Optimal stopping time formulation of adaptive image filtering

I. Capuzzo Dolcetta [*], R. Ferretti [†]

19.04.2000

Abstract

This paper presents an approach to image filtering based on an optimal stopping time problem for the evolution equation describing the filtering kernel. Well–posedness of the problem and convergence of fully discrete approximations are proved and numerical examples are presented and discussed.

1 Introduction

It is well–known (see [M]) that the filtering of a noisy image modelled by a function $x \to y_0(x)$ ($x \in R$, a rectangle in $\mathbb{R}^2$) can be performed by the convolution

$$y(t,x) = y_0(x) * h(t,x) \tag{1.1}$$

where

$$h(t,x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x|^2}{4t}} \tag{1.2}$$

Here, $t > 0$ plays the role of the scale factor of the filter and $2t$ is the second moment of h which is inversely related to the bandwidth.

From a pde's point of view, $y(t,x)$ given by (1.1) can be seen as the solution of the heat equation

$$y^{'}(t,x) = \Delta y(t,x) \quad , \quad t > 0 \; , \quad x \in R \tag{1.3}$$

[*]Dipartimento di Matematica, Università di Roma "La Sapienza", Piazzale A.Moro 2 - I 00185 Roma, Italy; e-mail: `capuzzo@mat.uniroma1.it`

[†]Dipartimento di Matematica, Università di Roma "Tor Vergata", via della Ricerca Scientifica - I 00133 Roma, Italy; e-mail: `ferretti@mat.uniroma2.it`

$$y(0,x) = y_0(x). \tag{1.4}$$

The determination of the scale of the filter (i.e. for which value of t the evolution (1.3),(1.4) should be stopped) is a crucial issue in this preliminary step of image processing. It is known (see [M]) that the size of the details in the filtered image $y(t,x)$ is of the order $\sqrt{t}$ but an optimal choice of t would also require in general the knowledge of the noise level.

The aim of this paper is to formulate the problem of determining in an adaptive way an "optimal" scale interval $[0,\theta^*]$ depending of course on the data $y_0(x)$ in the framework of optimal control theory. Some preliminary results in this direction have already been given by the authors in [CF].
More precisely, we consider $\theta \in [0,+\infty]$ as a control of stopping time nature acting on the dynamical system governed by (1.3),(1.4) and, as optimality criterion, we take the minimization with respect to $\theta \in [0,+\infty]$ of the performance index

$$J(y_0,\theta) = \int_0^\theta f(s,y(s,x))ds + \Phi(\theta,y(\theta,x)) \tag{1.5}$$

for any initial condition $y_0(x)$. Of course, the choice of functions f and Φ in (1.5) has to be suitably made in order to ensure the desired adaptive behavior of the filter. A basic example is given by

$$f(s,y(s,x)) \equiv c \tag{1.6}$$

$$\Phi(\theta,y(\theta,x)) = -\left(\int_R |y(\theta,x)-y_0(x)|^2 dx\right)^{\frac{\alpha}{2}} \tag{1.7}$$

for some positive constants c, α. With these choices, the integral term takes into account the computing cost whereas the stopping cost $\Phi(\theta,y(\theta,x))$ encourages filtering for small values of the scale factor. As we will show below, a careful balancing of the two terms is necessary in order to achieve the desired result.

We further observe that the approach proposed in the present paper is of quite general nature and therefore applies to different models in image smoothing. In the present paper, this approach will also be applied to the selective filtering model of Perona–Malik type proposed by Alvarez, Lions and Morel (see [ALM]), which is described by the evolution equation

$$y'(t,x) = div(g(|G_\sigma * \nabla_x y(t,x)|)\nabla_x y(t,x) \tag{1.8}$$

where

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|x|^2}{2\sigma^2}} \tag{1.9}$$

and $g : [0, +\infty) \to [0, +\infty)$ is a smooth function such that

$$g(0) = 1, \quad \lim_{s \to +\infty} g(s) = 0, \quad (sg(s))' \le 0 \quad \forall s > 0 \tag{1.10}$$

The above properties of function g imply that the evolution governed by (1.8) behaves roughly as the heat equation, hence producing a strong filtering effect, in those parts of R where $|\nabla_x y(t,x)|$ is small, whereas parts of R where $|\nabla_x y(t,x)|$ is large (i.e. the edges of the image) are almost kept fixed during the evolution, a pleasant feature as far as shape recognizing is concerned.

