IMAGE PROCESSING THROUGH REACTION
COMBINED WITH NONLINEAR DIFFUSION

G.-H. COTTET AND L. GERMAIN

ABSTRACT. We propose a method based on nonlinear diffusion and reaction for edge detection and contrast enhancement in image processing. We prove that the mathematical model is well posed and show numerically that the processed image can be observed on the asymptotic state of its solution. We illustrate the methods on test images and show on medical images how it can help to draw contours and detect one-dimensional coherent signals.

1. INTRODUCTION

The goal of this paper is to present new techniques based on partial differential equations for image processing. PDEs have been known for a long time as a useful tool in image processing, and pure diffusion equations are indeed underlying classical filtering techniques. The challenging aspect of image processing is to design methods which can filter selectively the noise without affecting the interesting features of the original image. In [7], Perona and Malik proposed a method toward that goal. Their idea is to penalize the diffusion where the gradients of the signal are large. This is desirable in order to keep the edges of the signal, but obviously, if the noise is to be removed, that principle must be modified. Later Alvarez et al. [1] derived a method based upon the same principle which overcomes that difficulty and is mathematically well posed. They show that by properly choosing the time of integration, their method is able to remove a large amount of noise without introducing a significant smearing at the edges of the signal. The choice of this integration time can be made in terms of the minimal scale which is to be kept in the original image. For a different mathematical approach, also based on partial differential equations, we refer to [6].

The method we propose here is different in spirit. It is based on reaction diffusion equations. The idea behind these models is to combine diffusion for noise filtering and reaction for contrast enhancement. On small time scales simple reaction diffusion models indeed perform well. They average out the noise on a scale which is related to the diffusion coefficient; through the reaction term, the result is then compared to a threshold and either amplified or damped out. However, for longer times these interesting features disappear: the edges of

the signal start shrinking with a velocity proportional to their mean curvature, and the image eventually completely vanishes.

Our goal was to derive a method where the processed image could be observed on the asymptotic state of the solution of the mathematical model. We believe that this goal is rather natural in view of the importance of attractors in biological systems. One possible direction would be to replace scalar equations by systems of reaction diffusion equations. These models are more likely to produce nontrivial equilibrium states which could be our desired asymptotic images. Some results exist in specific situations (see [8]). However, if one wishes to stay within the mathematically simpler framework of single reaction diffusion equations, then some nonlinearity must be introduced in the diffusion. This approach is natural to some extent, in view of some formal links which exist between reaction-diffusion equations and neural networks (see [3]). In this reference we trivially observe that neural networks whose transitions are governed by translation-invariant connections can be considered as space-time discretization of reaction-diffusion equations. The diffusion scale is, for obvious reasons, related to the range of the connections, while the shape of the reaction term is, up to a linear correction, given by the gain function which correlates input and output of a neuron. It is worth pointing out that in this analogy the uniqueness of an invariant measure (in a statistical physics framework) attached to the neural network, which is the uniform measure in the case of a synaptic matrix invariant with respect to translation, reflects the fact that the asymptotic state for a reaction diffusion is constant when the diffusion is isotropic. However, in real life, neural networks learn at the same time they process information; as a result, during a transient stage, synaptic weights move from a uniform distribution into one which is strongly affected by the correlations of the activities of the neurons. In other words, neurons which have experienced correlations in their output during a transient stage, become strongly connected. If one wishes to pursue the analogy with reaction-diffusion in this case, then nonlinearities have somehow to enter in the diffusion term.

For the construction of the kind of nonlinearity we have in mind, we used, as in [1], the fact that it is very important in image processing to have a good resolution of the edges. More specifically, in many applications in medical imaging it is desirable to be able to detect one-dimensional objects, which would be considered as noisy signals if isotropic diffusion were used. A natural idea to build diffusion operators which could overcome this difficulty is to determine at each point of the image the gradient of the signal and to diffuse only on the orthogonal direction. Obviously, this program cannot be strictly implemented, since it would lead to a diffusion which would vanish everywhere. So we propose instead to compute the direction of diffusion through a regularized version of the signal; this gives the following diffusion operator:

$$\Delta_\varepsilon u = \operatorname{div}([A_\varepsilon(u)][\nabla u]),$$

where $A_\varepsilon(u)$ denotes the orthogonal projection onto the direction orthogonal to the gradient of a regularization u_ε of u. This actually amounts to calculating the derivatives of u on two different scales: one to compute ∇u and one for the diffusion matrix. This diffusion term can also be interpreted as follows: it will be small if $\nabla u_\varepsilon \simeq \nabla u$; this is the case, in particular, when u has support around a one-dimensional curve, provided its curvature does not vary too much on a

scale of the order of ε. Therefore, it can be expected that our diffusion will not affect smooth one-dimensional objects, and ε appears again as a scale parameter which will determine the minimal size of the details that one wishes to keep in the image. On the other hand, if the signal is noisy, then the directions of ∇u_ε and ∇u will not show any coherence, and Δ_ε will act as an isotropic diffusion. In §2 we prove, along with the well-posedness of the initial-boundary value problem, the attractivity of piecewise constant states in the case where spatial derivatives are replaced by finite differences and the diffusion matrix is only computed at time 0. In view of this attractivity property, our method can also be discussed from the point of view of quantification. For a description and an algorithmic approach of the quantification problem we refer to [5]. We give in §3 numerical examples based on test images which illustrate these quantification properties of our method; we also give examples of medical images, and one, in particular, where it is crucial to detect one-dimensional objects.

