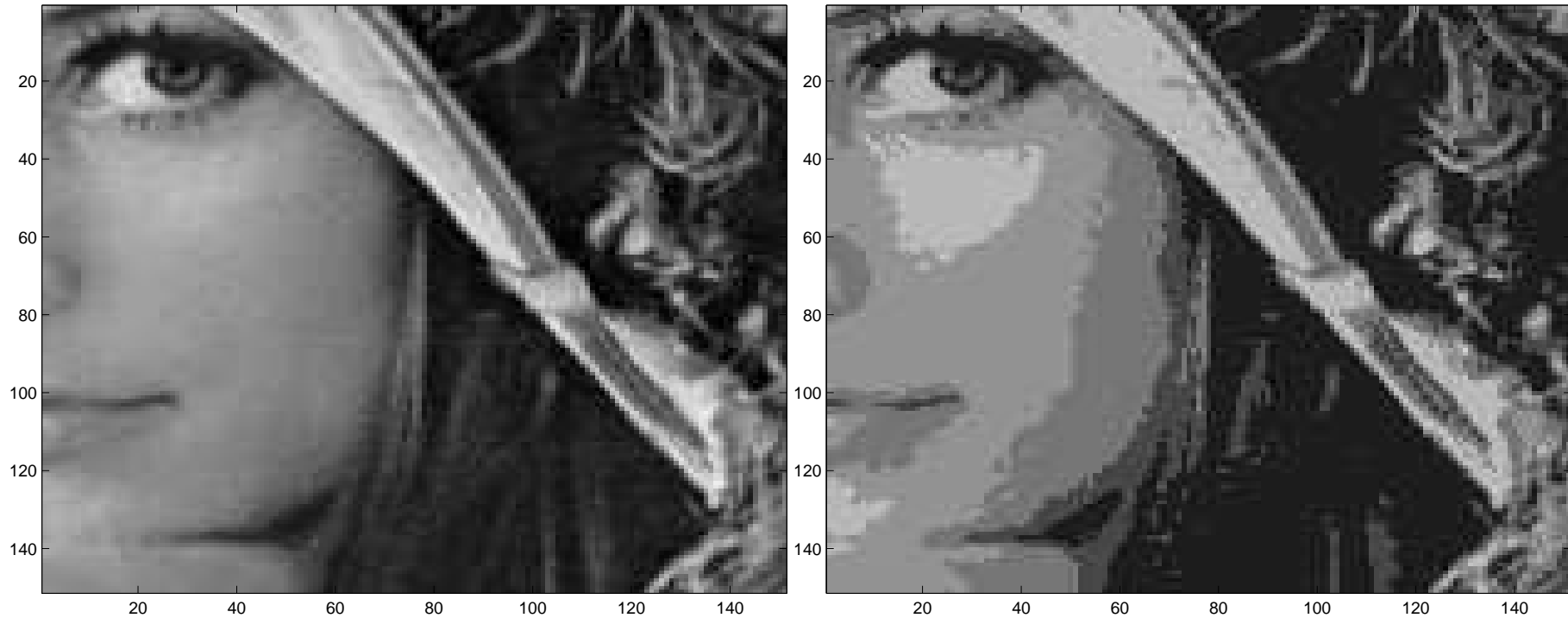


FIG. 6 – Original image (left) and classified image (without regularization) (right).



Image classification



We have to minimize with respect to Ω_i

$$J_1 = \sum_{i=1}^K \int_{\Omega_i} (w - \mu_i)^2 dx$$

and

$$J_2 = \sum_{i \neq j} |\Gamma_{ij}|.$$

The main difficulty comes from the fact that the unknowns are sets and not variables \implies define a **topological gradient for each class**.

For each pixel, the most negative topological gradient gives the subset and the class to which it should be reassigned in order to minimize the cost function.

Remark : in the present case, the asymptotic expansion is in fact an exact variation.

