SIAM J. NUMER. ANAL.
Vol. 31, No. 2, pp. 590-605, April 1994

SIGNAL AND IMAGE RESTORATION USING SHOCK FILTERS AND ANISOTROPIC DIFFUSION*

LUIS ALVAREZ† AND LUIS MAZORRA†

Abstract. The authors define a new class of filters for noise elimination and edge enhancement by using shock filters and anisotropic diffusion. Some nonlinear partial differential equations used as models for these filters are studied. The authors develop recursive and unconditional stable schemes which drastically reduce the computational effort of the algorithms. A new fast recursive approach to linear Gaussian filters is also shown by using the heat equation.

Key words. anisotropic diffusion, image restoration, shock filters, deconvolution, edge enhancement

AMS subject classifications. 49F22, 53A10, 82A60, 76T05, 49A50, 49F10, 80A15

1. Introduction.

This paper deals with signal and image restoration; by this we mean noise elimination, edge enhancement, and deconvolution. We present different models according to the signal dimension (one- or two-dimensional signals).

In the case of one-dimensional signals we introduce as a model the following hyperbolic partial differential equation.

$$u_t + F(G_\sigma^* u_{xx}, G_\sigma^* u_x)u_x = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^+, \tag{1}$$

where $G_\sigma(\,.\,)$ is a family of smoothing kernels which depends on a parameter σ (for instance, a family of Gaussians), $F(.\,,.)$ is a function which satisfies:

$$F(p,q)pq \geq 0 \quad \text{for any } p, q \in \mathbb{R}; \tag{2}$$

a simple choice for F is

$$F(p,q) = \text{sign}(p)\text{sign}(q), \tag{3}$$

where $\text{sign}(s) = 1$ if $s > 0$, $\text{sign}(s) = -1$ if $s < 0$, and $\text{sign}(0) = 0$.

Roughly speaking, if $u(x,0)$ is the initial signal, (1) develops shocks in the position of the zero crossing of $G_\sigma^* u_{xx}$, and in this way it produces an enhancement of the edges.

We discretize (1) by using a recursive scheme which is unconditionally stable, dissipative only in the interior of the homogeneous regions of the signals (meaning that discontinuities situated in the zero crossing of $G_\sigma^* u_{xx}$ are not smeared), and nonoscillatory.

We define the smoothing kernel $G_\sigma(\,.\,)$ as an approximation of the Gaussian filter by using a fast recursive discretization of the heat equation ($u_t - u_{xx} = 0$).

In the case of two-dimensional signals (i.e., images) we introduce as a model the following partial differential equation:

$$u_t = C\mathcal{L}(u) - F(G_\sigma^* u_{\eta\eta}, G_\sigma^* u_\eta)u_\eta \quad \text{in } \mathbb{R}^2 \times \mathbb{R}^+, \tag{4}$$

where $\eta = \eta(x)$ is the direction of $\nabla u(x)$, $F(.\,,.)$ verifies (2), $G_\sigma(.\,,.)$ is a family of two-dimensional smoothing kernels, $C > 0$ is any positive constant, and $\mathcal{L}(u)$ is any

*Received by the editors October 28, 1992; accepted for publication (in revised form) February 11, 1993. This research was partially supported by the C.I.C.Y.T., Spain, project TIC 92-1282.

†Departamento de Informatica y Sistemas, Universidad de Las Palmas, Campus Universitario de Tafira, 35017 Las Palmas, Spain.

directional smoothing operator; for instance, a simple choice for $\mathcal{L}(u)$ is

$$\mathcal{L}(u) = u_{\xi\xi},$$

where $\xi = \xi(x)$ is the direction perpendicular to the gradient $\nabla u(x)$.

Roughly speaking, this parabolic-hyperbolic equation diffuses the initial image $u(x, y, 0)$ in the parallel directions of the edges (noise elimination) and develops a shock in the perpendicular direction of the edge (edge enhancement and deconvolution).

We discretize (4) by using a recursive scheme which is unconditionally stable.

Due to the unconditional stability, our schemes produce interesting results with few iterations. On the other hand, the recursive implementation of the models drastically reduces the computational effort of each iteration. Therefore, our schemes produce stable, efficient, and very fast algorithms.

We remark that we have no boundary conditions in our models. In the practical cases we need to impose an artificial boundary condition. In this paper, we impose, in a natural way, that if $\partial u/\partial \nu = 0$ in the boundary (here ν represents the perpendicular direction to the boundary), it means that we minimize the boundary influence.

The organization of the paper is as follows: In §2 we develop the model (1) and we present some related models. In §3 we develop the model (4) for images (or two-dimensional signals) and we present some related models. Finally, in §4 we present some numerical experiences.

2. One-dimensional signal.

Related models. Rudin in [14] first applied the concepts and techniques developed in the numerical solution of nonlinear hyperbolic equations to image enhancement. There, the first experimental shock filter, based on a modification of the nonlinear Burgers equation was used. In a more recent paper [10], Osher and Rudin introduced an important improvement by using as a model the following hyperbolic equation:

$$u_t + F(u_{xx})|u_x| = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^+,$$

where $F(\,.\,)$ is a function such that $F(s)s \geq 0$. In order to discretize this equation they used an explicit monotone scheme which preserves the total variation and the size and location of local extrema. Therefore, this scheme cannot remove some kinds of noise, for example, "salt and pepper" noise. This model develops shocks in the position of the zero crossing of u_{xx}. Then if we have a noisy signal, it produces a lot of spurious shocks due to the influence of noise (see Fig. 1).

Our model. In a natural way (following the classic theory of Marr [8]) we introduce in the above equation a smoothing kernel $G_\sigma(\,.\,)$ in order to avoid the spurious shocks. We propose as a model the equation:

$$u_t + F(G_\sigma^* u_{xx}, G_\sigma^* u_x) u_x = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^+.$$

Since $G_\sigma^* u_x = (G_\sigma)_x^* u$, this equation is better posed because it becomes a first-order differential equation. Now the shocks are developed in the position of the zero crossing of $G_\sigma^* u_{xx}$.

To discretize (1), we first present a quick summary of numerical results concerning the linear hyperbolic equation.