In the sequel $\| \cdot \|$ will denote the L^2–norm, and the shorthand notation $y(t)$ will possibly be used to denote $y(t,x)$.

The outline of the paper is as follows: section 2 is concerned with existence and qualitative behaviour of the optimal stopping time, section 3 treats convergence of fully discrete numerical approximations for the optimal stopping time problem. Lastly, we present in section 4 numerical results on synthetic and real images.

2 Statement of the stopping criterion

In this section we assume that the evolution of $y(t)$ is given either by (1.3) or by (1.8), with initial condition (1.4), posed on a bounded rectangular set R and with Neumann boundary conditions on ∂R (in the linear model this preserves the average value of y). Here, G_σ is a regularizing kernel given by (1.9) and we will make the standing assumptions (1.10) on g.

We recall that both (1.3) and (1.8) have a unique solution for any $y_0 \in L^2(R)$. More precisely, in both models the following regularity result:

$$y \in C([0,T]; L^2(R))$$

holds for the solution y. Moreover, a parabolic maximum principle implying the boundedness of solutions for bounded initial data applies to both models. These results are classical for the heat equation, whereas for (1.8) they are proved in [CLMC], [ALM].

We will mainly focus on stopping criteria of the special form (1.6), (1.7). Indeed, we can easily check that the optimal stopping time θ^* for this form of the cost has, at least qualitatively, the correct behaviour, that is, it increases at the increase of the noise level. In fact, if the minimum θ^* for (1.6), (1.7) is positive and finite, it requires that

$$\frac{d}{d\theta} \| y(\theta^*) - y_0 \|^\alpha = c; \tag{2.11}$$

on the other hand, $\| y(\theta) - y_0 \|^\alpha$ has a finite limit as $\theta \to \infty$, and the left–hand side of (2.11), at least for small times, increases at fixed time for increasing noise level on the initial image y_0. Hence, at the increase of the noise on the image, the optimal stopping time θ^* becomes larger as adaptivity would require. We will show in figure 1 a computed example comparing $J(\theta)$ at different noise levels.

Let us now examine the dependence of the problem on the parameters c and α.

Theorem 2.1 *Assume that $y(t)$ is given by (1.3) or (1.8), and that $c > 0$. Then, (1.5) admits a finite optimal stopping time θ^*. Moreover,*
i) If y_0 is piecewise smooth with jumps, then $\theta^* > 0$ if $\alpha < 4$, $\theta^* \geq 0$ if $\alpha \geq 4$ and J has a local minimum in $\theta = 0$ if $\alpha > 4$;
ii) If y_0 has bounded discontinuous gradient, then $\theta^* > 0$ if $\alpha < 2$, $\theta^* \geq 0$ if $\alpha \geq 2$ and J has a local minimum in $\theta = 0$ if $\alpha > 2$.

Proof. The continuity of J with respect to θ is a consequence of the regularity results for (1.3) and (1.8) mentioned before. The assumption $c > 0$, along with the uniform boudedness of $\|y(t) - y_0\|$, implies that the minimum of J is achieved at a finite time. Moreover, for the heat equation we have, in the case of discontinuous but piecewise smooth initial data,

$$\|y(t) - y_0\|^2 \sim C\sqrt{t} \quad (t \to 0^+) \tag{2.12}$$

We will check (2.12) in a simplified case, that is a step function in $\mathbb{R}$. This seems a reasonably general indication at least if the set of discontinuities of y_0 is rectifiable.

Let us assume therefore that

$$y_0(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0. \end{cases}$$

Taking into account the representation formula (1.1) for $y(t, x)$, we have

$$y(t, x) - y_0(x) =$$

$$= \int_{\mathbb{R}} (y_0(x - \xi) - y_0(x)) \frac{e^{-\frac{\xi^2}{4t}}}{\sqrt{4\pi t}} d\xi = \begin{cases} \int_{-\infty}^{x} \frac{e^{-\frac{\xi^2}{4t}}}{\sqrt{4\pi t}} & \text{if } x < 0 \\ -\int_{x}^{\infty} \frac{e^{-\frac{\xi^2}{4t}}}{\sqrt{4\pi t}} & \text{if } x > 0 \end{cases}$$

that is, after some computation,

$$(y(t, x) - y_0(x))^2 = \left(1 - F\left(\frac{|x|}{\sqrt{4t}}\right)\right)^2 \tag{2.13}$$