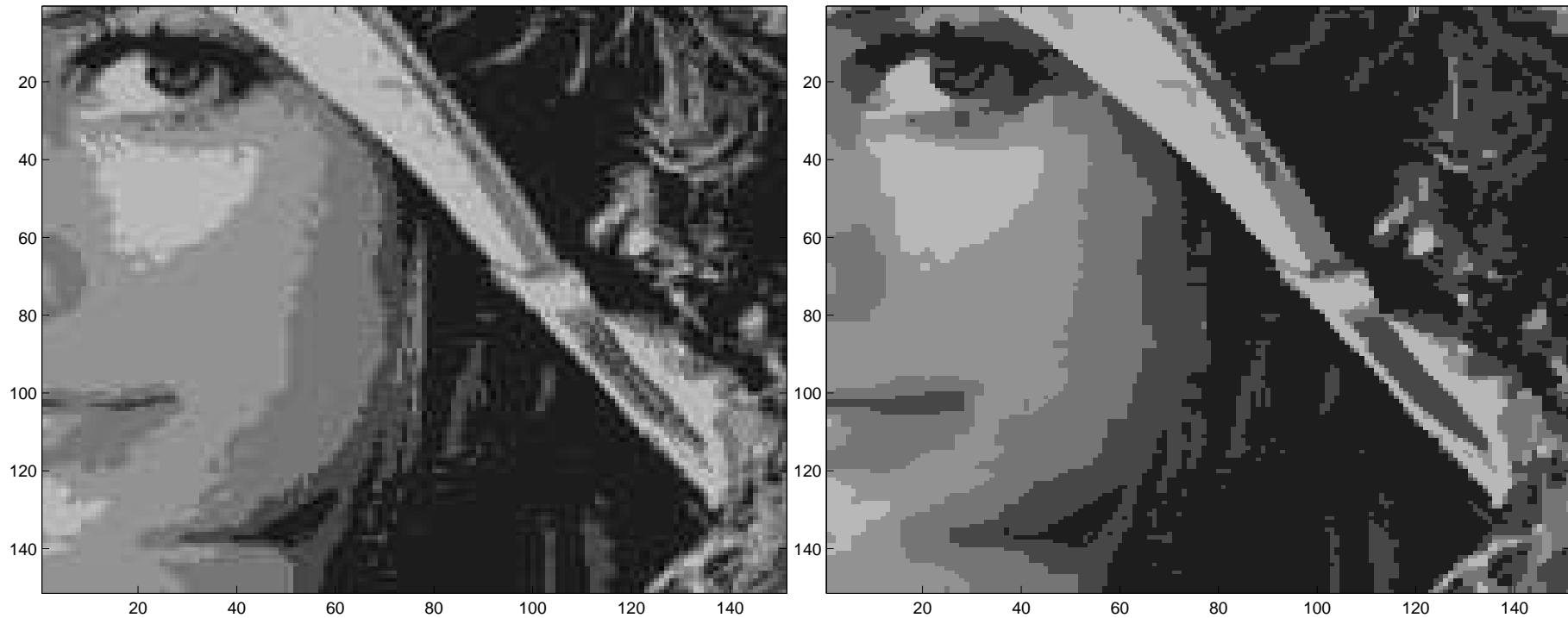


FIG. 7 – Classified images, without regularization (left) and with regularization (right).



Another way to use the topological gradient : we consider the following PDE

$$\begin{cases} -\operatorname{div}(c\nabla u) + u = w & \text{in } \Omega, \\ \partial_n u = 0 & \text{in } \Gamma = \partial\Omega, \end{cases}$$

where $c = \frac{1}{\varepsilon_c} \cdot \chi_{\Omega_1} + \varepsilon_c \cdot \chi_{\Omega_\varepsilon}$.

- if the pixel is on a **contour** ($c = \varepsilon_c$), the PDE is nearly equivalent to $u = w$,
- if it is **not on a contour** ($c = \frac{1}{\varepsilon_c}$), the PDE is nearly equivalent to $\Delta u = 0$ and we smooth the image.

We apply then the **topological gradient** method to this PDE.