Implicit schemes for hyperbolic equations. We consider the well-known linear hyperbolic equation:

$$u_t + cu_x = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^+, \tag{5}$$

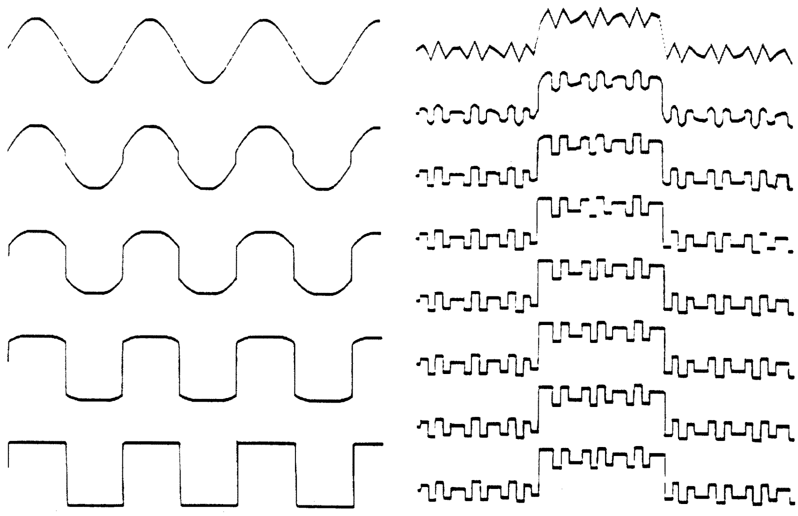

FIG. 1. *In the left part, we show the solution of* (1) ($\sigma = 0$) *for the initial signal* $u_0(x) = \cos x$. *We notice as discontinuities are developed in the location of the zero crossing of* u_{xx}. *In the right part, we show the evolution of a noisy step function* ($\sigma = 0$), *and we notice the occurrence of spurious shocks due to noise influence.*

where $c \in \mathbb{R}$ is a constant. Equation (5) is a pure transport equation, i.e., the solution $u(x,t)$ of (5) for an initial datum $f(x)$ is $u(x,t) = f(x - ct)$, where the direction of the travelling wave depends on the sign of c. In order to get an unconditional stable algorithm, we discretize (5) by using an implicit scheme. Let $h > 0$ and $k > 0$ be the spatial and temporal increments, and $u_i^n \simeq u(ih, kn)$. Assume that $c > 0$, and we use the following discretization:

$$\frac{u_i^{n+1} - u_i^n}{k} + c\frac{u_i^{n+1} - u_{i-1}^{n+1}}{h} = 0.$$

Let $\lambda - ck/h$. Then we have

$$(1+\lambda)u_i^{n+1} - \lambda u_{i-1}^{n+1} = u_i^n \quad \text{for any } i \in \mathbb{I} \text{ and } n \in \mathbb{R}.$$

Notice that u_i^{n+1} can be calculated with few operations (only two multiplications and one addition for each i). In the frequency domain the signal u_i^n can be written as the function $U^n(w) = \sum_i u_i^n e^{-iwnh}$. (See [6] for more details.) By standard techniques (see, for instance, [16]) and following the above discretization of (5), $U^{n+1}(w)$ can be computed as $U^{n+1}(w) = U^n(w)m_\lambda(w)$, where $m_\lambda(w)$ is defined by

$$m_\lambda(w) = \frac{1}{1 + \lambda - \lambda e^{iwh}}.$$

Since $|m_\lambda(w)| \leq 1$ for any $\lambda > 0$ and $w \in \mathbb{R}$, the scheme is stable for any value of k, $h > 0$, and we have:

$$m_\lambda(w) = \frac{1/(1+\lambda)}{1 - \lambda/(1+\lambda)e^{iwh}} = \sum_{m=0}^{+\infty} \frac{\lambda^m}{(1+\lambda)^{m+1}} e^{imwh}.$$

Using this equality we can explicitly compute the coefficients of the linear filter associated with the recursive scheme obtained by the implicit algorithm defined above. To compute u^{n+1} is equivalent to the convolution of u^n with the signal $h_i(\lambda)$ defined by

$$h_i(\lambda) = \begin{cases} \lambda^{|i|}/(1+\lambda)^{|i|+1} & \text{if } i \preceq 0, \\ 0 & \text{if } i \succ 0. \end{cases}$$

We notice that $m_\lambda(w)$ is the Fourier transform of the signal $h_i(\lambda)$. This scheme is dissipative (i.e., the discretization introduces an artificial diffusion; see [16] for more details). In fact, when λ tends to $+\infty$, u_i^{n+1} tends to the average of the signal u_j^n for $j \in (-\infty, i)$.

If $c < 0$, we use the following discretization:

$$\frac{u_i^{n+1} - n_i^n}{k} + c\frac{u_{i+1}^{n+1} - u_i^{n+1}}{h} = 0,$$

and by using the same arguments as in the case $c > 0$ we show that u^{n+1} can be interpreted as the convolution of u^n with the filter defined by:

$$h_i(\lambda) = \begin{cases} 0 & \text{if } i \prec 0, \\ \lambda^{i}/(1+\lambda)^{i+1} & \text{if } i \succeq 0, \end{cases}$$

where $\lambda = ck/h$.

Next, we show a fast algorithm in order to approach the Gaussian filters.

Gaussian filters approximation. We focus our attention on the one-dimensional case. In fact, the approximation of two-dimensional Gaussian filters is obtained by using the one-dimensional approximation in two perpendicular directions. It means that we can compute the two-dimensional convolution of a Gaussian function with a picture by the combination of the one-dimensional convolution of the picture in the direction of the x-axis and the y-axis. Let $G_\sigma(x)$ be the Gaussian filter

$$G_\sigma(x) = \frac{e^{-x^2/2\sigma^2}}{\sigma\sqrt{2\pi}}.$$

It is well known that the convolution of a function $f(x)$ with a Gaussian filter is equivalent to solving the heat equation $u_t - u_{xx} = 0$, for the initial datum $u(x,0) = f(x)$, i.e., let $t = \sigma^2/2$, then $u(x,t) = G_\sigma^* f(x)$. Therefore, an approximation of the solution of the heat equation is an approximation of Gaussian filters. We are going to discretize this equation by using the most simple unconditionally stable scheme. Let $h > 0$ and $k > 0$ be the spatial and temporal increments, and $u_i^n \simeq u(ih, kn)$. We use the following discretization:

$$\frac{u_i^{n+1} - u_i^n}{k} - \frac{u_{i+1}^{n+1} + u_{i-1}^{n+1} - 2u_i^{n+1}}{h^2} = 0.$$

Let $\lambda = k/h^2$. Then we have

$$(1+2\lambda)u_i^{n+1} - \lambda u_{i-1}^{n+1} - \lambda u_{i+1}^{n+1} = u_i^n \quad \text{for any } i \in \mathbb{I} \quad \text{and} \quad n \in \mathbb{N}.$$

By using the same arguments as in the case of the linear hyperbolic equation, we can interpret the signal u^{n+1} as the convolution of the signal u^n with a filter defined by the transference function

$$m_\lambda(w) = \frac{1}{1 + 2\lambda - 2\lambda\cos(wh)}. \tag{6}$$

As $|m_\lambda(w)| \le 1$ for any $\lambda > 0$ and $w \in \mathbb{R}$, the scheme is stable for any value of $k > 0$ and $h > 0$. Next, we show some interesting results about (6).