where

$$F(x) = \int_{-\infty}^{x} e^{-u^2} du.$$

Integrating (2.13), we obtain

$$\|y(t,x)-y_0(x)\|^2 = \int_{\mathbb{R}} \left(1 - F\left(\frac{|x|}{\sqrt{4t}}\right)\right)^2 dx =$$

$$4\sqrt{t}\int_0^\infty (1-F(u))^2 du.$$

Therefore, (2.12) implies that $J'(y_0, 0^+) = -\infty$ if $\alpha < 4$, $J'(y_0, 0^+) = c - C^2$ if $\alpha = 4$ and $J'(y_0, 0^+) = c$ if $\alpha > 4$. In the case of Lipschitz continuous initial data, by similar arguments,

$$\|y(t) - y_0\|^2 \sim Ct \quad (t \to 0^+) \tag{2.14}$$

and hence $J'(y_0, 0^+) = -\infty$ if $\alpha < 2$, $J'(y_0, 0^+) = c - C$ if $\alpha = 2$ and $J'(y_0, 0^+) = c$ if $\alpha > 2$.

Lastly, let us consider the case of the Perona–Malik model (1.8). Again by the regularity results in [ALM], [CLMC], for fixed y_0 and σ we have $|G_\sigma * \nabla_x y(t,x)| < C(y_0, \sigma)$ so that

$$0 < g(C(y_0,\sigma)) < g(|G_\sigma * \nabla_x y(t,x)|) \leq 1$$

and from standard arguments for parabolic equations we obtain the small–time behaviours (2.12), (2.14) for this model as well. □

Remark. The previous theorem shows that, although the optimal stopping time problem always admits a finite solution, this might be unsatisfactory under some conditions. If the image to be filtered, for example, is piecewise smooth but discontinuous, the criterion (1.5) would always perform filtering if $\alpha < 4$, even if the initial image y_0 is not noisy. On the other hand, if $\alpha > 4$, a search for the first local minimum of J would always yeld the stopping time $\theta^* = 0$. So if the image is discontinuous (this is the main situation of interest) it should be chosen $\alpha = 4$ and the constant c such as to obtain $\theta^* = 0$ for a non–noisy image, that is

$$c = C^2 = \frac{d}{dt}(\|y(t) - y_0\|^4)_{t=0}. \tag{2.15}$$

It could be noted that the result in (2.15) depends on the initial image y_0. In other terms, the value of the constant c should be computed performing (2.15) on a *typical* image to be filtered, that is, an image with the expected dimension of details and measure of the set of discontinuities. So a setup of the filter is still required, but it depends only on the characteristics of the image, not on the noise level (this would be the case, for example, in a Wiener filter).

3 Numerical approximation

In this section we will assume a state equation in the general form

$$\begin{cases} y'(t) = A(y(t)) \\ y(0) = y_0 \end{cases} \tag{3.16}$$

where y_0 and $y(t)$ belong to a Hilbert space H. For any initial state y_0, (3.16) is assumed to have a unique bounded continuous global solution $y \in C([0, +\infty[; H)$. Let now $H_n \subset H$ be a sequence of vector spaces, and $P_n : H \to H_n$ be a sequence of projections, and assume that $\dim H_n = k_n$ (with $k_n \to \infty$), and that for any $\bar{y} \in H$:

$$\lim_{n\to\infty} \|\bar{y} - P_n\bar{y}\| = 0. \tag{3.17}$$

A fully discrete approximation of (3.16) is in the form

$$\begin{cases} y_n^{k+1} = A_n(y_n^k) \\ y_n^0 = P_n y_0 \end{cases} \tag{3.18}$$

(here, $y_n^k \in H_n$ is intended to be an approximation of $y(k\Delta t)$) and a corresponding discretization for (1.5) is

$$J_{n,\Delta t}(P_n y_0, k) = \Delta t \sum_{j=0}^{k} f(j\Delta t, y_n^j) + \Phi(k\Delta t, y_n^k). \tag{3.19}$$