2. The mathematical model

We are dealing here with 2-dimensional images, but the same tools can be used in any dimension. We denote by Ω the domain $(0, 1) \times (0, 1)$ and by Γ its boundary. Although boundary conditions do not really matter here, we will focus for simplicity on homogeneous Dirichlet boundary conditions. Let ζ be a smooth cutoff function with support in the ball of radius 1, and ε a (small) positive parameter. Classically we write $\zeta_\varepsilon(x) = \varepsilon^{-2}\zeta(x/\varepsilon)$. If $u \in L^2(\Omega)$, we extend u in the whole plane by setting

$$\bar{u}(x) = \begin{cases} u(x) & \text{if } x \in \Omega \text{ and } d(x, \Gamma) > 2\varepsilon, \\ 0 & \text{otherwise,} \end{cases}$$

and we set $u_\varepsilon = \bar{u} * \zeta_\varepsilon$. For the reaction term, we will deal with functions which are of class C^1 and satisfy

$$f(\pm 1) = 0, \quad xf(x) > 0 \quad \text{for } x \neq 0,$$

but the analysis below applies as well to functions having more zeros, a possibility of practical interest if one wishes to obtain asymptotic states where more than two grey levels are extracted. We finally denote by A_ε the nonlinear operator whose value is the 2×2 matrix

$$A_\varepsilon(u)_{i,j} = \frac{\tilde{\partial}_i u_\varepsilon \tilde{\partial}_j u_\varepsilon}{|\nabla u_\varepsilon|^2 + \varepsilon^2},$$

where $\tilde{\partial}_1 = \partial/\partial x_2$ and $\tilde{\partial}_2 = -\partial/\partial x_1$. Up to the term ε^2 in the denominator, whose role is to avoid singularities in the derivatives of this matrix, $A_\varepsilon(u)$ is the orthogonal projection onto the direction which is perpendicular to the gradient of u_ε. Our reaction diffusion model is then

$$(2.1) \qquad \frac{\partial u}{\partial t} - \sigma\varepsilon^2 \operatorname{div}([A_\varepsilon(u)][\nabla u]) = f(u) \quad \text{in } \Omega,$$

$$(2.2) \qquad u(\cdot, 0) = u_0 \quad \text{in } \Omega,$$

$$(2.3) \qquad u = 0 \quad \text{on } \Gamma.$$

Note that although the solution of the above problem depends on ε, we denote it by u to simplify the notations. The key ingredients to prove the well-posedness of this problem are summarized in the following

Lemma 2.1. *Let* $\varepsilon > 0$, u *and* v *in* $L^\infty(\Omega)$. *The following assertions hold*:
(i) *if* $u \in L^\infty(\Omega)$, *then* $A_\varepsilon(u)_{i,j} \in C^\infty(\Omega)$ *for* $i, j \in \{1, 2\}$;
(ii) *there exists a constant* $C_\varepsilon > 0$ *such that*

$$\|A_\varepsilon(u) - A_\varepsilon(v)\|_{L^\infty} \leq C_\varepsilon \|u - v\|_{L^\infty};$$

(iii) *for* $\phi, \psi \in (L^\infty(\Omega))^2$

$$\langle [A_\varepsilon(u)]\phi, \psi \rangle = \frac{\langle \operatorname{curl} u_\varepsilon, \phi \rangle \langle \operatorname{curl} u_\varepsilon, \psi \rangle}{|\nabla u_\varepsilon|^2 + \varepsilon^2}.$$

The notation $\operatorname{curl} v$ is used here to denote the vector of components $(\partial_2 v, -\partial_1 v)$. The proof of these assertions is left to the reader. Assertions (i) and (iii) are consequences of the regularization step involved in the construction of $A_\varepsilon(u)$, while assertion (ii) results from simple algebraic calculations. Observe in particular that choosing $\phi = \psi$ gives the positivity of $A_\varepsilon(u)$. With that we can prove

Theorem 2.1. *Let* $u_0 \in H_0^1(\Omega)$ *and* $\varepsilon > 0$. *The system* (2.1)–(2.3) *has at least one solution in* $H_0^1(\Omega)$. *If* $-1 \leq u_0 \leq 1$, *then* $-1 \leq u(\cdot, t) \leq 1$ *for all time* $t > 0$.

Proof. Throughout the following, the parameter ε is fixed, and for simplicity we will drop it in the notations to avoid confusion with the vanishing parameter Δt, which will be introduced (however Δ will still denote the usual Laplace operator). We will exhibit a solution to (2.1)–(2.3) as a limit of solutions of approximate linear problems. Let $\Delta t > 0$ and (u^n) be the sequence obtained by the following implicit time discretization of (2.1)–(2.3):

$$(2.4) \quad u^{n+1} - \sigma\varepsilon^2 \Delta t \operatorname{div}([A(u^n)][\nabla u^{n+1}]) - \Delta t^2 \Delta u^{n+1} = u^n + \Delta t f(u^n) \quad \text{in } \Omega,$$

$$(2.5) \qquad\qquad\qquad\qquad u^{n+1} = 0 \quad \text{on } \Gamma.$$

In view of the assertion (iii) of Lemma 2.1 and the coercivity of the Laplacian in $H_0^1(\Omega)$, the above problems are uniformly parabolic in $H_0^1(\Omega)$, and u^{n+1} is therefore well defined. Moreover, since $f(+1) = f(-1) = 0$, if $|u^n| \leq 1$ and $\Delta t \leq 1/\sup|f'|$, then clearly $|u^n + \Delta t f(u^n)| \leq 1$. By the maximum principle, we thus have $|u^{n+1}| \leq 1$. If we denote by $u_{\Delta t}$ the piecewise linear (with respect to time) function such that $u_{\Delta t}(\cdot, t_n) = u^n$, we conclude that

$$u_{\Delta t} \text{ is bounded in } L^\infty(0, T; L^\infty(\Omega)).$$

To pass to the limit in the reaction term, we need some compactness in L^2. This results from a bound of u^n in $H_0^1(\Omega)$ that we now derive. By differentiating (2.1), if $v^n = \nabla u^n$, we obtain the following system in v:

$$(2.6) \quad \begin{aligned} v^{n+1} - \sigma\Delta t \operatorname{div}([A(u^n)][\nabla v^{n+1}]) - \Delta t^2 \Delta v^{n+1} \\ = v^n + \Delta t[f'(u^n)v^n + \sigma R(u^n, v^{n+1})], \end{aligned}$$

where $R = (R_i)$, with

$$R_i(u, v) = \sum_j \frac{\partial}{\partial x_j} \left(\sum_k \frac{\partial A_{jk}(u)}{\partial x_i} v_k \right).$$