LEMMA 1. *Let $h, k, t > 0, 1 = k/h$, and $m_\lambda(w)$ be defined by* (6); *then we have*

$$m_\lambda(w) = \frac{\nu/\lambda}{(1 - \nu e^{iwh})(1 - \nu e^{-iwh})}, \tag{7}$$

$$m_\lambda(w) = (1 + 4\lambda)^{-1/2} \sum_{m=-\infty}^{+\infty} \nu^{|m|} e^{iwhm}, \tag{8}$$

$$\int_{-\pi/h}^{+\pi/h} m_\lambda^n(w)\, dw = \frac{2\pi\nu^n}{h\lambda^n(1 - \nu^2)^{2n}} \sum_{m=0}^{n-1} \binom{n + m - 1}{m} \binom{n - 1}{m} \nu^{2m}(1 - \nu^2)^{n-m}, \tag{9}$$

$$\textit{where} \quad \nu = \frac{1 + 2\lambda - \sqrt{1 + 4\lambda}}{2\lambda}.$$

Proof. Equations (7) and (8) follow by a straight calculus. Equation (9) follows by using complex integration in the unit circle (we take $z = e^{iwh}$).

Notice that following (7), $m_\lambda(w)$ is the combination of a causal filter and an anticausal filter. Therefore, we can decompose the calculus of u^{n+1} in the following steps.

$$u_i^{n+1/3} = u_i^n + \nu u_{i-1}^{n+1/3} \qquad \text{for any } i \in \mathbb{I}, \tag{i}$$

$$u_i^{n+2/3} = u_i^{n+1/3} + \nu u_{i+1}^{n+2/3} \quad \text{for any } i \in \mathbb{I}, \tag{ii}$$

$$u_i^{n+1} = (\nu/\lambda) u_i^{n+2/3} \qquad \text{for any } i \in \mathbb{I}. \tag{iii}$$

Indeed, (i) corresponds to the convolution of the signal u^n with the filter defined by the transference function $m_\lambda^1(w) = 1/(1 - \nu e^{iwh})$, (ii) corresponds to the convolution of the signal computed in the step (i), $u_i^{n+1/3}$, with the filter defined by the transference function $m_\lambda^2(w) = 1/(1 - \nu e^{-iwh})$; finally, (iii) corresponds to multiplying the signal computed in step (ii), $u_i^{n+2/3}$, by (ν/λ). The combination of these three steps is equivalent to the convolution with $m_\lambda(w)$. Notice that the calculus of u_i^{n+1} needs only three multiplications and two additions for each $i \in \mathbb{I}$. Moreover, in many applications we can avoid step (iii) because it is only a normalization process.

Notice that by (8) we can interpret u^{n+1} as the convolution of u^n with the filter $h_i(\lambda)$ defined by:

$$h_i(\lambda) = (1 + 4\lambda)^{-1/2} \nu^{|i|} \quad \text{for any } i \in \mathbb{I};$$

since ν tends to 1 when λ tends to infinity, the effective support of this filter goes to infinity; however, notice that the number of operations in order to calculate u^{n+1} remains fixed independently of λ.

The transference function of the filter associated to the nth iteration of the implicit discretization defined above is $m_\lambda^n(w)$. Therefore, u^n can be obtained as the convolution of the original signal u^0 with this filter. Equality (9) gives a formula to compute the energy of such a filter.

In terms of frequencies, we can interpret the solution of the heat equation $u(x,t)$ as the convolution of the initial datum $f(x)$ with a Gaussian filter, which has the transference function $m(w) = e^{-tw^2}$; when we discretize the signal $f(x)$ and the Gaussian filter $G_\sigma(x)$ by using a spatial increment h, the filter $m(w)$ becomes a $2\pi/h$ periodic function defined by:

$$m(w,h) = \sum_{m=-\infty}^{+\infty} e^{-t(w+2\pi m/h)^2} \quad \text{for any } t > 0$$

(see [6] for more details). On the other hand, u_i^n is an approximation of $u(ih, nk)$ and it can be interpreted as the convolution of the initial signal u_i with the filter defined by the transference function $m_\lambda^n(w)$. Therefore, $m_\lambda^n(w)$ is an approximation of $m(w,h)$. From a theoretical point of view, λ must be taken small so that u_i^n will be near the continuous solution u. Therefore, if t is taken "big" many iterations are needed for approaching $u(ih,t)$ and then the computational effort grows quickly. Thus from a practical point of view we need to take λ big in order to reduce the computational effort. On the other hand, if h is small enough and t big enough the support of the function $m(w) = e^{-tw^2}$ is nearly included in $[-2\pi/h, 2\pi/h]$ and then $m(w,h) \cong m(w)$ is in $[-2\pi/h, 2\pi/h]$. Notice that $m_\lambda(w)$ and $m_\lambda^n(w)$ are nonnegative even though $2\pi/h$ periodic functions decrease in $[0,\pi/h]$ and $m_\lambda(0) = m_\lambda^n(0) = 1$. Moreover, $m_\lambda^n(w)$ decreases faster than $m_\lambda(w)$ and in this way better approaches the function $m(w)$. The theoretical relation between λ, n, and t is given by

$$h^2 \lambda n = t. \tag{10}$$

But as we show in the numerical experiments (see Fig. 2), in this case $m_\lambda^n(w)$ converges slowly to $m(w)$ as n tends to infinity. In order to accelerate this convergence we use a different criterion for choosing λ and n. We take λ and n so that $m(w)$ and $m_\lambda^n(w)$ will have the same energy:

$$F(\lambda,n) = \int_{-\pi/h}^{\pi/h} m_\lambda^{2n}(w)\,dw = \int_{-\pi/h}^{\pi/h} m^2(w)\,dw \cong (\pi/2t)^{1/2}. \tag{11}$$

Notice that $F(\lambda,n)$ is a decreasing continuous function with respect to λ which satisfies:

$$F(0,n) = 2\pi/h \quad \text{and} \quad \lim F(\lambda,n) = 0 \quad \text{if } \lambda \longrightarrow +\infty.$$

Therefore, if $2\pi/h > (\pi/2t)^{1/2}$ we have that for any $n \in \mathbb{N}$ there exist a unique $\lambda > 0$ such that $F(\lambda,n) = (\pi/2t)^{1/2}$. Moreover, we can make this relation explicit by using (9). For instance, for fixing ideas we take $n = 1$; then we have the relation

$$(2\pi/h)(1+2\lambda)(1+4\lambda)^{-3/2} = (\pi/2t)^{1/2}.$$

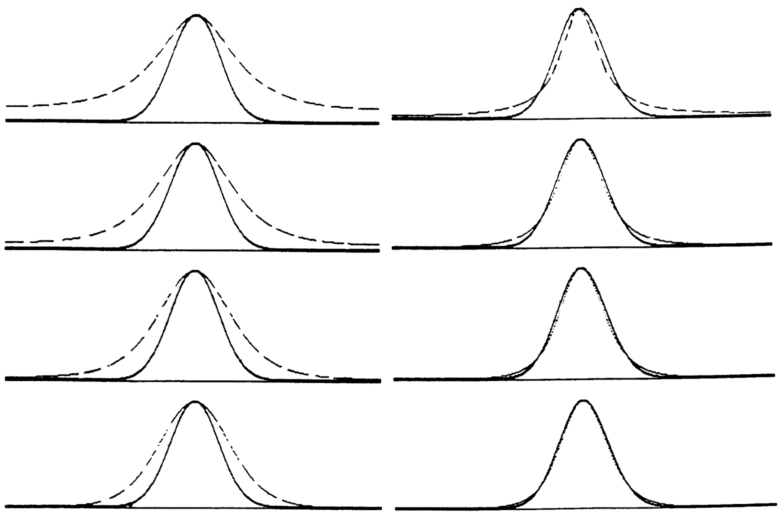

FIG. 2. *Gaussian approximation. We present an approximation of the Gaussian filter* $m(w) = e^{-3w^2}$ *by using iterations of* $m_\lambda(w)$ *(the solid line corresponds to* $m(w)$ *and the dashed line corresponds to the iterations of* $m_\lambda(w)$*). In the left part we use the classical approach given by* (10). *In the right part we use the energy conservation criterion* (11).

