SIAM J. APPL. MATH.
Vol. 57, No. 1, pp. 153–175, February 1997

008

IMAGE QUANTIZATION USING REACTION-DIFFUSION EQUATIONS*

LUIS ALVAREZ† AND JULIO ESCLARÍN†

Abstract. In this paper we present an image quantization model based on a reaction-diffusion partial differential equation. The quantized image is given by the asymptotic state of this equation. Existence and uniqueness of the solution are proved in the framework of viscosity solutions. We introduce an L^∞ stable algorithm in order to compute numerically the solution of the equation, and some experimental results are shown. A new energy functional based on the classical Lloyd method is used to compute the quantizer codewords.

Key words. image quantization, partial differential equation, viscosity solutions, multiscale vision model

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

PII. S0036139994277580

1. Introduction.

In recent years, partial differential equations (PDEs) have become solid and useful tools for image processing. We present new techniques based on PDEs for image processing and more specifically for image quantization. We consider an image as any bounded function $u : \mathbb{R}^2 \longrightarrow \mathbb{R}$. A quantizer Q_S is a rule to associate two finite sets with u : $\{u_k\}_{k=1,\dots,S}$, which represents the quantizer codewords, and $\{t_k\}_{k=1,\dots,S+1}$, which represents the quantizer separators, satisfying

$$t_1 < u_1 < t_2 < \cdots < u_S < t_{S+1}.$$

Each quantizer Q_S then generates a quantized image of u by replacing any value of the image in the interval $[t_k, t_{k+1})$ by u_k. In what follows we will consider t_1 and t_{S+1} to be fixed a priori, following the bounds of the image u.

In this paper we address two problems related to image quantization. The first one is the choice of the quantizer Q_S for a given number of codeword levels S. The second one is to introduce a denoising procedure which acts at the same time as the quantization procedure.

The classical method (see Lloyd [15]) of choosing the quantizer Q_S consists of minimizing the average quadratic error. Let $H(s)$ be the probability distribution associated to the image u given by

$$H(s) = P\{u(x,y) < s\}.$$

We consider that u is periodic with period $[a,b) \times [c,d)$, which represents the original image domain; then P can be considered as the Lebesgue measure on $[a,b) \times [c,d)$. We notice that the histogram $h(s)$ of u is given by $h(s) = H'(s)$. The Lloyd quantizer $\tilde{Q}_S$ minimizes the quadratic energy:

$$E(Q_S) = \sum_{k=1}^{S} \int_{t_k}^{t_{k+1}} (s - u_k)^2 dH(s). \tag{1}$$

*Received by the editors November 24, 1994; accepted for publication (in revised form) November 3, 1995. This research was supported in part by C.E. Human Capital and Mobility program ERBCHRXCT930095 and CICYT Spanish Project TIC92-1282.

http://www.siam.org/journals/siap/57-1/27758.html

†Departamento de Informática y Sistemas, Universidad de Las Palmas de Gran Canaria, Campus de Tafira, 35017 Las Palmas, Spain (luis@amihp710.dis.ulpgc.es and jesclarin@dis.ulpgc.es).

In the framework of image processing, it is not desirable that two levels t_k, t_{k+1} or u_k, u_{k+1} be too close to each other because perceptually they represent the same level. We introduce two terms in the Lloyd energy in order to penalize quantization levels for being too close. We will use in our experiments the quantizer $\hat{Q}_S$ which minimizes the energy

$$E(Q_S) = \sum_{k=1}^{S}\left(\int_{t_k}^{t_{k+1}} (s-u_k)^2 dH(s) + \frac{L}{t_{k+1}-t_k} + C\left(\frac{t_k+t_{k+1}}{2} - u_k\right)^2\right), \tag{2}$$

where L, C are positive constants. We will use dynamic programming techniques to compute the minimum $\hat{Q}_S$ of the above energy.

The second and main step in our quantization method is to smooth out the noise in the original image. This will be done by solving the initial value problem for PDEs of the reaction-diffusion type. Before explaining this procedure, let us observe that there is a way of generating a quantized image via solving an ordinary differential equation. Let $f \in C^0(\mathbb{R})$ be a function satisfying that $f(s) = 0$ if and only if $s = t_1, u_1, \ldots, u_S, t_{S+1}$ and that $\frac{\partial f}{\partial s}(t_k) > 0$ and $\frac{\partial f}{\partial s}(u_k) < 0$ for any k. We solve the initial value problem for the ordinary differential equation

$$u_t = f(u)$$

with the original image $u_0(x, y)$ as the initial datum. That is, at the end of the process of solving the initial value problem for the above ordinary differential equation the gray levels of the image are reduced to one of the values $u_1, \ldots, u_S$, which are attractors of the above ordinary differential equation, and the values $t_1, \ldots, t_{S+1}$ among the zeros of $f(u)$ act as separator levels in the process. Of course, this method does not take care of reducing any noise in the image.