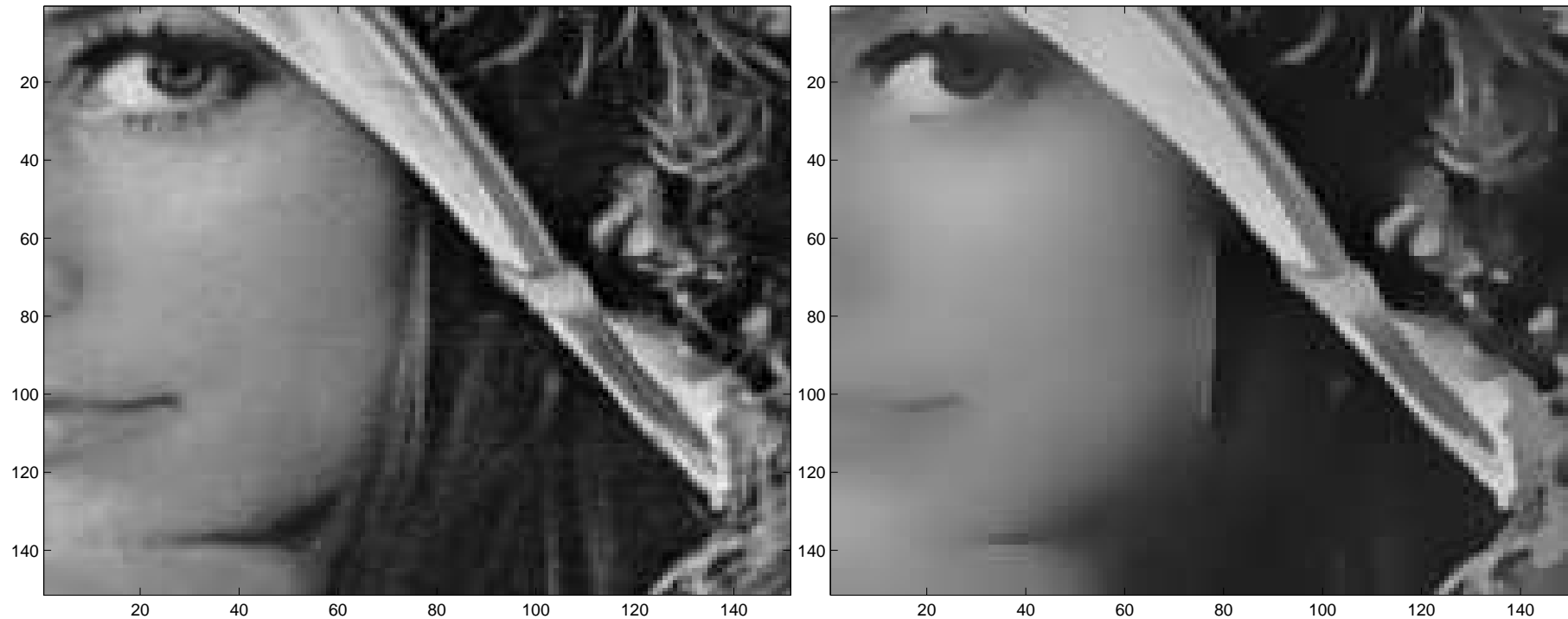


FIG. 8 – Original image (left) and smooth image (right).



Image classification

If we assume that the color classes are given, we only have to apply the previous **unregularized classification** method : each pixel is reassigned to its closest class.

If the **classes are not given**, it is possible to determine them in an **optimal way**, still by using the topological gradient method. The idea is to study the impact of changing a class $\mu_i := \mu_i + 1$ or $\mu_i - 1$ on the cost function.

If the **number of classes is not given**, we can add a penalization term in the cost function, measuring the number of classes and the previous algorithm provides the *optimal* number of classes, and their *optimal* values.