As we show in the numerical experiments (see Fig. 2), criterion (11) is more efficient than (10) in order to approach $m(w)$.

Remark 1. The recursive filtering structure drastically reduces the computational effort required for filter implementations. The operations are done with a fixed number of multiplications and additions per output point independently of the size of the neighborhood considered. (See, e.g., Shen and Castan [15] and Deriche [4] for overviews of this subject.)

Discretization of (1). We now consider the equation (1):

$$u_t + F(G_\sigma^* u_{xx}, G_\sigma^* u_x) u_x = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^+;$$

let $h > 0$ and $k > 0$ be the spatial and temporal increments, and $u_i^n \simeq u(ih, kn)$. We use the following discretization:

(12)

$$\frac{u_i^{n+1} - u_i^n}{k} + \frac{F^-(u_i^n)(u_{i+1}^{n+1} - u_i^{n+1}) + F^+(u_i^n)(u_i^{n+1} - u_{i-1}^{n+1})}{h} = 0, \quad i \in \mathbb{I}, \quad n \in \mathbb{N},$$

where $F^-(s) = \min\{0, F(s)\}$, $F^+(s) = \max\{0, F(s)\}$, $F(u_i^n) = F((G_\sigma^* u_i^n)_{xx}, (G_\sigma^* u_i^n)_x)$, $G_\sigma^* u_i^n$ is calculated using the filter (6), and finally

$$(13) \qquad \begin{aligned} (G_\sigma^* u_i^n)_{xx} &= \frac{G_\sigma^* u_{i+1}^n + G_\sigma^* u_{i-1}^n - 2G_\sigma^* u_i^n}{h^2}, \\ (G_\sigma^* u_i^n)_x &= \frac{G_\sigma^* u_{i+1}^n - G_\sigma^* u_{i-1}^n}{2h}. \end{aligned}$$

In order to solve (12) we use the following fixed-point equation:

$$w_i = \lambda_i b_i + \nu_i w_{i+1} + \eta_i w_{i-1} \quad \text{for any } i \in \mathbb{I}, \tag{14}$$

where $w_i = u_i^{n+1}$, $\lambda_i = (1 + (k/h)|F(u_i^n)|)^{-1}$, $b_i = u_i^n$, $\nu_i = -\lambda_i F^-(u_i^n)k/h$, and $\eta_i = \lambda_i F^+(u_i^n)k/h$.

Next we show that this fixed-point equation has a unique, stable, and nonoscillatory solution.

THEOREM 1. *Let b_i be a bounded sequence, and λ_i, ν_i, η_i such that $\nu_i\eta_i = 0, \lambda_i > \varepsilon, 0 \leq \nu_i, \eta_i < 1$, and $\lambda_i + \nu_i + \eta_i = 1$ for any $i \in \mathbb{I}$ and some $\varepsilon > 0$. Then we have the following*:

(i) *The fixed-point scheme* (14) *has a unique bounded solution w_i.*

(ii) *Let $m = \inf\{b_i, i \in \mathbb{I}\}$ and $M = \sup\{b_i, i \in \mathbb{I}\}$. Then $m \leq w_i \leq M$ for any $i \in \mathbb{I}$.*

(iii) *If F is defined by* (3), *then* (14) *is monotonicity preserving* (*i.e., if b_i is a monotone sequence then w_i is also monotone*).

Proof. (i) Let E be the space of bounded sequence $z = \{z_i \in R,\ i \in \mathbb{I}\}$ and $Q : E \longrightarrow E$ defined by $Q(z)_i = \lambda_i b_i + \nu_i z_{i+1} + \eta_i z_{i-1}$. Let $z, v \in E$; then we have

$$\sup_{i\in\mathbb{I}}\{|Q(z)_i - Q(v)_i|\} \leq \sup_{i\in\mathbb{I}}\{|\nu_i(z_{i+1} - v_{i+1}) + \eta_i(z_{i-1} - v_{i-1})|\} \leq (1-\varepsilon)\sup_{i\in\mathbb{I}}\{|z_i - v_i|\}.$$

Therefore, Q is a contractive function and (i) follows by using the classic Banach fixed-point theorem. Moreover, if $z^0 \in E$, then the sequence $z^k \in E$ defined by $z^k = Q(z^{k-1})$ converges to the solution w of (14).

In order to prove (ii) and (iii) we notice that there exist two special types of points $i \in \mathbb{I}$ defined by the sets

$$A = \{i \in \mathbb{I} : \eta_{i+1} = \nu_i = 0\},$$

$$B = \{i \in \mathbb{I} : \eta_i = \nu_{i+1} = 0\}.$$

A represents the set of points i where the solution w of (14) can develop shocks. In fact, if $i \in A$, the solution w_j for $j \leq i$ is completely independent of w_j for $j \geq i+1$.

In $i \in B$, we have (by using (14)):

$$w_i = \lambda_i b_i + (1 - \lambda_i)w_{i+1},$$

$$w_{i+1} = \lambda_{i+1}b_{i+1} + (1 - \lambda_{i+1})w_i,$$

$$w_i = [\lambda_i b_i + \lambda_{i+1}(1 - \lambda_i)b_{i+1}]/[1 - (1 - \lambda_{i+1})(1 - \lambda_i)],$$

$$w_{i+1} = [\lambda_{i+1}b_{i+1} + \lambda_i(1 - \lambda_{i+1})b_i]/[1 - (1 - \lambda_{i+1})(1 - \lambda_i)].$$

We remark that w_i and w_{i+1} are an average of b_i and b_{i+1}.

Let $i_1, i_2 \in A$, $i_1 < i_2$; then there exist $j_1 \in B$ such that $i_1 < j_1 < i_2$. Moreover, if $j_1, j_2 \in B$, then there exists $i_2 \in A$ such that $j_i < i_2 < j_2$.