In place of the ordinary differential equation in this simple method, as mentioned above, our model utilizes a PDE of the reaction-diffusion type. We make use of the diffusion effect of the PDE as a way of reducing the noise in the data of the image. The diffusion effect is considered to act on the data as a regularization or homogenization. In other words, we require that the gray levels of neighbor pixels verify a kind of homogeneity, and this is realized via solving the PDE in our model. This kind of homogenization of data is carried out by means of filtering. The linear filtering is usually done by convolution with Gaussian kernels of increasing variance, as proposed in [12], [16], [20]. Koenderink [13] noticed that in each scale the convolution of a signal with a Gaussian is equivalent to the solution of the heat equation, with the signal as the initial datum.

This datum being called u_0, the following images are obtained by solving the heat equation

$$\frac{\partial u(x,y,t)}{\partial t} = \Delta u(x,y,t) \qquad \text{in } \mathbb{R}^2 \times [0,+\infty[,$$

$$u(x,y,0) = u_0(x,y) \qquad \text{in } \mathbb{R}^2.$$

The resolution of this equation for an initial datum with bounded quadratic norm is $u(t,x,y) = G_t * u_0$, where

$$G_t(x,y) = Ct^{-1}\exp\left(-\frac{(x^2+y^2)}{4t}\right)$$

is the Gaussian function; here t represents a scale parameter.

Along the scales, the linearity of the Laplacian operator produces a displacement or loss of the edges in the regions that are present in the image, which makes it necessary to find nonlinear operators.

Perona and Malik [18] have proposed that the Laplacian operator should be replaced by the operator

$$\operatorname{div}(g(\|\nabla u\|)\nabla u),$$

and they use the equation

$$u_t = \operatorname{div}(g(\|\nabla u\|)\nabla u)$$

as a model, where $g(\cdot)$ is a positive nonincreasing function.

With this model they tried to penalize diffusion where the gradient is large. However, this model presents some disadvantages: the first one is that it does not work properly when noise is present (noise remains). The second disadvantage is that from a theoretical point of view, the equation is well posed only if $sg(s)$ is a nondecreasing function.

Later Alvarez, Lions, and Morel [2] proposed a model based on the equation of the mean curvature flow

$$u_t = g(|G * \nabla u|)\|\nabla u\| \operatorname{div}\left(\frac{\nabla u}{\|\nabla u\|}\right).$$

This equation produces a diffusion in the direction of the edges which corresponds to the orthogonal direction to the gradient. It does not produce a blurring effect in the image as the previous equations did because it does not alter the contrast between different objects present in the image.

In a similar way, Cottet and Germain [7] have introduced a model based on reaction-diffusion equations. The idea behind this model consists of combining the diffusion for noise filtration and the reaction to improve the contrast. Its purpose is to describe a model in which a processed image can be seen as the asymptotic state of the resolution of the mathematical model

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

where Ω is a bounded domain in $\mathbb{R}^2$, $u \in L^2(\Omega)$, and $f(x) \in C^1$ satisfies

$$f(\pm 1) = 0, xf(x) > 0 \text{ for } x \in (-1,0) \cup (0,1).$$

The nonlinear operator A_ε is defined by the 2×2 matrix

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

where $\tilde{\partial}_1 = \frac{\partial}{\partial_{x_2}}$ and $\tilde{\partial}_2 = -\frac{\partial}{\partial_{x_1}}$. Hence, the term ε^2 is used to avoid singularities when $\|\nabla u\| = 0$ in the derivatives of this matrix. A_ε projects vectors onto the orthogonal complement of the gradient.

The selection of the ε parameter poses some problems: if ε tends to zero, u_ε tends to u, and as a consequence there is no diffusion since the term of diffusion becomes zero. On the other hand, if ε is large a distortion is introduced. Moreover, the geometrical interpretation of the operator is not clear.

Of course, there are a lot of ways to define nonlinear filtering using PDEs. In [1] and [3], the reader can find a complete study of the applications of PDEs to multiscale analysis theory using an axiomatic approach.