We point out that (3.16) is a general form which includes both models into consideration. In the usual implicit setting, the heat equation is discretized by

$$\frac{y_n^{k+1} - y_n^k}{\Delta t} = -R_n y_n^{k+1}$$

where R_n is the stiffness matrix associated to the Laplace operator (see [RT]). In this case, the function $A_n(y_n^k)$ in (3.18) is defined by

$$A_n(y_n^k) = (I + \Delta t R_n)^{-1} y_n^k.$$

On the other hand, Perona–Malik model (1.8) might be discretized (see [ALM]) by

$$\frac{y_n^{k+1} - y_n^k}{\Delta t} = -B_n(y_n^k) y_n^{k+1}$$

where $B_n(y_n^k)$ is a suitable, consistent approximation of the term

$$-div(g(|G_\sigma * \nabla y(k\Delta t)|)\nabla y(k\Delta t)),$$

which results in setting

$$A_n(y_n^k) = (I + \Delta t B_n(y_n^k))^{-1} y_n^k.$$

General results for semi–discrete approximations of a large class of non–quadratic control problems (including optimal stopping time problems) has been studied in [F1]–[F2]. We will now extend this approach to fully discrete approximations.

Approximation (3.18) is assumed to be convergent and uniformly stable, in the sense that, for all $k \in N$, $t \in [k\Delta t, (k+1)\Delta t)$ and $y_0 \in H$:

$$\|y(t) - y_n^k\|_H \leq C(y_0, n, \Delta t)(1+t)^{-\beta} \tag{3.20}$$

with $\lim_{n,\Delta t} C(y_0, n, \Delta t) = 0$, and $\beta > 1$ (we note that this limit is intended for $n \to \infty$, $\Delta t \to 0$ on the couples $(n, \Delta t)$ which satisfy the convergence conditions for the fully discrete scheme).

Given the evolution equation (3.16) and the initial state y_0, the optimal control problem into consideration is to find a stopping time $\theta^* \in [0, +\infty]$ minimizing the cost (1.5), assuming the boundedness of f and the (uniform in t) local Lipschitz continuity of both f and Φ:

$$|f(t, y_1)| \leq M_f \ , \quad |f(t, y_1) - f(t, y_2)| \leq L_f \|y_1 - y_2\| \tag{3.21}$$

$$|\Phi(t_1, y_1) - \Phi(t_2, y_2)| \leq L_\Phi(|t_1 - t_2| + \|y_1 - y_2\|) \tag{3.22}$$

for any $y_1, y_2 \in H$ and $t, t_1, t_2 \in R^+$, with $L_f = L_f(\|y_1\|, \|y_2\|)$, $L_\Phi = L_\Phi(\|y_1\|, \|y_2\|)$.

In the approximate version of this problem, given the fully discrete approximation (3.18) with initial state $P_n y_0$, one looks for a stopping time $k^* \Delta t$ minimizing the cost (3.19) (for simplicity, we drop the dependence of k^* on n and Δt).

The value functions for both the original and the approximate problem are defined as:

$$v(x) := \inf_{\theta \geq 0} J(x, \theta) \ , \quad v_{n,\Delta t}(x_n) := \inf_{k \geq 0} J_{n,\Delta t}(x_n, k). \tag{3.23}$$

Such value functions will be used in the proof of the following main convergence result.

Theorem 3.1 *Assume (3.20)–(3.23). Assume moreover that an optimal stopping time $k^* \Delta t$ exists for problem (3.18), (3.19). Then, for any $y_0 \in H$, $|J(y_0, k^* \Delta t) - v(y_0)| \to 0$ as $n \to \infty$, $\Delta t \to 0$.*

Proof. By the definition of v, $v_{n,\Delta t}$, for $y_0 \in H$ and any given $\varepsilon > 0$, it is possible to find two finite stopping times θ^ε, $k^\varepsilon \Delta t$ (depending on ε) such that:

$$v(y_0) \leq J(y_0, \theta^\varepsilon) \leq v(y_0) + \varepsilon \tag{3.24}$$

$$v_{n,\Delta t}(P_n y_0) \leq J_{n,\Delta t}(P_n y_0, k^\varepsilon) \leq v_{n,\Delta t}(P_n y_0) + \varepsilon. \tag{3.25}$$