Therefore, we can decompose $\mathbb{I} = \bigcup_m [i_m + 1, i_{m+1}]$, where i_m, $i_{m+1} \in A$ (or $i_m = -\infty$, $i_{m+1} = +\infty$) and $[i_m + 1, i_{m+1}] \cap B = \{j_m\}$. Then we can explicitly compute the solution w_i in $[i_m + 1, i_{m+1}]$. Since $j_m \in B$, w_{j_m}, w_{j_m+1} are computed by using the above equality. If $j \in [i_m + 1, j_m]$, then j is not in A and B and since $\eta_{im+1} = 0$, then $\eta_j = 0$ and we compute w_j in a recursive way by using the fact that in this region

$$w_j = \lambda_i b_i + (1 - \lambda_i) w_{j+1}.$$

By using the same argument if $j \in [j_m, j_{m+1}]$ then $\nu_j = 0$. We compute w_j in a recursive way by using the fact that in this region

$$w_j = \lambda_i b_i + (1 - \lambda_i) w_{j-1}.$$

Numerically, this method is very efficient (see Fig. 3). Notice that if B is empty then A is empty or $A = \{i_0\}$. When A is empty, we have that $\eta_i = 0$ for any i or $\nu_i = 0$ for any i and we can apply the above argument to obtain w_i in a recursive way. If $A = \{i_0\}$ then $\eta_i = 0$ if $i > i_0$ and $\nu_i = 0$ if $i \preceq i_0$, then we can compute w_i separately in $(-\infty, i_0]$ and $[i_0 + 1, \infty)$.

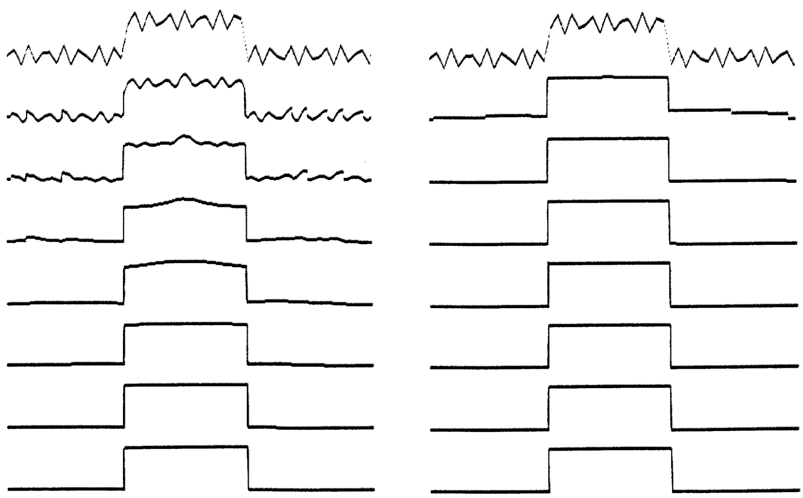

FIG. 3. *We show (from top to bottom) some iterations of scheme* (12) *for a noisy initial signal. In the left part* $\sigma = 999$ *and* $k = 5$. *In the right part* $\sigma = 9999$ *and* $k = 50$. *Notice that the initial step function is reconstructed perfectly and in a stable way with respect to* k *and* σ.

Finally, (ii) and (iii) follow because we can interpret the solution w of (14) as

$$w_i = \sum_j a_{ij} b_j \quad \text{for any } i \in \mathbb{I},$$

where $a_{ij} \succeq 0$ and $\Sigma a_{ij} = 1$ for any $i \in \mathbb{I}$.

Remark 2. Notice that if we consider $F(p, q)$ defined by (3) then $\lambda_i = \lambda$ for any $i \in \mathbb{I}$ and if we take $k \longrightarrow +\infty$ then $\lambda \longrightarrow 0$ and the solution w of (14) converges to a step function w^∞ defined by

$$w_i^\infty = (b_{j_m} + b_{j_m+1})/2 \quad \text{for any } i \in [i_m + 1, i_{m+1}],$$

where j_m, i_m, i_{m+1} are defined above.

Remark 3. It can be interesting to take the variance σ_2 of the filter $G_\sigma(\,.\,)$ as a function of t (i.e., $\sigma = \sigma(t)$). In this way we introduce a multiscale interaction in the calculus of the location of the zero crossing of $G_\sigma^* u_{xx}$ and therefore scheme (12) develops shocks in the location of the zero crossing which remain stable in different scales given by $\sigma(t)$.

3. Two-dimensional model.

Related models. The "low level" analysis of images presents two opposite requirements. Generally, we wish to smooth the homogeneous regions of the picture with two scopes: noise elimination and image interpretation.

The low-pass filtering is generally made by convolution with Gaussians of increasing variance. We understand easily the necessity of a previous low-pass filtering: if the signal is noisy, the gradient will have many irrelevant maxima which must be eliminated. Koenderink [9] noticed that the convolution of the signal with Gaussians at each scale was equivalent to the solutions of the heat equation ($u_t - \Delta u = 0$) with the signal as initial datum. Denoting by u_0 this datum, the "scale space" analysis associated with u_0 consists in solving the Cauchy problem:

$$u_t = u_{xx} + u_{yy} = \Delta u,$$

$$u(x, y, 0) = u_0(x, y).$$

The solution of this equation for an initial datum with bounded quadratic norm is given by the convolution of u_0 with a Gaussian function ($u(x, y, t) = G_t^* u_0$), where

$$G_t(x, y) = \frac{e^{-(x^2+y^2)/4t}}{4\pi t};$$

here, the variance of this Gaussian filter is given by the relation $t = \sigma^2/2$. Then (x, y) is an edge point for the "scale" t at points where Δu changes sign and $|\nabla u|$ is "big." Of course, this last condition introduces some a priori defined threshold. Unfortunately, it is well known that the "edges" at low scales give an inexact account of the boundaries which, according to our perception, should be regarded as correct.

An important improvement of the edge detection theory was proposed by Perona and Malik [12]. Their main idea is to introduce a part of the edge detection step in the filtering itself, allowing an interaction between scales from the beginning of the algorithm. They propose replacing the heat equation by a nonlinear equation:

$$u_t = \operatorname{div}(g(|\nabla u|)\nabla u) \qquad \text{in } \mathbb{R}^N \times \mathbb{R}^+,$$

$$u(x, 0) = u_0(x) \qquad \text{in } \mathbb{R}^N.$$

In this equation, g is a smooth nonincreasing function with $g(0) = 1$, $g(s) \geq 0$, and $g(s)$ tending to zero at infinity, while $u_0(\,.\,)$ represents the original image. The idea is that the smoothing process obtained by the equation is "conditional": if $\nabla u(x)$ is big, then the diffusion will be low, and therefore the exact localization of the edges will be kept. If $\nabla u(x)$ is small, then the diffusion will tend to smooth still more around x. This model has been improved in Catte, Coll, Lions, and Morel [3], where the authors avoid some practical and theoretical difficulties by using as a model

$$u_t = \operatorname{div}(g(|\nabla(G_\sigma^* u)|)\nabla u) \qquad \text{in } \mathbb{R}^N \times \mathbb{R}^+,$$

$$u(x,0) = u_0(x) \qquad \text{in } \mathbb{R}^N,$$

where G_σ is a smoothing kernel (for instance, a Gaussian). This model avoids the occurrence of spurious edges produced by noise influence, and it is more consistent from a theoretical point of view.