The method proposed in this paper is also based on reaction-diffusion equations. In particular, we propose the following equation:

$$\frac{\partial u}{\partial t} - g(|u * \nabla G|)\|\nabla u\| \operatorname{div}\left(\frac{\nabla u}{\|\nabla u\|}\right) - f(u) = 0 \qquad \text{in } [0,+\infty[\times\mathbb{R}^N, \tag{3}$$
$$u(0,x) = u_0(x),$$

where $g(x) \geq 0$ is a nonincreasing function satisfying $g(0) = 1$, G is a convolution kernel (for example, a Gaussian function), and $f(u)$ is a Lipschitz function with a finite number of zeros.

The interpretation of the terms in the equation is as follows:

(a) The term

$$\|\nabla u\| \operatorname{div}\left(\frac{\nabla u}{\|\nabla u\|}\right) = \triangle u - \frac{\nabla^2 u(\nabla u, \nabla u)}{\|\nabla u\|^2}$$

represents an effect of degenerate diffusion that diffuses u in the orthogonal direction to the gradient ∇u and does not diffuse in any other direction. The role of this degenerate term of diffusion is to regularize u on both sides of an edge with the minimum loss of the edge itself (an edge is defined as a line along which the gradient is large).

(b) The term $g(|G * \nabla u|)$ is used for "enhancement" of the edges. It certainly controls diffusion speed: if ∇u has a small mean in a neighborhood of a point x, this point x is considered as an interior point of a smooth region of the image and the diffusion is therefore strong. If ∇u has a large mean value in the neighborhood of x, x is considered as an edge point and the diffusion speed is lowered, since $g(s)$ is small for large s.

(c) The function $f(s)$ determines the asymptotic state of the equation using the quantizer $\hat{Q}_S$ which minimizes the energy (2), as we have explained above.

Using equation (3), we introduce a smoothing procedure in the quantization rule. So the asymptotic solution of this equation when t tends to ∞ represents a smooth quantization of the original picture.

The paper is organized as follows. In section 2 the mathematical validation of our model is presented. We show existence and uniqueness of the solution of equation (3) in the framework of viscosity solution theory. In section 3 we develop the numerical analysis of our model. We develop an algorithm to compute the global minimum $\hat{Q}_S$ of the functional energy (2) using dynamic programming techniques and a discretization scheme for equation (3). We also show the L^∞ stability of the scheme. Finally, in section 4, we give some experimental results.

2. Mathematical model analysis. In what follows we will use the notation $\partial_i u = \frac{\partial u}{\partial x_i} = u_i$ and $\partial_{ij} u = \frac{\partial u^2}{\partial x_i \partial x_j} = u_{ij}$.

The mathematical model that we propose is the following:

$$\frac{\partial u}{\partial t} - g(u * \nabla G) \sum_{i,j} a_{ij}(\nabla u)\partial_{ij} u - f(u) = 0 \quad \text{in } [0,+\infty[\times\mathbb{R}^N, \tag{4}$$
$$u(0,x) = u_0(x),$$

where

(1) $g^{\frac{1}{2}}(\mathbb{R}^N, \mathbb{R})$ is Lipschitz, $g(p) \geq 0 \; \forall p \in \mathbb{R}^N$;

(2) $D^\alpha G \in L^1(\mathbb{R}^N) \; \forall |\alpha| \leq 2$;

(3) $\sum_{i,j} a_{ij}(p)\xi_i\xi_j \geq 0 \;\; \forall p \in \mathbb{R}^N \setminus \{0\}, \xi \in \mathbb{R}^N, \quad a_{ij} = a_{ji}$;

(4) a_{ij} are continuous and bounded on $\mathbb{R}^N \setminus \{0\}$;
(5) $f(x) \in C^0(\mathbb{R})$, $f'(x) \in L^\infty(\mathbb{R})$ and an interval $[a, b]$ exists such that

$$f(x) > 0 \text{ if } x < a \text{ and } f(x) < 0 \text{ if } x > b.$$

We begin with a brief summary of the definition of viscosity solutions of (4) on $\mathbb{R}^N$.