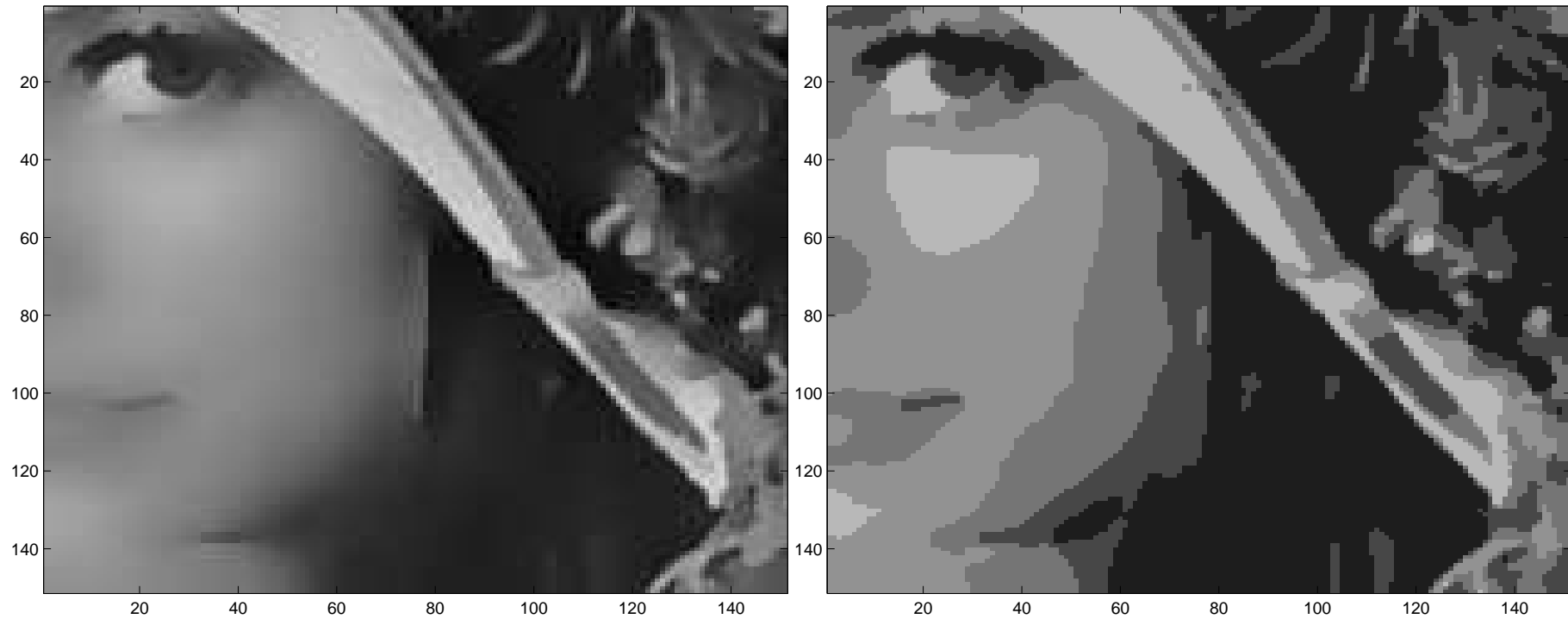


FIG. 9 – Smooth image (left) and classified image (right).



FIG. 10 – Top : original image (left), unregularized classified image (right); Bottom : regularized classified image (left), smooth and classified image (right).



Image classification



TAB. 1 – Computational cost (time in seconds) and length of the interfaces for the different algorithms and number of classes.

Number of classes	$n = 2$	$n = 3$	$n = 5$
Algorithm	$C = \{0; 255\}$	$C = \{34; 112; 165\}$	$C = \{29; 71; 117; 146; 184\}$
Closest class (unregularized)	$t = 0.02$ $ \Gamma = 2358$	$t = 0.06$ $ \Gamma = 4513$	$t = 0.05$ $ \Gamma = 7913$
Topological gradient (regularized)	$t = 12.67$ $ \Gamma = 2069$	$t = 45.63$ $ \Gamma = 3872$	$t = 81.78$ $ \Gamma = 5770$
Smothering - closest class (+ topological gradient)	$t = 37.16$ $ \Gamma = 1566$	$t = 37.17$ $ \Gamma = 2870$	$t = 37.17$ $ \Gamma = 4839$