Observe that the maximum principle does not apply directly to (2.6), because R is mixing derivatives of order 1 of different components of v. To overcome this

difficulty, we observe that if v satisfies $\partial v_i / \partial x_j = \partial v_j / \partial x_i$ (which is the case because of our definition of v), we can apply the following integration by parts (recall that A vanishes in a neighborhood of Γ, so there is no contribution from the boundary):

$$
\begin{aligned}
\int \langle R(u, v), v \rangle \, dx &= \int \sum_{i,j,k} \frac{\partial A_{jk}(u)}{\partial x_i} v_k \frac{\partial v_i}{\partial x_j} \\
&= \int \sum_{i,j,k} \frac{\partial A_{jk}(u)}{\partial x_i} v_k \frac{\partial v_j}{\partial x_i} \\
&= - \int \sum_{i,j,k} \frac{\partial^2 A_{jk}(u)}{\partial x_i^2} v_k v_j - \int \sum_{i,j,k} \frac{\partial A_{jk}(u)}{\partial x_i} \frac{\partial v_k}{\partial x_i} v_j .
\end{aligned}
$$

But exchanging j and k and using again the symmetry of the Jacobian matrix of v, along with the symmetry of A, we see that the last piece of the integral above is equal to $\int \langle R(u, v), v \rangle \, dx$. Therefore, we obtain

$$
\int \langle R(u, v), v \rangle \, dx = -\frac{1}{2} \int \sum_{i,j,k} \frac{\partial^2 A_{jk}(u)}{\partial x_i^2} v_k v_j .
$$

Now multiply (2.6) by v^{n+1} and integrate by parts to get

$$
(1 - C\sigma \Delta t \| A(u^n) \|_{W^{2,\infty}(\Omega)}) \| v^{n+1} \|_{L^2(\Omega)}^2 \leq \| v^{n+1} \|_{L^2(\Omega)} (1 + C\Delta t) \| v^n \|_{L^2(\Omega)}.
$$

We already have that u^n is bounded, which, in view of Lemma 2.1, assertion (i), implies that $\| A(u^n) \|_{W^{2,\infty}(\Omega)} \leq C$. This finally gives, if Δt is small enough,

$$
\| v^{n+1} \|_{L^2(\Omega)} \leq (1 + C\Delta t) \| v^n \|_{L^2(\Omega)}
$$

and therefore

$$
\| v^n \|_{L^2(\Omega)} \leq C(T) \| v^0 \|_{L^2(\Omega)} \quad \text{for } t_n \leq T.
$$

In other words,

$$
u_{\Delta t} \text{ is bounded in } L^\infty(0, T; H_0^1(\Omega)).
$$

By means of (2.4) it is also readily seen that

$$
\frac{\partial u_{\Delta t}}{\partial t} \text{ is bounded in } L^\infty(0, T; H^{-1}(\Omega)).
$$

From standard compactness arguments we thus obtain the existence of u in $L^\infty(0, T; H_0^1(\Omega))$ such that, for a subsequence,

$$
u_{\Delta t} \to u \quad \text{in } L^2(0, T; L^2(\Omega)).
$$

This already ensures that $f(u_{\Delta t}) \to f(u)$ in the sense of distributions in $[0, T) \times \Omega$. Next, since all the derivatives of $u_{\Delta t} * \zeta$ are bounded, we also get, for instance,

$$
u_{\Delta t} * \zeta \to u_\varepsilon * \zeta \quad \text{in } L^2(0, T; W_0^{1,4}(\Omega)),
$$

which implies

$$
[A(u_{\Delta t})][\nabla u_{\Delta t}] \to [A(u)][\nabla u]
$$

still in the sense of distributions. From this it is not difficult to check that u is a solution to the original equation. $\quad \square$

Regarding uniqueness, it seems that more smoothness has to be required (namely $u \in H^2(\Omega)$). We feel that this smoothness is not natural in the context of image processing, where one has basically to deal with signals which are only bounded. For the remainder of this discussion, we prefer to focus on stability and asymptotic results. We now give two results illustrating on the one hand the possibility of having almost steady piecewise constant states and, on the other hand, the attractivity of these solutions. We first show that anisotropic diffusion is small for piecewise constant u, provided the curvature of the boundary separating the domains were u takes constant values is not too large and does not change too abruptly.

Theorem 2.2. *Let D be a domain whose boundary ∂D is of class C^3 and γ a normal parametrization of ∂D. Assume that D is locally on one side of ∂D and let $\bar{u}$ be the characteristic function of D. Then we have for ε small enough*

$$[A_\varepsilon(\bar{u})][\nabla \bar{u}] = \alpha \delta_{\partial D},$$

where α is a bounded vector function defined on ∂D which satisfies

$$(2.7) \qquad \|\alpha\|_{L^\infty} \leq C\varepsilon^2.$$

If u is such that

$$\|u - \bar{u}\|_{W^{1,\infty}(D)} + \|u - \bar{u}\|_{W^{1,\infty}(\Omega - D)} \leq \kappa < 1/4,$$

then $[A_\varepsilon(u)][\nabla \bar{u}] = \alpha \delta_{\partial D}$ with

$$(2.8) \qquad \|\alpha\|_{L^\infty} \leq C\varepsilon[\varepsilon + \kappa].$$

In the above estimates, C is a positive constant depending on ∂D (namely the second- and third-order derivatives of γ) and ζ.

Proof. For simplicity we give the proof in the case when ζ is the characteristic function of the ball with center 0 and radius 1. Let ν and τ denote respectively the normal (directed outward) and the tangent along ∂D. We have $\nabla \bar{u} = \nu \delta_{\partial D}$ and thus

$$[A_\varepsilon(\bar{u})][\nabla \bar{u}] = \frac{\partial \bar{u}_\varepsilon / \partial \tau}{|\nabla \bar{u}_\varepsilon|^2 + \varepsilon^2} \delta_{\partial D}.$$

Let $x \in \partial D$. We can assume, without restricting generality, that $x = 0$, that at this point the tangent is horizontal and ∂D is parametrized by the first coordinate. So we can write

$$\frac{\partial \bar{u}_\varepsilon}{\partial \tau}(x) = \int \bar{u}(y) \frac{\partial \zeta_\varepsilon}{\partial y_1}(x - y)\,dy = \varepsilon^{-2} \int_{-\varepsilon}^{\varepsilon} \bar{u}(\alpha(y_2), y_2) - \bar{u}(-\alpha(y_2), y_2)\,dy_2,$$

where $\alpha(y_2) = \sqrt{\varepsilon^2 - y_2^2}$. The parameter y_2 gives a nonzero contribution to the integral only if the points $(\alpha(y_2), y_2)$ and $(\alpha(-y_2), y_2)$ lie on opposite sides of ∂D, in which case its contribution is 1, that is, if

$$\gamma(-\alpha(y_2)) \leq y_2 \leq \gamma(\alpha(y_2)).$$

But an expansion to the order 3 of γ shows that

$$|\gamma(\alpha(y_2)) - \gamma(-\alpha(y_2))| \leq \varepsilon^3 \sup |\gamma'''|,$$

whence we get

$$(2.9) \qquad \left| \frac{\partial \bar{u}_\varepsilon}{\partial \tau}(x) \right| \leq \varepsilon \sup |\gamma'''|.$$