Let u be in $C([0,T]\times\mathbb{R}^N)$ for some T in $]0,+\infty[$. Then, u is a viscosity subsolution of (4) if for all Φ in $C^2([0,T]\times\mathbb{R}^N)$ the following condition holds at any point (t_0, x_0) in $]0,T]\times\mathbb{R}^N$, which is a local maximum point of $(u-\Phi)$:

$$\frac{\partial\Phi}{\partial t}(t_0,x_0) - g(u*\nabla G(t_0,x_0))\sum_{i,j} a_{ij}(\nabla\Phi(t_0,x_0))\partial_{ij}\Phi(t_0,x_0) - f(u(t_0,x_0)) \le 0$$
$$\text{if } \nabla\Phi(t_0,x_0) \neq 0,$$

$$\frac{\partial\Phi}{\partial t}(t_0,x_0) - g(u*\nabla G(t_0,x_0))\limsup_{p\to 0}\sum_{i,j} a_{ij}(p)\partial_{ij}\Phi(t_0,x_0) - f(u(t_0,x_0)) \le 0$$
$$\text{if } \nabla\Phi(t_0,x_0) = 0.$$

We define a viscosity supersolution in a similar manner by replacing "local maximum point" by "local minimum point," "$\le$" by "$\ge$," and "lim sup" by "lim inf." Finally, a viscosity solution is a function which is a subsolution and a supersolution. In what follows, we will suppose that there exists $h = (h_1, \dots, h_N)$ such that

$$(5) \qquad u_0(x + (0,0,\dots,h_k,\dots,0)) = u_0(x) \;\; \forall x \in R^N \;\text{ and }\; k = 0,1,\dots,N.$$

It means that u_0 is a periodic function.

LEMMA 2.1. *Let u be a viscosity solution of* (4) *satisfying* (5). *Then for any $t > 0$ and $x \in \mathbb{R}^N$*

$$\min\Big\{\inf_{\mathbb{R}^N}|u_0|, a\Big\} \le u(t,x) \le \max\Big\{\sup_{\mathbb{R}^N}|u_0|, b\Big\}.$$

Proof. Fix $T > 0$. We will show that the above inequality holds for $(t,x) \in [0,T]\times R^N$. Note that the periodicity assumption on u guarantees that u has a maximum over $[0,T]\times R^N$.

Let (t_0,x_0) be a maximum point of $u(t,x)$. Let us suppose that $t_0 > 0$. In the viscosity subsolution definition, we take $\Phi \equiv 0$, and therefore $u-\Phi$ has a local maximum at the point (t_0,x_0). So

$$-f(u(t_0,x_0)) \le 0 \Longrightarrow f(u(t_0,x_0)) \ge 0 \Longrightarrow u(t_0,x_0) \le b.$$

By repeating the same argument with a viscosity supersolution, if (t_0,x_0) is a local minimum of $u-\Phi$ we obtain

$$-f(u(t_0,x_0)) \ge 0 \Longrightarrow f(u(t_0,x_0)) \le 0 \Longrightarrow a \le u(t_0,x_0).$$

If $t_0 = 0$ we have

$$\inf_{\mathbb{R}^N}|u_0| \le u(x,0) \le \sup_{\mathbb{R}^N}|u_0|;$$

then joining together both inequalities we have

$$\min\Big\{\inf_{\mathbb{R}^N}|u_0|, a\Big\} \le u(x,t) \le \max\Big\{\sup_{\mathbb{R}^N}|u_0|, b\Big\}. \qquad \square$$

THEOREM 2.2. *Let u_0 be Lipschitz on $\mathbb{R}^N$.*

(i) *Equation* (4) *has a unique viscosity solution u in $C([0,T]\times\mathbb{R}^N)\cap L^\infty(0,T;W^{1,\infty}(\mathbb{R}^N))$ for any $T<\infty$.*

(ii) *Let $u\in C([0,T]\times\mathbb{R}^N)\cap L^\infty(0,T;W^{1,\infty}(\mathbb{R}^N))$, $v\in C([0,\infty)\times\mathbb{R}^N)$ be viscosity solutions of problem* (4) *satisfying* (5) *with u_0, v_0 as initial data. Then for each $T\in[0,+\infty)$ there exists a constant K which depends only on $\|u_0\|_{W^{1,\infty}}$,$\|v_0\|_{L^\infty}$ and on the Lipschitz constant of u (and also on T) such that*

$$\sup_{0\le t\le T}|u(t,\cdot)-v(t,\cdot)|_{L^\infty(\mathbb{R}^N)}\le K\|u_0-v_0\|_{L^\infty(\mathbb{R}^N)}. \tag{6}$$

To prove the theorem we shall use the techniques of viscosity solutions theory [2], [4], [8], [9], [10], and [19]. We shall start by proving the result of uniqueness (inequality (6)). In order to do that, we shall study some properties of the maximum point (t_0,x_0,y_0) of the function

$$u(t,x)-v(t,y)-\frac{|x-y|^4}{4\varepsilon}-\lambda t \qquad t\in[0,T],\ \ x,y\in\mathbb{R}^N,\ T,\varepsilon,\lambda\in]0,+\infty[. \tag{7}$$

Notice that since $u,v\in C([0,\infty[\times\mathbb{R}^N)$ are periodic functions, such a maximum point (t_0,x_0,y_0) is always attained.