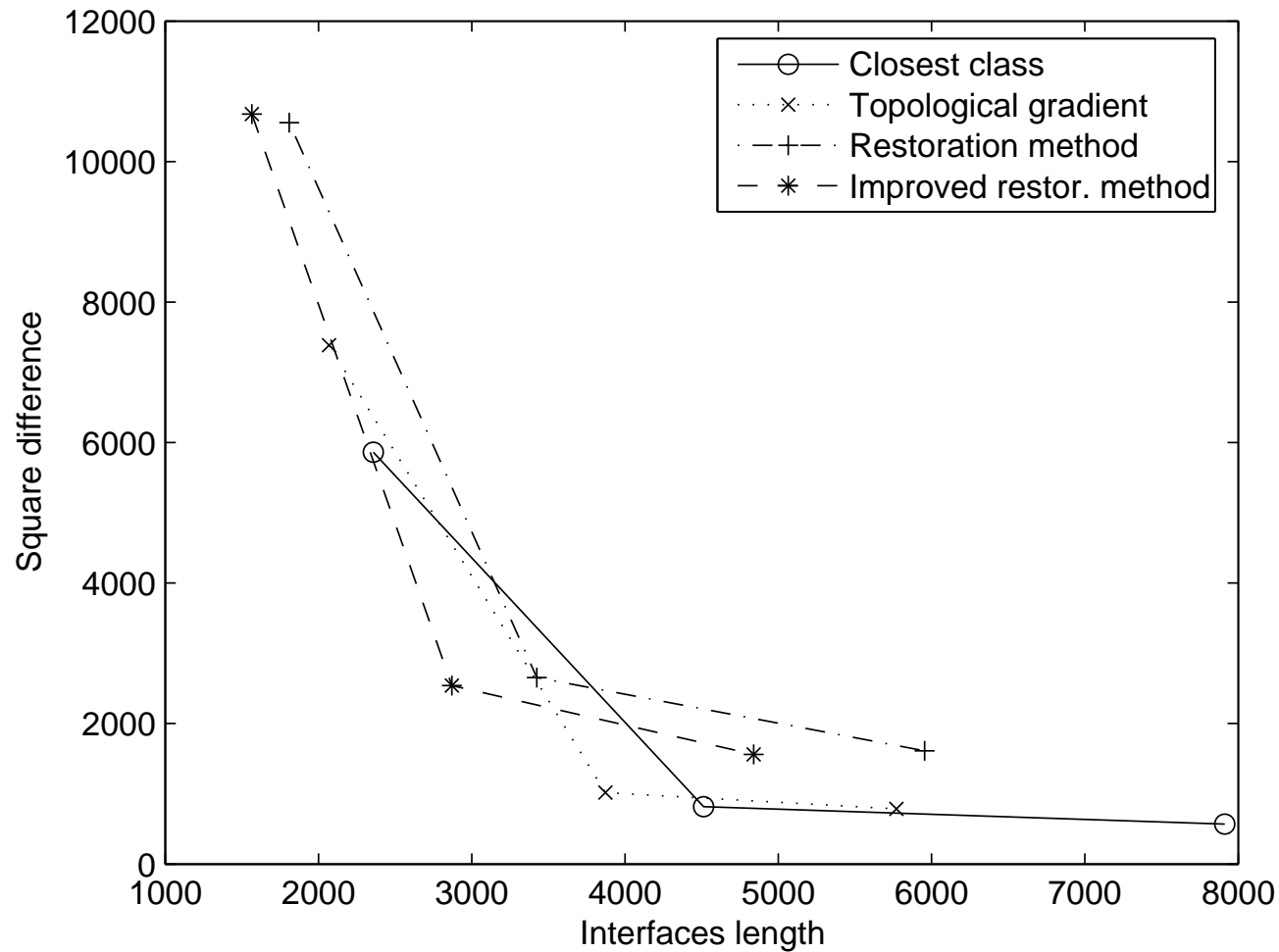


FIG. 11 – Square difference between the classified images and the original image versus the length of the interfaces for the different algorithms, and for 2, 3 and 5 classes.