Next, integrating over y_1 yields

$$\frac{\partial \bar{u}_\varepsilon}{\partial \nu}(x) = \varepsilon^{-2} \int_{-\varepsilon}^{\varepsilon} \bar{u}(y_1, \alpha(y_1)) - \bar{u}(y_1, -\alpha(y_1)) \, dy_1.$$

Since D is on one side of ∂D, the quantity in the integral above has a constant sign and an expansion of γ to the order 2 shows that its absolute value is larger than 1 if $|y_1| \sup |\gamma''| \leq \varepsilon$, which implies that

$$(2.10) \qquad \left| \frac{\partial \bar{u}_\varepsilon}{\partial \nu}(x) \right| \geq \frac{C}{\varepsilon \sup |\gamma''|}.$$

Combining (2.9) and (2.10) finally gives (2.7). The proof of (2.8) goes along the same lines, with the additional observation that, if $(\alpha(y_2), y_2)$ and $(-\alpha(y_2), y_2)$ are on the same side of ∂D, then

$$|\bar{u}(\alpha(y_2), y_2) - \bar{u}(-\alpha(y_2), y_2)| \leq \varepsilon \kappa,$$

which leads to

$$\left| \frac{\partial \bar{u}_\varepsilon}{\partial \tau}(x) \right| \leq \varepsilon [\kappa + \varepsilon \sup |\gamma'''|].$$

As for (2.10), this inequality remains unchanged since, if $(y_1, \alpha(y_1))$ and $(y_1, -\alpha(y_1))$ lie on opposite sides of ∂D, we still have

$$|\bar{u}(y_1, -\alpha(y_1)) - \bar{u}(y_1, -\alpha(y_1))| \geq 2\kappa \geq 1/2. \quad \square$$

We point out that the above arguments can be easily modified to deal with domains which are locally contained between two curves, no matter how close these curves are to each other, provided they are smooth enough. This is an important feature in view of selective smoothing which would not affect one-dimensional objects. The result of Theorem 2.2 naturally raises the question of the behavior of our method in the presence of singularities, for instance, when the boundaries of the patches have corners. In this case it can be expected that the algorithm will first produce a smoothing of the edge, until its curvature reaches an acceptable value, at which time the diffusion will stop working.

To illustrate further that sufficiently coherent signals are not destroyed by our reaction-diffusion process, we will now prove that, at least for a semidiscretized and linearized version of our model, piecewise constant states are *almost steady-state* solutions for our model. To simplify, let us restrict our discussion to a specific finite element discretization. Let $h > 0$ and $\mathcal{T}_h$ be a uniform triangulation of Ω with triangles of side h. Let x_i be the nodes of this triangulation. Define $\mathcal{V}_h$ to be the space of continuous functions on Ω which vanish on Γ and are linear on each triangle. Finally, let $u_0 \in H_0^1(\Omega)$, let u_0^h be its projection on $\mathcal{V}_h$ and $u^h \in \mathcal{V}_h$ such that for all $\phi \in \mathcal{V}_h$

$$(2.11) \qquad \frac{d}{dt} \langle u^h, \phi \rangle + \sigma h^2 \int_\Omega \langle [A_\varepsilon(u_0^h)][\nabla u^h], \nabla \phi \rangle = \sum_i c_i f(x_i) \phi(x_i),$$

$$(2.12) \qquad u^h(\cdot, 0) = u_0^h.$$

A few comments are in order about this new problem. Observe first that the singular perturbation parameter ε has been replaced by h, which to some extent is equivalent since in practice ε and h are of the same order (the regularization

is done over a few pixels). More important is the fact that we reduce our problem to a linear one, by evaluating the diffusion matrix only at the initial step. In §3 we will see that in practice this linearization can be efficient.

Let D be a polygonal domain of Ω which is the union of triangles of $\mathcal{T}_h$. We denote by N the number of vertices of D and by $\bar{u}$ the function in $\mathcal{V}_h$ which is equal to 1 in D and -1 in $\Omega - D$. Then we can prove

Theorem 2.3. *If u_0^h satisfies $\|u_0^h - \bar{u}\|_{W^{1,\infty}(D)} + \|u_0^h - \bar{u}\|_{W^{1,\infty}(\Omega-D)} \le 1/2$, there exists a constant $C > 0$ such that, if $\sigma N\sqrt{\varepsilon h^{-1}} \le C$, then*

$$(2.13) \qquad \limsup_{t \to \infty} \|u^h - \bar{u}\|_{L^2(\Omega)} \le C\sigma N\sqrt{\varepsilon h}.$$

Proof. Let $T^* = \max\{t > 0;\ \|u^h - \bar{u}\|_{L^\infty(\Omega)} \le 1/2\}$,

$$\lambda_1 = \inf\left\{ \left| \frac{f(v)}{v-1} \right|\ ;\ v \in [-1,\, 1],\ |1-v| \le 1/2 \right\},$$

$$\lambda_2 = \inf\left\{ \left| \frac{f(v)}{v+1} \right|\ ;\ v \in [-1,\, 1],\ |1+v| \le 1/2 \right\},$$

and $\lambda = \min(\lambda_1,\, \lambda_2)$. For $t \le T^*$ we have

$$(u^h(x_i) - \bar{u}(x_i))(f(u^h(x_i)) \le \lambda |u^h(x_i) - \bar{u}(x_i)|^2,$$

so that, setting $\phi = e = u^h - \bar{u}$ in (2.11), we obtain

$$(2.14) \qquad \begin{aligned} &\frac{d}{dt}\|e\|_{L^2(\Omega)} + \sigma h^2 \int_\Omega \langle [A_\varepsilon(u_h^0)][\nabla e],\, \nabla e \rangle + C\lambda\|e\|_{L^2(\Omega)} \\ &\le \sigma h^2 \int_\Omega \langle [A(u_\varepsilon^0)][\nabla\bar{u}],\, \nabla e \rangle\, dx. \end{aligned}$$