However, these two models do not have a clear geometric interpretation because the term inside the divergence is hybrid and combines the estimate of the gradient and the gradient itself. The intuitive idea of edges is that they are generally piecewise smooth. Therefore, it seems natural to use a directional smoothing kernel such that it diffuses more in the direction parallel to the edge and less in the perpendicular one. Following this idea, Alvarez, Lions, and Morel [1] proposed as a model the equation

$$u_t = g(|\nabla(G_\sigma^* u)|)u_{\xi\xi} \qquad \text{in } \mathbb{R}^N \times \mathbb{R}^+,$$

$$u(x,0) = u_0(x) \qquad \text{in } \mathbb{R}^N,$$

where $\xi = \xi(x)$ represents the direction perpendicular to the gradient $\nabla u(x)$. In this model, the authors keep the improvement introduced in the above model and they introduce a new directional diffusion notion. Roughly speaking, this model conserves the exact location and sharpness of the edge, while smoothing the picture on both sides on this edge. In the particular case $g(s) = 1$ for any $s \geq 0$, this equation has received a lot of attention because of its geometrical interpretation; indeed (see, e.g., Osher and Sethian [11] and Evans and Spruck [5]) the level set of the solution moves in the normal direction with a speed proportional to their mean curvature.

In a more recent paper, Alvarez, Guichard, Lions, and Morel [2] generalize the above equation by using an axiomatic approach to the low-pass filtering theory. They show the exact shape of any local morphological low-pass filter.

Our model. In a natural way, we merge the shock capturing techniques shown for the one-dimensional case in §2 and the directional smoothing ideas presented above. We propose as a model the following parabolic-hyperbolic equation:

$$u_t = C\mathcal{L}(u) - u_\eta F(G_\sigma^* u_{\eta\eta}, G_\sigma^* u_\eta) \quad \text{in } \mathbb{R}^2 \times \mathbb{R}^+, \tag{15}$$

where $\eta = \eta(x)$ is the direction of $\nabla u(x)$, $F(.\ ,.)$ verifies (2), and $G_\sigma(.\ ,.)$ is a family of two-dimensional smoothing kernels, $C > 0$ is any positive constant, and $\mathcal{L}(u)$ is any directional smoothing operator presented above. In this paper we use, as a "canonical" example of $\mathcal{L}(u)$, the following operator:

$$\mathcal{L}(u) = u_{\xi\xi},$$

where $\xi = \xi(x)$ is the direction perpendicular to the gradient $\nabla u(x)$.

Roughly speaking, the interpretation of the terms of (4) with $\mathcal{L}(u)$ given above are as follows:

(a) The term $u_{\xi\xi}$ represents a degenerate diffusion term, which diffuses u in the direction orthogonal to its gradient ∇u. The aim of this term is to make u smooth on both sides of an edge with a minimal smoothing of the edge itself. Moreover this operator tends to move the level surfaces of the picture in the normal direction with a velocity proportional to the curvature of the level surfaces.

(b) The linear smoothing kernel G_σ is introduced in order to avoid noise influence in the process by merging different scales.

(c) The term $-u_\eta F(G_\sigma^* u_{\eta\eta}, G_\sigma^* u_\eta)$ produces edge enhancement and deconvolution by developing shocks in the position of the zero crossing of $G_\sigma^* u_{\eta\eta}$ in the direction of the gradient.

(d) Finally, the constant $C > 0$ represents the "balance" between the effect of anisotropic diffusion and shock filters. We remark that these two procedures are complementary in a natural way.

Discretization of (15). We discretize (15) by using a fast, recursive, and unconditionally stable scheme. With respect to the shock filter term, we follow the same ideas that we developed for the one-dimensional case. For the anisotropic diffusion term, we use the kind of schemes shown in Alvarez, Lions, and Morel [1].

Let k and h be the temporal and spatial increments, $i = (i_1, i_2) \in \mathbb{I} \times \mathbb{I}$, and $x_i = (i_1 h, i_2 h)$. Let u_i^n be an approximation of $u(x_i, nk)$ and ∇u_i be an approximation of $\nabla u(x_i, nk)$. Let $B \subset \mathbb{I} \times \mathbb{I}$ be a neighborhood of $(0,0)$. Let $j : \mathbb{I} \times \mathbb{I} \longrightarrow B$ be such that for each i, $j(i)$ maximizes the function $E_i : B \longrightarrow \mathbb{R}$:

$$E_i(j) = \frac{(\nabla u_i^n, j)}{\| j \|}.$$

This means that $j(i)$ is the best approximation to the orientation of ∇u_i^n in the set B. We also define the function $1 : \mathbb{I} \times \mathbb{I} \longrightarrow \mathbb{I} \times \mathbb{I}$, where $1(i) = (j_2(i), -j_1(i))$ (i.e., the direction perpendicular to $j(i)$).

Remark 4. The set $B \subset \mathbb{I} \times \mathbb{I}$ represents the number of directions that we use in order to approximate the vector ∇u_i^n in the lattice $\mathbb{I} \times \mathbb{I}$. For instance, in the numerical experiments that we show in §4 we use as B the following set:

$$B = \{ i \in \mathbb{I} \times \mathbb{I} : \| i \|^2 \leq 5 \}.$$

Next, we discretize (15) in the following way:

$$\begin{aligned} \frac{u_i^{n+1} - u_i^n}{k} &= C \frac{u_{i-1(i)}^{n+1} - 2u_i^{n+1} + u_{i+1(i)}^{n+1}}{\| 1(i) \|^2 h^2} \\ &\quad - \frac{F^-(u_i^n)(u_{i+j(i)}^{n+1} - u_i^{n+1}) + F^+(u_i^n)(u_i^{n+1} - u_{i-j(i)}^{n+1})}{\| j(i) \| h} \end{aligned}$$

for any $i \in \mathbb{I} \times \mathbb{I}$, where $F^-(s) = \min\{0, F(s)\}$, $F^+(s) = \max\{0, F(s)\}$, and $F(u_i^n) = F((G_\sigma^* u_i^n)_{\eta\eta}, (G_\sigma^* u_i^n)_\eta)$.

In order to calculate the sequence u_i^{n+1} for any $i \in \mathbb{I} \times \mathbb{I}$ we use the following fixed-point equation:

$$(16) \qquad w_i = \lambda_i b_i + \alpha_i (w_{i-1(i)} + w_{i+1(i)}) + \nu_i w_{i+j(i)} + \eta_i w_{i-j(i)} \quad \text{for any } i \in \mathbb{I} \times \mathbb{I}.$$

where $b_i = u_i^n$, $w_i = u_i^{n+1}$,

$$\lambda_i = \frac{1}{1 + 2kC/\|1\|^2 h^2 + (kF^+(u_i^n) - kF^-(u_i^n))/\| j \| h}, \qquad \alpha_i = \lambda_i \frac{kC}{\|1\|^2 h^2},$$

$$\nu_i = -\lambda_i \frac{kF^-(u_i^n)}{\| j \| h} \quad \text{and} \quad \eta_i = \lambda_i \frac{kF^+(u_i^n)}{\| j \| h} \quad \text{for any } i \in \mathbb{I} \times \mathbb{I}.$$

Therefore, we can interpret u_i^{n+1} as the solution w_i of (16).