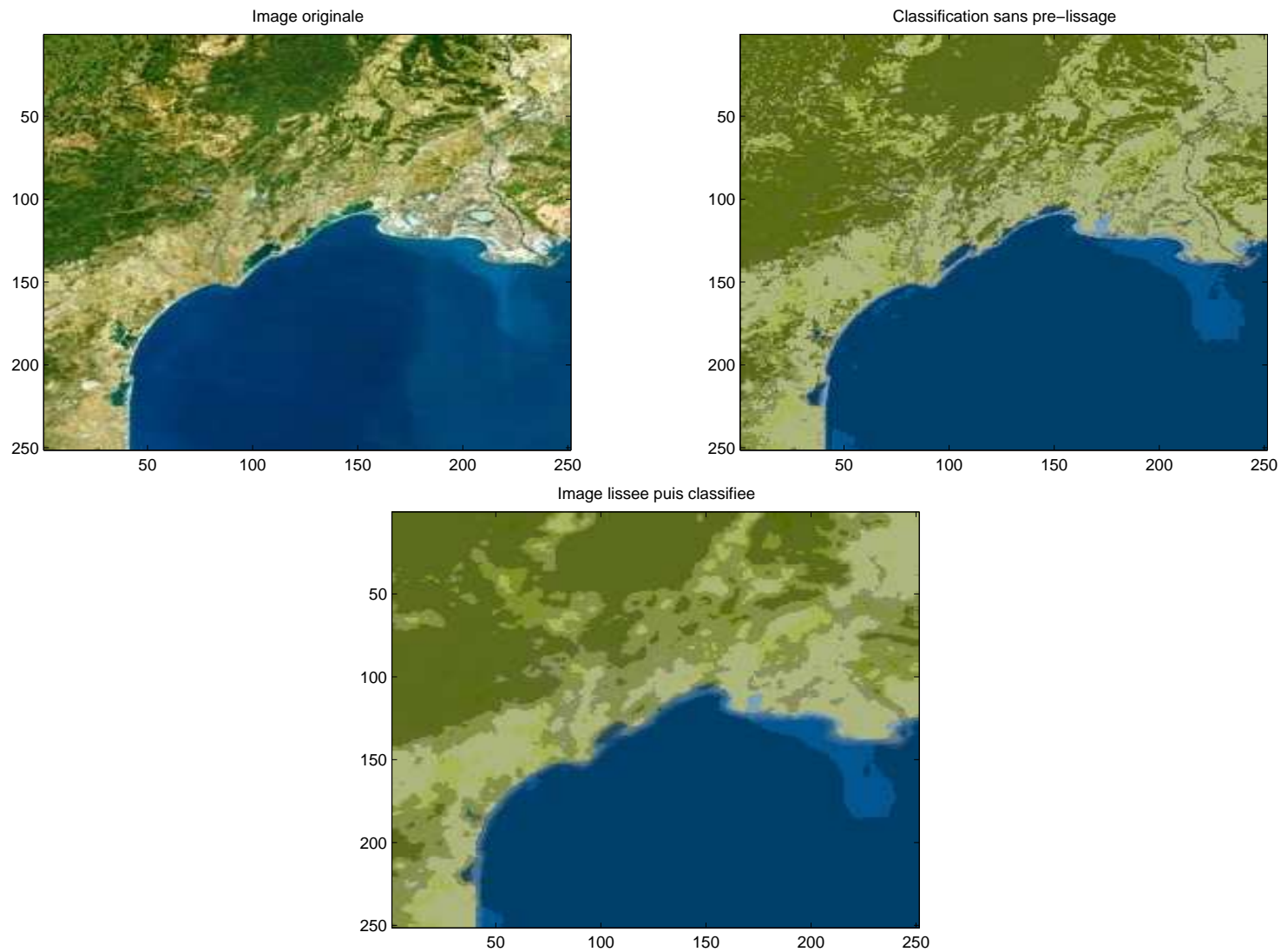


FIG. 12 – Top : original image (left), unregularized classified image (right) ; Bottom : regularized and classified image.



RESTORATION OF 3D-IMAGES, OR MOVIES



Algorithm



- **Initialization** : $c = c_0$.
- Computation of u_0 and v_0 , solutions of the **direct** and **adjoint** problems, using a preconditioned conjugate gradient method (preconditioner = discrete cosine transform).
- Computation of the **3×3 matrix M** and its lowest eigenvalue λ_{min} at each point of the domain.
- $c_1 = \begin{cases} \varepsilon_c & \text{if } x \in \Omega \text{ such that } \lambda_{min} < \alpha < 0, \quad \varepsilon_c > 0 \\ c_0 & \text{elsewhere.} \end{cases}$
- Calculation of u_1 solution to the **perturbed** problem with $c = c_1$, still using a PCG method (preconditioned by a DCT).
- u_1 is the **restored movie**.



Small resolution movies



- Original movie
- Noised movie
- Restored movie
- All together

Size of the movie : 180×72 pixels, 30 frames.

↪ 388.000 points, 5 minutes.



Small resolution movies



- Original movie
- Noised movie
- Restored movie
- All together

Size of the movie : 288×176 pixels, 52 frames.

↪ 2.6 million points, 1 hour.



Small resolution movies



- Original movie
- Noised movie
- Restored movie
- All together

Size of the movie : 320×144 pixels, 100 frames.

↪ 4.5 million points, 2 hours.



Large resolution movies



- Original movie
- Noised movie
- Restored movie
- All together

Size of the movie : 512×288 pixels, 110 frames (24 fps).

↪ more than 16 million points, 12 hours.