We now bound the integral in the right-hand side. Let V^h be the union of the triangles which intersect ∂D and V_h^ε, the part of V^h consisting of the points which are at a distance less than ε from one of the vertices of D. Then $\nabla\bar{u}$ is bounded by $2h^{-1}$ and clearly vanishes outside V^h. In view of our assumption on $u_0^h - \bar{u}$, the argument used in the proof of (2.7) shows that $[A_\varepsilon(u_0^h)][\nabla\bar{u}]$ has support in V^h and satisfies

$$[A_\varepsilon(u_0^h)][\nabla\bar{u}] \le \begin{cases} \dfrac{C}{h} & \text{in } V_\varepsilon^h, \\[2mm] \dfrac{C\varepsilon}{h} & \text{in } V^h - V_\varepsilon^h. \end{cases}$$

Therefore, we can write

$$\int_\Omega \langle [A_\varepsilon(u_h^0)][\nabla\bar{u}],\, \nabla e \rangle\, dx \le C\kappa\frac{\varepsilon}{h}\int_{V^h} |\nabla e(x)|\, dx + \frac{1}{h}\int_{V_\varepsilon^h} |\nabla e(x)|\, dx$$

$$\le \frac{CN\sqrt{\varepsilon}}{\sqrt{h}}\|\nabla e\|_{L^2(\Omega)} \le \frac{CN\sqrt{\varepsilon}}{h\sqrt{h}}\|e\|_{L^2(\Omega)}.$$

By the positivity of A we then deduce from (2.14) the following inequality:

$$\frac{d}{dt}\|e\|_{L^2(\Omega)} + C\lambda\|e\|_{L^2(\Omega)} \le C\sigma N\sqrt{h\varepsilon}, \qquad t \in [0,\, T^*].$$

By Gronwall's lemma this yields

$$\|e(t)\|_{L^2(\Omega)} \le \|e(0)\|_{L^2(\Omega)} \exp(-C\lambda t) + C\sigma N\sqrt{h\varepsilon}(1 - \exp(-C\lambda t)), \qquad t \in [0,\, T^*].$$

To conclude, it remains to deduce from the inverse inequality $\|e\|_{L^\infty} \leq h^{-1}\|e\|_{L^2}$ that, if for a proper positive constant C, $\sigma N\sqrt{\varepsilon}h^{-1} \leq C$, then $T^* = +\infty$. $\quad\square$

This result is obviously very incomplete for, unless the triangulation is adapted to the particular patterns present in the image, it is very likely that, when $h \to 0$, the number of vertices will increase. However, it illustrates, like the estimates proved in Theorem 2.2, how strongly our model contrasts with isotropic reaction-diffusion, and it indicates that signals which are constant on smooth patches are stable. Moreover, the role of the scale parameter ε is to select those asymptotic states which are in some sense coherent on a scale of the order of ε.

3. NUMERICAL EXPERIMENTS

We first describe how the nonlinear equation (2.1) has been discretized. The cutoff function ζ has been chosen to be the classical hat function (although this function has not the smoothness required by Theorem 2.1), and the convolution u_ε is approximated by a midpoint quadrature rule. The divergence and gradient involved in the diffusion were approximated by using the following formulas:

$$\text{div}_h(v_1, v_2) \approx h^{-1}[(v_{i+1,j} - v_{i,j}) + (v_{i,j+1} - v_{i,j})],$$
$$\text{grad}_h v \approx h^{-1}(v_{i,j} - v_{i-1,j}, v_{i,j} - v_{i,j-1}),$$

so that, if $A \equiv \text{Id}$, we find the usual five-point box scheme for the operator Δ. Moreover, the ellipticity of the second-order continuous operator is preserved by this discretization. However, we must observe that it does not seem that any discrete maximum principle holds, which requires some care to ensure stability. We used an explicit time discretization with a time step ensuring stability for the reaction term on the one hand and, on the other hand, limited by the usual stability condition related to the case $A \equiv \text{Id}$ for the diffusion. If one wishes to quantify the original image over N grey level, and if we set $\delta g = 2N^{-1}$ and $g_i = -1 + (i - .5)\delta g$, the reaction function takes the form

$$f(x) = -\delta g^{-2}(x - g_i)\left(x - \frac{g_i + g_{i+1}}{2}\right)(x - g_{i+1}) \quad \text{for } x \in [g_i, g_{i+1}],$$

and the time step is $\Delta t = \min(.5\sigma ; 2)$. The resulting scheme can be summarized by the formula:

$$u^{n+1} = \max(-1 ; \min(1 ; u^n + \sigma h^2 \Delta t \, \text{div}_h([A_\varepsilon^h(u^n)][\nabla_h u^n]) + \Delta t f(u^n))).$$