LEMMA 2.3. *Let (t_0,x_0,y_0) be a maximum point of* (7) *and let L be the Lipschitz constant of u. Then*

$$|x_0-y_0|\le(4\varepsilon L)^{\frac{1}{3}}.$$

Proof. $u(t_0,x_0)-v(t_0,y_0)-\frac{|x_0-y_0|^4}{4\varepsilon}-\lambda t_0\ge u(t_0,y_0)-v(t_0,y_0)-\lambda t_0\Longrightarrow\frac{|x_0-y_0|^4}{4\varepsilon}\le|u(t_0,x_0)-u(t_0,y_0)|\le L|x_0-y_0|\Longrightarrow|x_0-y_0|\le(4\varepsilon L)^{\frac{1}{3}}$. □

LEMMA 2.4. *Let (t_0,x_0,y_0) be a maximum point of* (7). *For all $\delta>0$ there exists $M>0$ such that if $\varepsilon=(\delta\ \sup_{\mathbb{R}^N\times[0,T]}|u-v|)^{\frac{1}{3}}$ and $\lambda=M\ \sup_{\mathbb{R}^N\times[0,T]}|u-v|$ then the maximum point of* (7) *is obtained for $t_0=0$.*

Proof. First let us suppose that $t_0>0$. Using the proof of the uniqueness and existence theorem of [9] we find for all $\lambda,\mu>0$ real numbers c,d, and two $(N\times N)$ symmetric matrices X,Y such that

$$\begin{gathered} c-d=\lambda,\\ \begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix}\le\begin{pmatrix} A+\mu A^2 & -A-\mu A^2\\ -A-\mu A^2 & A+\mu A^2\end{pmatrix}.\end{gathered} \tag{8}$$

$$\begin{aligned} &c-g(u*\nabla G)(t_0,x_0)\sum_{i,j}a_{ij}(\varepsilon^{-1}|x_0-y_0|^2(x_0-y_0))X_{ij}-f(u(t_0,x_0))\le 0,\\ &d-g(v*\nabla G)(t_0,y_0)\sum_{i,j}a_{ij}(\varepsilon^{-1}|x_0-y_0|^2(x_0-y_0))Y_{ij}-f(v(t_0,y_0))\ge 0,\end{aligned} \tag{9}$$

where

$$A=\frac{1}{\varepsilon}|x_0-y_0|^2Id+\frac{2}{\varepsilon}(x_0-y_0)\otimes(x_0-y_0)$$

and

$$A^2=\frac{1}{\varepsilon^2}|x_0-y_0|^4Id+\frac{8}{\varepsilon^2}|x_0-y_0|^2(x_0-y_0)\otimes(x_0-y_0).$$

Now we shall demonstrate that if we suppose

$$\lambda > L_f \sup_{\mathbb{R}^N \times [0,T]} |u - v|,$$

then $x_0 \neq y_0$ where L_f is the Lipschitz constant of f in the interval,

$$[\min\{a, \inf |u_0|, \inf |v_0|\}, \max\{b, \sup |u_0|, \sup |v_0|\}].$$

By Lemma 2.1 we have that

$$\min\{a, \inf |u_0|, \inf |v_0|\} \leq u, v \leq \max\{b, \sup |u_0|, \sup |v_0|\}.$$

If $x_0 = y_0$ then $A \equiv 0$, which means that $X \leq 0$ and $Y \geq 0$. Then from (9) we obtain

$$c - f(u(t_0, x_0)) \leq 0,$$
$$d - f(v(t_0, y_0)) \geq 0$$

and therefore

$$\lambda = c - d \leq f(u(t_0, x_0)) - f(v(t_0, y_0)) \leq L_f \sup_{\mathbb{R}^N \times [0,T]} |u - v|.$$

Hence if

$$\lambda > L_f \sup_{\mathbb{R}^N \times [0,T]} |u - v|$$

then

$$x_0 \neq y_0.$$

If we put $\mu = \varepsilon |x_0 - y_0|^{-2}$ in (8), we obtain

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \frac{2}{\varepsilon} \begin{pmatrix} B & -B \\ -B & B \end{pmatrix}, \tag{10}$$

where B is $|x_0 - y_0|^2 Id + 5(x_0 - y_0) \otimes (x_0 - y_0)$.