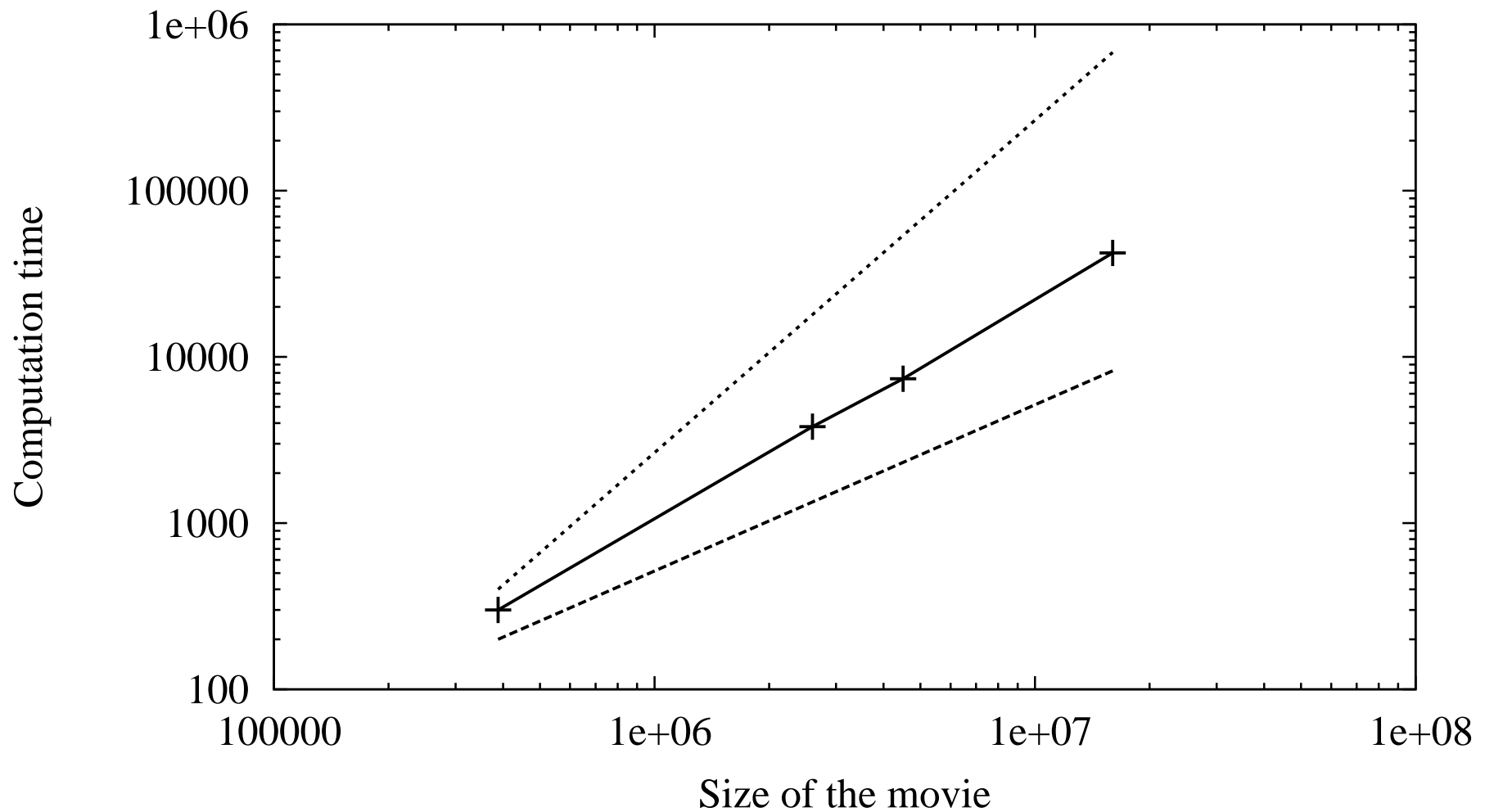


FIG. 13 – Computation time versus the size of the movie ; Topological gradient : $\mathcal{O}(n^{1.3})$.



Conclusion



- The **topological gradient** is very efficient.
- The image **restoration** (and **classification**) is performed in **only one iteration** : only 3 resolutions of a PDE are performed.
- The quality of the obtained images is good.
- Next step : color movies, restoration of a missing frame, ...



One missing image



- Original movie
- Restored movie
- All together



THANK YOU FOR YOUR ATTENTION