We split the proof into three steps.
Step 1: $\limsup_{n,\Delta t} v_{n,\Delta t}(P_n y_0) \leq v(y_0)$. To prove this step, note that by the definition of $v_{n,\Delta t}$ one has:

$$v_{n,\Delta t}(P_n y_0) \leq \Delta t \sum_{j=0}^{h^\varepsilon} f(j\Delta t, y_n^j) + \Phi(h^\varepsilon \Delta t, y_n^{h^\varepsilon}) \tag{3.26}$$

with $h^\varepsilon = [\theta^\varepsilon / \Delta t]$. Adding the terms $\pm J(y_0, \theta^\varepsilon)$, and using (3.20), (3.21), (3.22) and (3.24), one obtains:

$$v_{n,\Delta t}(P_n y_0) \leq \Delta t \sum_{j=0}^{h^\varepsilon} f(j\Delta t, y_n^j) - \int_0^{\theta^\varepsilon} f(t, y(t))dt + \Phi(h^\varepsilon \Delta t, y_n^{h^\varepsilon}) -$$

$$-\Phi(\theta^\varepsilon, y(\theta^\varepsilon)) + v(y_0) + \varepsilon \tag{3.27}$$

which gives the bound

$$v_{n,\Delta t}(P_n y_0) - v(y_0) \leq \Delta t \sum_{j=0}^{h^\varepsilon} f(j\Delta t, y_n^j) - \int_0^{h^\varepsilon \Delta t} f(t, y(t))dt -$$

$$- \int_{h^\varepsilon \Delta t}^{\theta^\varepsilon} f(t, y(t))dt + L_\Phi[\Delta t + C(y_0, n, \Delta t)(1+t)^{-\beta}] + \varepsilon \leq$$

$$\leq C(y_0, n, \Delta t) \left[L_f \int_0^{h^\varepsilon \Delta t} (1+t)^{-\beta} dt + L_\Phi (1 + \theta^\varepsilon)^{-\beta} \right] +$$

$$+\Delta t(M_f + L_\Phi) + \varepsilon \tag{3.28}$$

thus proving the step.
Step 2: $\liminf_{n,\Delta t} v_{n,\Delta t}(P_n y_0) \geq v(y_0)$. By the definition of v,

$$v(y_0) \leq \int_0^{k^\varepsilon \Delta t} f(t, y(t))dt + \Phi(k^\varepsilon \Delta t, y(k^\varepsilon \Delta t)). \tag{3.29}$$

Adding $\pm J_{n,\Delta t}(P_n y_0, k^\varepsilon)$ and operating as before, one has:

$$v(y_0) - v_{n,\Delta t}(P_n y_0) \leq \int_0^{k^\varepsilon \Delta t} f(t, y(t))dt - \Delta t \sum_{j=0}^{k^\varepsilon} f(j\Delta t, y_n^j) +$$

$$+L_\Phi \|y(k^\varepsilon \Delta t) - y_n^{k^\varepsilon}\| + \varepsilon \le$$

$$\le C(y_0, n, \Delta t) \left[L_f \int_0^\infty (1+t)^{-\beta} dt + L_\Phi \right] + \varepsilon. \tag{3.30}$$

This completes the step and ensures that $\lim_{n,\Delta t} v_{n,\Delta t}(P_n y_0) = v(y_0)$.
Step 3: $\lim_{n,\Delta t} |J(y_0, k^* \Delta t) - v(y_0)| = 0$. It suffices to note that

$$|J(y_0, k^* \Delta t) - v(y_0)| \le$$

$$\le |J(y_0, k^* \Delta t) - J_n(P_n y_0, k^*)| + |v_{n,\Delta t}(P_n y_0) - v(y_0)|$$

and to apply the same arguments of step 2, noting that $v_{n,\Delta t}(P_n y_0) \to v(y_0)$. □

Remark. It is clear that in practice the space discretization step cannot vanish in image processing, since it naturally comes from a previous discretization of fixed resolution. Nevertheless, theorem 3.1 shows conditions under which results obtained by different discretization steps can give comparable results (since they converge to the same limit problem).

4 Numerical tests

This section provides some numerical examples. We present in figure 2 the original and noisy images for all tests. We also present for each test the results of processing with optimal stopping time θ^*, and (for comparison) with $\theta^*/2$ and $2\theta^*$ with both heat and Perona–Malik equations. The parameters of the stopping criterion are $\alpha = 4$, $c = 600$ for the linear filter and $c = 25$ for the nonlinear filter. They have been determined for both models as shown in the remark to theorem 2.1, and have been kept constant in all numerical tests with the same model. The resolution is 150×150 pixels.

To give an idea about the behaviour of the filter, we plot in figure 1 $J(y_0, \theta)$ vs θ at different noise levels for an initial image as in test 2. The uppermost curve (whose minimum is at $\theta = 0$) refers to a zero noise level, and the noise increases going downwards. It is apparent that at the increase of the term $\|y(t) - y_0\|^4$ the minimum of the curve is shifted away from the origin.