Let Λ be the $n \times n$ matrix given by $a_{ij}(\varepsilon^{-1}|x_0 - y_0|^{-2}(x_0 - y_0))$; then $\Lambda \geq 0$.

Let $g_1 = g(u * \nabla G)(t_0, x_0)$ and $g_2 = g(v * \nabla G)(t_0, y_0)$.

Now we shall prove for the $(2n \times 2n)$ matrix

$$\Gamma = \begin{pmatrix} g_1 \Lambda & \sqrt{g_1 g_2} \Lambda \\ \sqrt{g_1 g_2} \Lambda & g_2 \Lambda \end{pmatrix}$$

that $\Gamma \geq 0$: let $\xi, \eta \in \mathbb{R}$ and compute that

$${}^t(\xi, \eta) \Gamma (\xi, \eta) = g_1 \, {}^t\xi \Lambda \xi + 2\sqrt{g_1 g_2} \, {}^t\xi \Lambda \eta + g_2 \, {}^t\eta \Lambda \eta = (\sqrt{g_1}\xi + \sqrt{g_2}\eta) \, \Lambda \, (\sqrt{g_1}\xi + \sqrt{g_2}\eta).$$

Therefore, since $\Lambda \geq 0$ we have that $\Gamma \geq 0$. If we multiply (10) by Γ and take the trace of the result, we have

$$g_1 \sum_{i,j} a_{ij} \xi_{ij} - g_2 \sum_{i,j} a_{ij} \eta_{ij} \leq 2\varepsilon^{-1} \left(g_1 - 2\sqrt{g_1 g_2} + g_2\right) \operatorname{tr}(\Lambda B)$$
$$\leq 2\varepsilon^{-1} \left(\sqrt{g_1} - \sqrt{g_2}\right)^2 \sum_{i,j} a_{ij} b_{ij} \leq 2\varepsilon^{-1} \left(\sqrt{g_1} - \sqrt{g_2}\right)^2 C_0 |x_0 - y_0|^2$$

for a certain C_0 which depends only on $(a_{ij})_{1 \leq i,j \leq N}$.

Then, taking into account (9) and that $\lambda = c - d$, we have

$$\begin{aligned} \lambda &\leq g_1 \sum_{i,j} a_{ij}\xi_{ij} - g_2 \sum_{i,j} a_{ij}\eta_{ij} - (f(u) - f(v)) \\ &\leq 2\varepsilon^{-1} C_0 \left(\sqrt{g_1} - \sqrt{g_2}\right)^2 |x_0 - y_0|^2 + L_f \sup_{[0,T]\times\mathbb{R}^N} |u - v|. \end{aligned} \tag{11}$$

Let us see an estimation of $(\sqrt{g_1} - \sqrt{g_2})$.
Since

$$\sqrt{g_1} - \sqrt{g_2} = \sqrt{g(u * \nabla G)(t_0, x_0)} - \sqrt{g(v * \nabla G)(t_0, y_0)}$$

and $\sqrt{g}$ is Lipschitz, we have

$$(\sqrt{g_1} - \sqrt{g_2}) \leq L_g |(u * \nabla G)(t_0, x_0) - (v * \nabla G)(t_0, y_0)|,$$

where L_g is the Lipschitz constant of $\sqrt{g}$.
In the inequality

$$\begin{aligned} &|(u * \nabla G)(t_0, x_0) - (v * \nabla G)(t_0, y_0)| \\ &\leq |(u * \nabla G)(t_0, x_0) - (u * \nabla G)(t_0, y_0)| + |(u * \nabla G)(t_0, y_0) - (v * \nabla G)(t_0, y_0)| \end{aligned}$$

we see

$$\begin{aligned} &|(u * \nabla G)(t_0, x_0) - (u * \nabla G)(t_0, y_0)| \\ &= \left| \int_{\mathbb{R}^N} (u(t_0, x_0 - z) - u(t_0, y_0 - z))(\nabla G)(z)\, dz \right| \\ &\leq L|x_0 - y_0| \int_{\mathbb{R}} |(\nabla G)(z)|\, dz \leq L \cdot C_G |x_0 - y_0| \end{aligned}$$