3.1. Test images. In all of this section the *exact* image that we want to recover consists of a black triangle above a black narrow rectangle (see Figure 3 on p. 670) on a white background (so the reaction function is chosen to extract two grey levels). In a first experiment (top picture in Figure 1) this is the initial condition, which ideally should be preserved by our image processing, and we measure the distance between the computed solution and this exact solution by the L^2 norm of the difference. The curves A and B correspond to 32×32 and 64×64 images, where ε was kept at the same value (5 pixels for the first image, 10 pixels for the second). The curve C represents the evolution
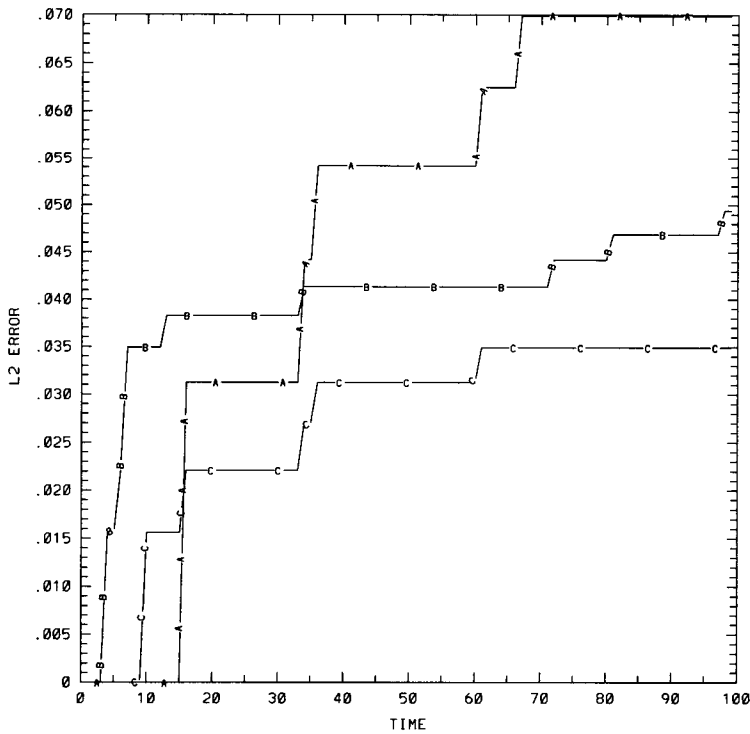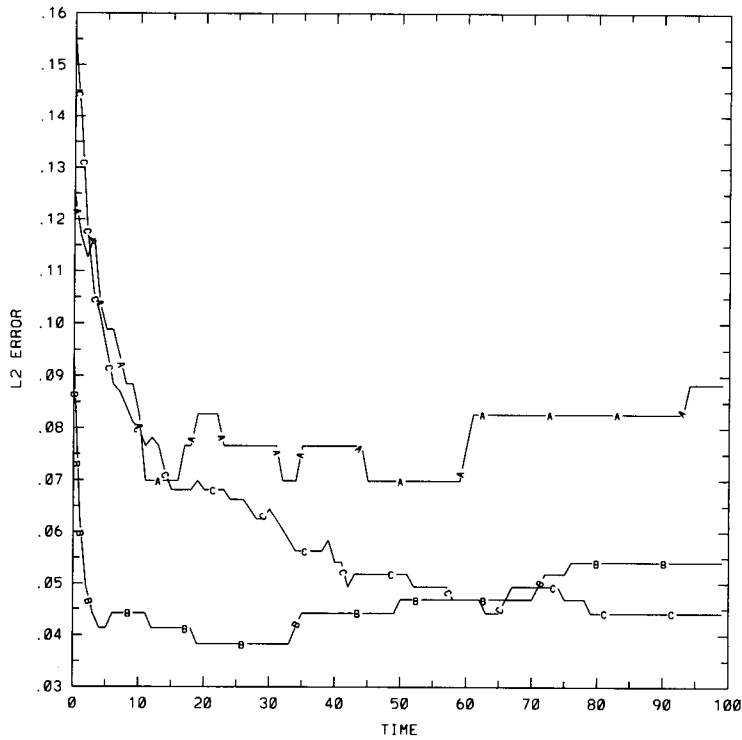
FIGURE 1. The effect of refinement on the asymptotic state of the model, for an *exact* (top) and perturbed (bottom) initial condition
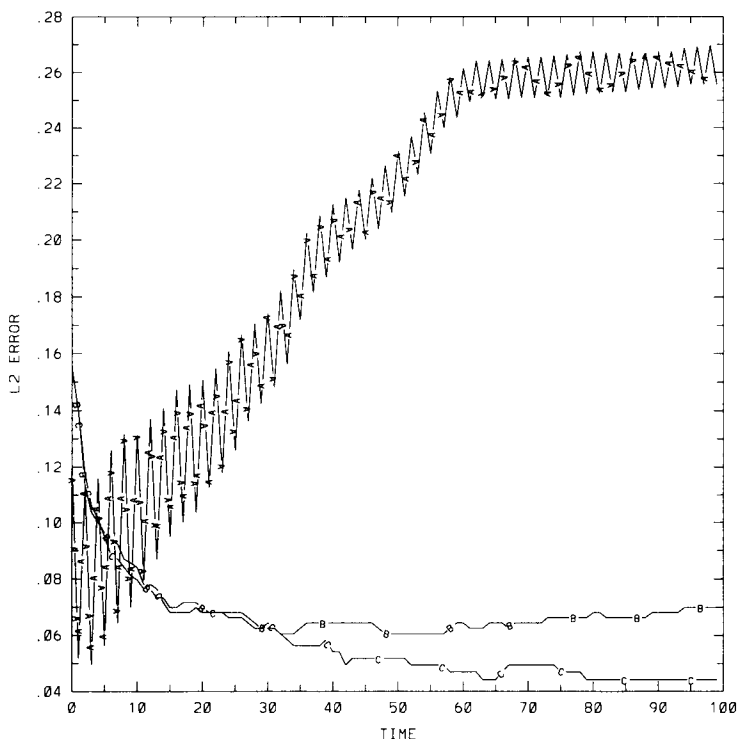
FIGURE 2. Isotropic reaction-diffusion and pure nonlinear diffusion compared to our method for a 64×64 image

of the error for the 64×64 image, where ε has been reduced to the width of 5 pixels. The successive improvements confirm the result of Theorem 2.3. In a second experiment (bottom picture of Figure 1) the same images have been slightly perturbed by replacing 10% of the pixels by random values. In this situation, the processed images basically fall into the same neighborhood of the desired state as in the previous experiments. In the second test (Figure 2) we have compared in the same situation as previously (64×64 image with a 10% noise) our method with, on the one hand an isotropic reaction-diffusion equation, and, on the other hand, our model where the reaction is turned off. All calculations were done with the same diffusion coefficient σ (about 10). At the end of the computation for the isotropic reaction-diffusion model, the image is almost completely lost. Although the reaction improves the convergence when combined with nonlinear diffusion in our model, it seems that the pure nonlinear diffusion model has the same kind of asymptotic properties. In the third test we applied a more severe alteration of a 128×128 image (Figure 3, left picture). In this case, 70% of the pixels have been destroyed, and we have represented the processed image obtained after 150 iterations and a thresholding dividing the image between black and white. The regularization has been done over 14 pixels in each direction. Since in this case most of the computation is consumed in the evaluation of u_ε, we have computed this term only every 20 iterations. As can be seen on the right picture, all the noise has been removed and the edges of the patterns have been reasonably preserved. In light of the analysis
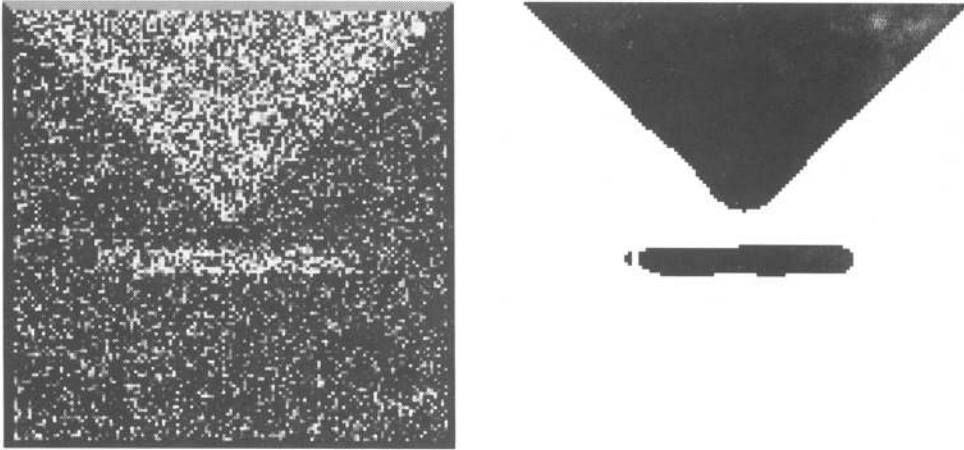
FIGURE 3. Image processing on a 70% destroyed image