THEOREM 2. *Let b_i be a bounded sequence, and $\lambda_i, \alpha_i, \nu_i, \eta_i$ such that $\nu_i \eta_i = 0, \lambda_i > \varepsilon, 0 \leq \alpha_i, \nu_i, \eta_i < 1$, and $\lambda_i + 2\alpha_i + \nu_i + \eta_i = 1$ for any $i \in \mathbb{I} \times \mathbb{I}$ and some $\varepsilon > 0$. Then we have the following:*

(i) *The fixed-point scheme* (16) *has a unique bounded solution w_i.*

(ii) *Let $m = \inf\{b_i, i \in \mathbb{I} \times \mathbb{I}\}$ and $M = \sup\{b_i, i \in \mathbb{I} \times \mathbb{I}\}$. Then $m \leq w_i \leq M$ for any $i \in \mathbb{I} \times \mathbb{I}$.*

Proof. (i) Let E be the space of bounded sequence $z = \{z_i \in \mathbb{R}, i \in \mathbb{I} \times \mathbb{I}\}$ and $Q : E \longrightarrow E$ defined by

$$Q(z)_i = \lambda_i b_i + \alpha_i(z_{i-1(i)} + z_{i+1(i)}) + \nu_i z_{i+j(i)} + \eta_i z_{i-j(i)}.$$

Let z, $v \in E$; then we have

$$\sup_{i \in \mathbb{I} \times \mathbb{I}} \{|Q(z)_i - Q(v)_i|\}$$

$$\leq \sup_{i \in \mathbb{I} \times \mathbb{I}} \{|\alpha i(z_{i-1(i)} - v_{i-1(i)} + z_{i+1(i)} - v_{i+1(i)}) + \nu_i(z_{i+j(i)} - v_{i+j(i)}) + \eta_i(z_{i-j(i)} - v_{i-j(i)})|\}$$

$$\leq (1-\varepsilon) \sup_{i \in \mathbb{I} \times \mathbb{I}} \{|z_i - v_i|\}.$$

Therefore, Q is a contractive function and (i) follows by using the classic Banach fixed-point theorem. Moreover, if $z^0 \in E$, then the sequence $z^k \in E$ defined by $z^k = Q(z^{k-1})$ converges to the solution of (14).

(ii) Let $z_i^0 = b_i$ for any $i \in \mathbb{I}$, $m = \inf\{b_i, i \in \mathbb{I} \times \mathbb{I}\}$, $M = \sup\{b_i, i \in \mathbb{I} \times \mathbb{I}\}$, and z^k defined above. Since λ_i, α_i, ν_i, $\eta_i \geq 0$ and $\lambda_i + 2\alpha_i + \nu_i + \eta_i = 1$ for any $i \in \mathbb{I} \times \mathbb{I}$, we have that $m \leq z_i^k \leq M$ for any $i \in \mathbb{I} \times \mathbb{I}$ and $k \in \mathbb{N}$, and (ii) follows by passing to the limit when $k \to +\infty$.

Remark 5. In order to accelerate the convergence to the solution $w \in E$ of the fixed-point equation (16), we use a recursive scheme based on the Gauss–Seidel algorithm to solve linear systems. Let $z^0 = b \in E$; we define $z^k \in E$, for any $k \in \mathbb{N}$, to be an approximation of w. Assuming that z^{k-1} is known, we define z^k as

$$z_i^k = \lambda_i b_i + \alpha_i(z_{i-1(i)}^{k'} + z_{i+1(i)}^{k'}) + \nu_i z_{i+j(i)}^{k'} + \eta_i z_{i-j(i)}^{k'} \tag{17}$$

for any $i = (i_1, i_2) \in \mathbb{I} \times \mathbb{I}$, where k' is k or $k-1$ according to the method we use to go through the index $i \in \mathbb{I} \times \mathbb{I}$. For instance, in the numerical experiments that we show in §4, we go through $\mathbb{I} \times \mathbb{I}$ in a decreasing or increasing way according to the value of k (k odd or even).

Remark 6 (estimation of ∇u_i, removing noise influence). We need an estimation of ∇u_i in our scheme in order to obtain the functions $j(i)$ and $1(i)$. A simple estimate of ∇u_i is given by

$$u_{i_x} = \frac{u_{i+(1,0)} - u_{i-(1,0)}}{2h} \quad \text{and} \quad u_{i_y} = \frac{u_{i+(0,1)} - u_{i-(0,1)}}{2h}.$$

This estimate gives satisfactory results in the case of smooth functions. However, in the case of discontinuous functions as a noisy picture, it could be interesting to replace ∇u_i by a local dominant orientation. One may think that it is sufficient to filter (u_{ix}, u_{iy}) with a linear smoothing kernel. However, this will not work for one important reason: in our analysis the opposite directions (u_{ix}, u_{iy}) and $(-u_{ix} - u_{iy})$ are equivalents, but the convolution with a linear smoothing kernel will cancel each other out. One way to solve this problem is as follows.

Let $G_\sigma(. , .)$ be a two-dimensional smoothing kernel. We define $\theta(i)$ to be the orientation of ∇u_i as the angle θ that maximizes the function

$$E(\theta) = G_\sigma^*(u_{ix} \cos\theta + u_{iy} \sin\theta)^2.$$

This estimate eliminates noise influence if we assume, in a natural way, that the noise does not have any special orientation. On the other hand, it is not difficult to show that $\theta(i)$ satisfies

$$\tan(2\theta(i)) = \frac{2G_\sigma^*(u_{i_x} u_{i_y})}{G_\sigma^*(u_{i_x}^2 - u_{i_y}^2)}.$$

A similar estimate was used in Kass and Witkin [7] and Ravishankar and Schunck [13] for analyzing oriented structures in an image.

4. Experimental results. We present experiments on three images of varying difficulty which demonstrate the merits (or demerits) of the model (15) and its associated algorithms. The first image is a synthetic picture (128 $\times$ 128 pixels) which consists of a triangle above a narrow rectangle. We have introduced in this picture a noise generated by a random average of two convolutions of the original image with Gaussian functions $G_\sigma(. , .)$ and with different standard deviation ($\sigma = 7$ and $\sigma = 14$; see Fig. 4). Of course, this is a very special noise and it cannot be modelled by the classical methods for picture restoration. We have chosen this kind of noise to point out the fact that our method does not use any a priori information about the noise introduced in the image.

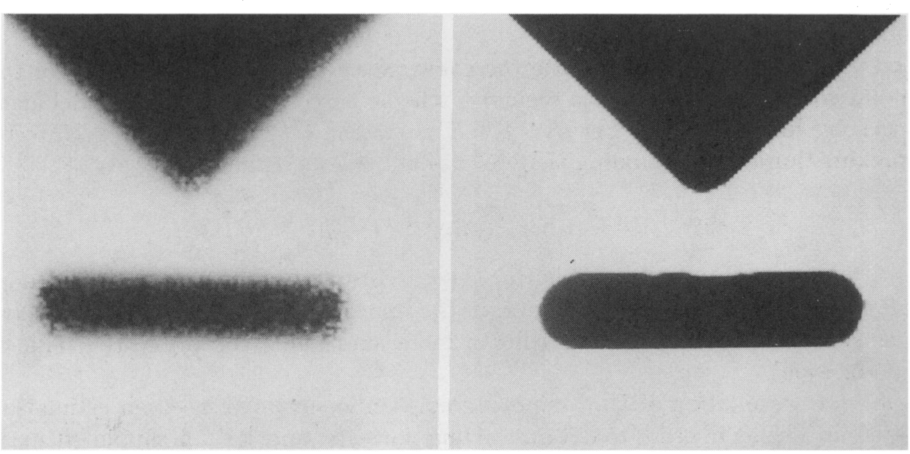

FIG. 4. *In the left part we present the noisy image and in the right part we present the result of* 15 *iterations of the scheme* (17). *The time discretization step is* $4, \sigma = 10$ *and* $C = 2$.