and

$$\begin{aligned} &|(u * \nabla G)(t_0, y_0) - (v * \nabla G)(t_0, y_0)| \\ &= \left| \int_{\mathbb{R}^N} (u(t_0, y_0 - z) - v(t_0, y_0 - z))(\nabla G)(z)\, dz \right| \\ &\leq \sup_{[0,T]\times\mathbb{R}^N} |u - v| \int_{\mathbb{R}} |(\nabla G)(z)|\, dz \leq C_G \sup_{[0,T]\times\mathbb{R}^N} |u - v|, \end{aligned}$$

where $C_G = \int_{\mathbb{R}^N} |(\nabla G)(z)|\, dz < +\infty$.
Therefore, $(\sqrt{g_1} - \sqrt{g_2})$ is estimated from above by

$$C_2 \Big(\sup_{[0,T]\times\mathbb{R}^N} |u - v| + |x_0 - y_0| \Big),$$

where C_2 depends only on g, the Lipschitz constant of u, $\int_{\mathbb{R}} |(\nabla G)(z)|\, dz$, and $\sup_{\mathbb{R}^N} \times [0,T] |u - v|$.
We can deduce from (11) that

$$\lambda \leq C\Bigg(\Big(\sup_{[0,T]\times\mathbb{R}^N} |u - v| \Big)^2 \frac{|x_0 - y_0|^2}{\varepsilon} + \frac{|x_0 - y_0|^4}{\varepsilon} \Bigg) + L_f \sup_{[0,T]\times\mathbb{R}^N} |u - v|,$$

where $C = 2C_0L_g^2C_2^2$. From Lemma 2.3 we can conclude the following:

$$\lambda \leq C_3\left(\|u-v\|^2_{L^\infty(0,T)}\varepsilon^{-\frac{1}{3}} + \varepsilon^{\frac{1}{3}}\right) + L_f\|u-v\|_{L^\infty(0,T)},$$

where $C_3 = \max((4L)^{\frac{2}{3}}, (4L)^{\frac{4}{3}})C$.

Without loss of generality, we can suppose that $\sup_{[0,T]\times\mathbb{R}^N} |u-v| > 0$ and take

$$\varepsilon^{\frac{1}{3}} = \delta\|u-v\|_{L^\infty(0,T)}.$$

We obtain

$$\lambda \leq \left(L_f + C_3\left(\delta + \frac{1}{\delta}\right)\right)\|u-v\|_{L^\infty(0,T)}.$$

Therefore if we take $M = L_f + C_3(\delta + \frac{1}{\delta}) + 1$ and $\lambda = M\|u-v\|_{L^\infty(0,T)}$ we have a contradiction and thus $t_0 = 0$. □

LEMMA 2.5. *If the function* (7) *attains a maximum at a point* $(t_0, x_{0,0})$ *with* $t_0 = 0$, *then*

$$\|u-v\|_{L^\infty(0,T)} \leq \|u_0 - v_0\|_{L^\infty} + \frac{3}{4}L^{\frac{4}{3}}\varepsilon^{\frac{1}{3}} + \lambda T,$$

where L is the Lipschitz constant of u.

Proof. We may assume that $\|u-v\|_{L^\infty(0,T)} = \sup(u-v)$. The other case can be treated similarly.

We have $u(t,x) - v(t,y) - \frac{|x-y|^4}{4\varepsilon} - \lambda t \ \leq \sup_{x,y}\{u_0(x) - v_0(y) - \frac{|x-y|^4}{4\varepsilon}\}$. If we take $x = y$, we have

$$\begin{aligned} \|u-v\|_{L^\infty(0,T)} &\leq \lambda T + \sup_{x,y}\left\{u_0(x) - v_0(y) - \frac{|x-y|^4}{4\varepsilon}\right\} \\ &\leq \lambda T + \|u_0 - v_0\|_{L^\infty} + \sup_{x,y}\left\{L|x-y| - \frac{|x-y|^4}{4\varepsilon}\right\}, \end{aligned} \tag{12}$$

and since the function

$$h(s) = Ls - \frac{s^4}{4\varepsilon}$$

attains its maximum at the point then $s = (L\varepsilon)^{\frac{1}{3}}$. Substituting this value in (12) the proof is concluded. □