done in Theorem 2.2, we can interpret the damage at the ends of the rectangle by the fact that at these points the curvature varies rapidly. Observe that the rectangle is only 7 pixels wide, so that simple regularization over a bandwidth of 14 pixels would have greatly affected it.

3.2. **Medical images.** We now illustrate the possibilities of our method in the context of medical imaging. The first image (top picture in Figure 4) is a 256×256 "classical" MRI image of the brain (see [2] for an approach via dynamical systems and neural networks of the same image). In this image, which is actually extracted from a series of slices, one would like to have a precise contouring of the tumor which appears as a slightly clearer zone in the top-left part of the picture. We have processed this image with the above method, using a regularization over only 3 pixels in both directions; this is actually possible because the original image is not very noisy. Moreover, the computation of the diffusion matrix has only been performed at the first step (as for the model analyzed in Theorem 2.3). We used a reaction function which has 8 stable zeros (which means that 8 grey levels are selected) and is piecewise cubic. The bottom picture in Figure 4 shows the processed image after rescaling; on this image it is easy to identify precisely the location of the tumor, as well as two darker zones around this white region, corresponding to compressed areas in the brain. Our last example deals with a more challenging aspect of medical imaging. The original image is a 512×512 image obtained through an angiogram, still in the brain, and the interesting information concerns the vessels which trace complicated one-dimensional curves starting from the core (top-left picture in Figure 5 on p. 672). We first checked that although the original image does not seem very noisy (somehow, the eye is able to follow the vessels far away from the core), simple thresholding fails to select these vessels on a black and white image. The dots which appear on the top-right picture correspond to inhomogeneous tissues which cannot be distinguished from the vessels. Of course, one could think that this difficulty results from a value of the threshold which has been chosen too small, but taking a bigger value has the effect of disconnecting the vessels themselves. The bottom-left picture is a processed version of the original image with 4 grey levels in the reaction function and a regularization over 7 pixels in each direction. As in the previous example, the
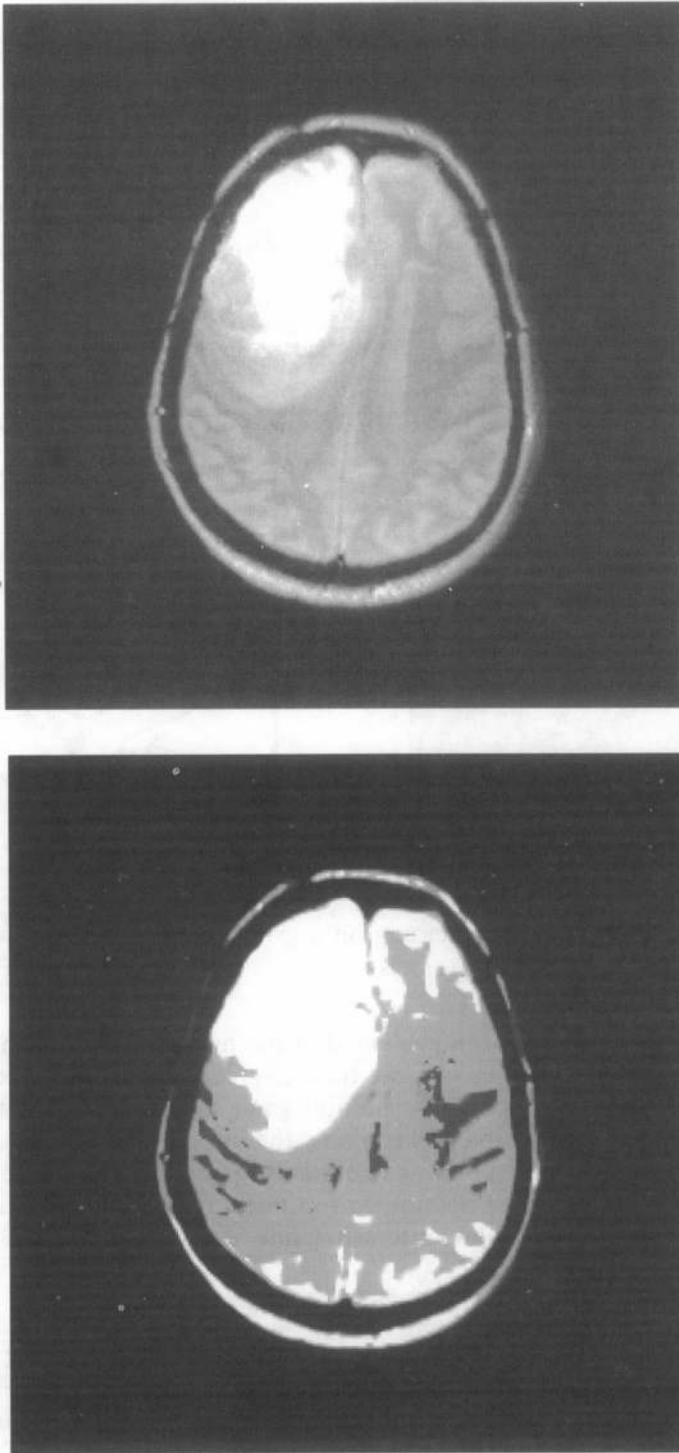
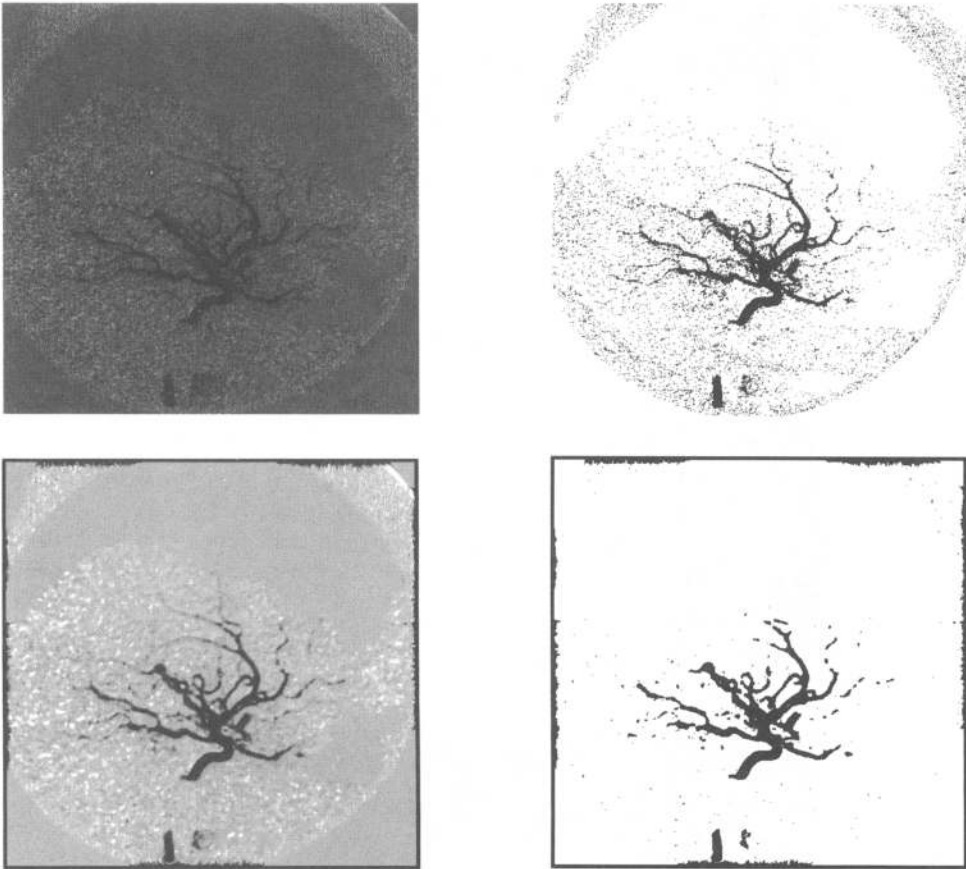FIGURE 4. Image processing with 8 grey levels on a 256×256 MRI image of the brain