In the second experiment we present a medical image of the brain (512 $\times$ 512 pixels). In this image (see Fig. 5), which is actually extracted from a series of slices, we have introduced a standard noise which involves both blur and noise. The blur is generated by convolving with a Gaussian function $G_\sigma(. \, , .)$ with standard deviation $\sigma = 6$. White Gaussian noise was added with means zero and standard deviation $\sigma = 10$.

Finally, in Fig. 6, we present a satellite image where we have not introduced any special noise. We point out here the nontrivial asymptotic state of the solution of the mathematical model (15).

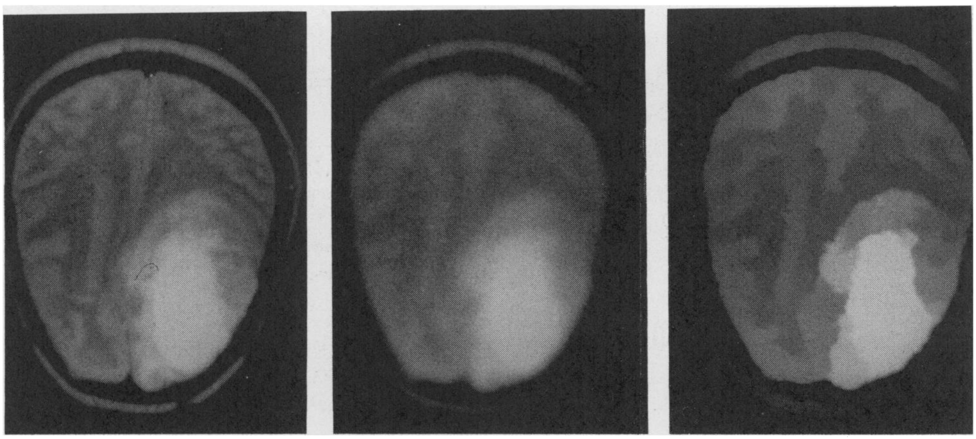

FIG. 5. *In the left we present the original image. In the center we present the noisy image, and in the right, we present the result of five iterations of the scheme* (17). *The time discretization step is* $5, \sigma = 3$ *and* $C = 1$.

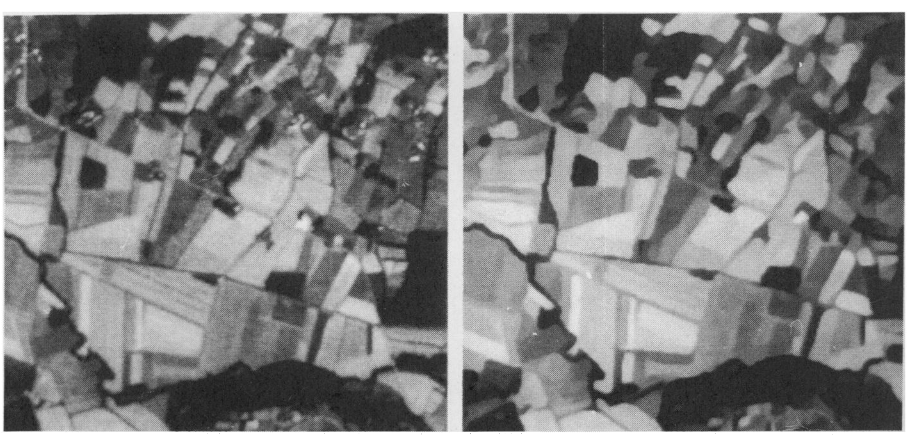

FIG. 6. *In the left part we present the original image. In the right part we present the asymptotic state image by using the scheme* (17). *The time discretization step is* $1, \sigma = 3$ *and* $C = 1$. (*The number of iterations to obtain the asymptotic state is* 20.)

REFERENCES

[1] L. ALVAREZ, P. L. LIONS, AND J. M. MOREL, *Image selective smoothing and edge detection by nonlinear diffusion* (II), SIAM J. Numer. Anal., 29 (1992), pp. 845–866.

[2] L. ALVAREZ, F. GUICHARD, P. L. LIONS, AND J. M. MOREL, *Axiomatic and fundamental equations of image processing*, Arch. Rational Mech. Anal., 123 (1993), pp. 199–257.

[3] F. CATTE, T. COLL, P. L. LIONS, AND J. M. MOREL, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal., 29 (1992), pp. 182–193.

[4] R. DERICHE, *Fast algorithms for low-level vision*, IEEE Trans. Pattern Anal. Machine Intelligence, 12 (1990), pp. 78–87.

[5] L. C. EVANS AND J. SPRUCK, *Motion of level sets by mean curvature* I, preprint.

[6] C. GASQUET AND P. WITOMSKI, *Analyse de Fourier et applications*, Masson Editeur, Paris, France, 1990.

[7] M. KASS AND A. WITKIN, *Analyzing oriented patterns*, Comput. Vision Graphics Image Process., 37 (1987), pp. 362–385.

[8] D. MARR, *Vision*, Freeman and Co., San Francisco, CA, 1982.

[9] J. KOENDERINK, *The structure of images*, Biol. Cybern., 50 (1984), pp. 363–370.

[10] S. OSHER AND L. RUDIN, *Feature-oriented image enhancement using shock filters*, SIAM J. Numer. Anal., 27 (1990), pp. 919–940.

[11] S. OSHER AND J. SETHIAN, *Fronts propagating with curvature dependent speed: Algorithms based on the Hamilton–Jacobi formulation*, J. Comp. Phys., 79 (1988), pp. 12–49.

[12] P. PERONA AND J. MALIK, *Scale space and edge detection for visual scene analysis*, Proc. IEEE Computer Soc. Workshop on Computer Vision, 1987.

[13] A. RAVISHANKAR AND B. G. SCHUNCK, *Computing oriented texture fields*, CVGIP, Graphical Models and Image Processing, 53 (1991), pp. 157–185.

[14] L. RUDIN, *Images, numerical analysis of singularities, and shock filters*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1987.

[15] J. SHEN AND S. CASTAN, *An optimal linear operator for edge detection*, in Proc. CVPR, Miami, FL, June 1986, pp. 109–114.

[16] G. A. SOD, *Numerical Methods in Fluid Dynamics*, Cambridge University Press, Cambridge, UK, 1985.