Figure 1. Plot of $J(y_0, \theta)$ vs θ for y_0 as in test 2 at different noise levels.

Test 1. The first test image is a dark square in lighter background. In figures 3 and 4 processed images with respectively the heat equation and the Perona–Malik equation are presented. In the first case, they present a remarkable diffusion of the edges which is caused by the use of the heat kernel. This effect is considerably reduced by the Perona–Malik model. However, in both cases the optimal stopping time actually seems to give a good compromise. Noise still has a high level at the stopping time $\theta^*/2$ for both models. Edges are completely diffused at time $2\theta^*$ in the case of the heat equation, whereas in the nonlinear case they are still kept (although the improvement in terms of noise with respect to the stopping time θ^*, is not worth doubling the processing time).

Test 2. This test presents a dark disk with a slight lightening towards its edge, on a uniform lighter background. The difficult point is to recover both the sharp edge of the disk and the more regular variation in its interior (note that this variation is almost unrecognizable in the noisy image). As in the previous example, linear filter produces (see figure 5) a high diffusion as soon as the smoothing of noise improves. The effect of nonlinear filter (figure 6) is remarkably good, and although the longer stopping time $2\theta^*$ gives a somewhat better result in terms of noise, the optimal time θ^* is a good compromise between quality of image and processing time.

Test 3. In this test a real photographic image has been used after the addition of artificial white noise as shown in figure 2. The result, obtained by nonlinear filtering, is good although it is clear that the stopping time θ^* is slightly too large. However, it must be noted that the parameters of the stopping criterion have been determined on a much simpler image (the disk of test 2). This results in a lower value for the constant c which in turn causes the stopping time to increase. On the other hand, determining the value of c on the same image, the optimal stopping gives a much better result which is shown in figure 8 compared to the previous one. The second value of c is about one order of magnitude larger than the first and this shows that it is important to "calibrate" the filter on an image of proper complexity; however this also shows that a different choice of this parameter has no dramatic effect, provided its order of magnitude is correct.

Figure 2. Initial (left) and noisy (right) images for tests 1–3.

Figure 3. Filtered image (at $t = \theta^*$, $t = \theta^*/2$, $t = 2\theta^*$) for test 1, linear filter.

Figure 4. Filtered image (at $t = \theta^*$, $t = \theta^*/2$, $t = 2\theta^*$) for test 1, nonlinear filter.

Figure 5. Filtered image (at $t = \theta^*$, $t = \theta^*/2$, $t = 2\theta^*$) for test 2, linear filter.

Figure 6. Filtered image (at $t = \theta^*$, $t = \theta^*/2$, $t = 2\theta^*$) for test 2, nonlinear filter.

Figure 7. Filtered image (at $t = \theta^*$, $t = \theta^*/2$, $t = 2\theta^*$) for test 3, nonlinear filter.

Figure 8. Filtered image (at $t = \theta^*$, $c = 25$ and $c = 210$) for test 3, nonlinear filter.

References

[ALM] L. Alvarez, P. L. Lions and J. M. Morel, *Image selective smoothing and edge detection by nonlinear diffusion II*, SIAM J. Num. Anal., **29** (1992), 845–866.

[BL] A. Bensoussan and J. L. Lions, *Impulse control and quasi–variational inequalities*, Gauthier–Villars, Paris, 1984

[CF] I. Capuzzo Dolcetta and R. Ferretti, *On an optimal stopping time formulation of adaptive signal filtering*, proceedings of the conference SSPCS '97, Pereslavl–Zalessky, 1997

[CLMC] F. Catté, P. L. Lions, J. M. Morel and T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Num. Anal., **29** (1992), 182–193.

[F1] R. Ferretti, *Internal approximation schemes for optimal control problems in Hilbert spaces*, J. Math. Sist. Est. Contr., **7** (1997), 115–118

[F2] R. Ferretti, *Dynamic programming techniques in the approximation of optimal stopping time problems in Hilbert spaces*, proceedings of the conference IFIP 7.2 "Boundary control and variations", Rabat, 1993

[M] D. Marr, *Vision*, Freeman, New York, 1982

[RT] P. A. Raviart and J. M. Thomas, *Introduction a l'analyse numérique des équations aux derivées partielles*, Masson, Paris, 1983