FIGURE 5. Image processing on a 512×512 angiogram with recognition of $1 - D$ vessels

diffusion matrix has only been evaluated at the first step. Although this image seems better than the original one, we tried to give a less subjective comparison criterion, and played the same thresholding game as on the top-right picture. It appears now that it is possible to keep only the interesting features and remove the part of the image corresponding to either brighter or inhomogeneous tissues. Although it has not yet been possible to reconnect the vessels far away from the core, the processed image shows interesting fine details in the core of the vessels.

4. CONCLUSION

We have presented a scalar reaction-diffusion model for image processing which is able to provide nontrivial stable asymptotic states. The test images show that this model retains the interesting features of isotropic diffusion for noise reduction, without leading to a destruction of the edges of the signal. Moreover, the fact that one-dimensional smooth coherent signals are not affected by this kind of processing, predicted by the analysis from the particular shape of the diffusion operator, is confirmed on real examples. So far, and for simplicity, the reaction function uses uniformly distributed thresholds. Future

plans consist in studying the possibility of an interactive choice of these thresh-olds. Also, the extension to 3-dimensional images will be considered, as well as the coupling between these methods and segmentation methods, which are currently used at the medical school in Grenoble.

BIBLIOGRAPHY

1. L. Alvarez, J.-M. Morel, and P.-L. Lions, *Image selective smoothing by nonlinear diffusion*, SIAM J. Numer. Anal. **29** (1992), 845–866.

2. F. Berthomier, O. François, D. Francillard, T. Coll, P. Cinquin, I. Marque, and J. De-mongeot, *Asymptotic behavior of neural networks and image processing*, Self-Organization, Emerging Properties and Learning (A. Babloyantz, ed.), Plenum Press, NATO Series, New York, 1991, pp. 219–230.

3. G.-H. Cottet, *Modèles de reaction-diffusion pour des réseaux de neurones stochastiques et déterministes*, C. R. Acad. Sci. Paris Sér. I Math. **312** (1991), 217–221.

4. O. François, *Systèmes de spins synchronisés, modèles de Hopfield stochastiques*, Thesis, Université Joseph Fourier, Grenoble, 1992.

5. T. Kurosawa, H. Tsuchiya, Y. Maruyama, H. Ohtsuka, and K. Nakazato, *A new bi-level reproduction of continuous tone images*, Proc. 2nd Internat. Conf. Image Processing and its Applications, IEE Publ., London, 1986, pp. 82–86.

6. S. Osher and L. Rudin, *Feature-oriented image enhancement using shock filters*, SIAM J. Numer. Anal. **27** (1990), 919–940.

7. P. Perona and J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Analysis and Machine Intelligence **12** (1990), 629–639.

8. C. B. Price, P. Wanback, and A. Oosterlinck, *Application of reaction-diffusion equations to image processing*, Proc. 3rd Internat. Conf. Image Processing and its Applications, IEE Publ., London, 1989, pp. 49–53.

LMC-IMAG, UNIVERSITÉ JOSEPH FOURIER, BP 53X, 38041 GRENOBLE CÉDEX, FRANCE
E-mail address: cottet@imag.fr

TIMC-IMAG, FACULTÉ DE MÉDECINE DE GRENOBLE, 38700 LA TRONCHE, FRANCE