Proof of the uniqueness. If we take two viscosity solutions for problem (4), by using Lemma 2.5,

$$\|u-v\|_{L^\infty(0,T)} \leq \|u_0 - v_0\|_{L^\infty} + \frac{3}{4}L^{\frac{4}{3}}\varepsilon^{\frac{1}{3}} + \lambda T$$

taking $\varepsilon^{\frac{1}{3}} = \delta\|u-v\|_{L^\infty(0,T)}$, where $\delta = L^{-\frac{4}{3}}$, we obtain

$$\|u-v\|_{L^\infty(0,T)} \leq \|u_0 - v_0\|_{L^\infty} + \frac{3}{4}\|u-v\|_{L^\infty(0,T)} + \lambda T$$

and thus

$$\|u-v\|_{L^\infty(0,T)} \leq 4\|u_0 - v_0\|_{L^\infty} + 4\lambda T.$$

Let t_1 be defined by the relation $4Mt_1 = \frac{1}{2}$, where M is the constant given by Lemma 2.4 for a fixed δ. Then, taking $\lambda = M\|u-v\|_{L^\infty(0,T)}$ we obtain

$$\frac{1}{2}\|u-v\|_{L^\infty(0,t_1)} \le 4\|u_0-v_0\|_{L^\infty}$$
$$\Rightarrow \|u-v\|_{L^\infty(0,t_1)} \le 8\|u_0-v_0\|_{L^\infty}.$$

Now let us take an interval $(t_1,\ 2t_1)$. Applying the same argument we obtain

$$\|u-v\|_{L^\infty(t_1,t_2)} \le 8\|u_1-v_1\|_{L^\infty(0,t_1)} \le 8\cdot 8\|u_0-v_0\|_{L^\infty},$$

where u_1, v_1 are the values taken in t_1.

Next we obtain

$$\|u-v\|_{L^\infty(0,T)} \le 8^n\|u_0-v_0\|_{L^\infty},$$

where n is such that

$$n\cdot t_1 \ge T,$$

and therefore we have

$$\|u-v\|_{L^\infty(0,T)} \le C_T\|u_0-v_0\|_{L^\infty}.$$

If both solutions have the same initial datum u_0, they coincide. This concludes the uniqueness statement of the theorem.

For the proof of the existence of solutions, we shall use the following approach.

Let us consider a family of periodic functions $u_0^\varepsilon \in C^\infty(\mathbb{R}^N)$ such that

$$u_0^\varepsilon \longrightarrow u_0 \text{ uniformly in } \mathbb{R}^N \text{ as } \varepsilon \to 0.$$
$$\|\nabla u_0^\varepsilon\|_{L^\infty} \le \|\nabla u_0\|_{L^\infty};\ \|u_0^\varepsilon\|_{L^\infty} \le \|u_0\|_{L^\infty}.$$

Let us introduce $g_\varepsilon = g + \varepsilon$; $a_{ij}^\varepsilon = \varepsilon\delta_{ij} + \alpha_{ij}^\varepsilon$, where

(i) $\alpha_{ij}^\varepsilon \to a_{ij}$ uniformly on compact subsets of $R^n \setminus \{0\}$;
(ii) $\sup_{i,j,\varepsilon}\sup_{R^n\setminus\{0\}} |\alpha_{ij}^\varepsilon| < \infty$ (uniform boundedness!);
(iii) $\sum_{i,j} \alpha_{ij}^\varepsilon(p)\xi_i\xi_j \ge 0\ \ \forall p \in \mathbb{R}^N \setminus \{0\},\ \xi \in \mathbb{R}^N$;
(iv) α_{ij}^ε has compact support on $\mathbb{R}^N \setminus \{0\}$.

We need the following lemma.

LEMMA 2.6. *Let $A = (a_{ij})$, $U = (u_{ij})$ be symmetric matrices with $A \ge 0$; then*

$$\Big|\sum_{i,j} a_{ij}u_{ij}\Big| \le C\Big(\sum_{i,j} a_{ij}\sum_k u_{ki}u_{kj}\Big)^{\frac{1}{2}}.$$

Proof. As

$$\sum_{i,j} a_{ij}u_{ij} = \operatorname{tr} AU \qquad \text{and} \qquad \sum_{i,j} a_{ij}\sum_k u_{ki}u_{kj} = \operatorname{tr} AU^2,$$

to prove the lemma, one chooses B so that it diagonalizes U (and not A); i.e.,

$${}^tBB = Id, \quad U = {}^tBDB.$$

Then, writing $(c_{ij}) = BA^tB$ and $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$, one computes

$$\begin{aligned} |\operatorname{tr} AU|^2 &= |\operatorname{tr}(A^tBDB)|^2 = |\operatorname{tr}(BA^tB)D|^2 \\ &= \left(\sum_i c_{ii}\lambda_i\right)^2 \le \sum_i c_{ii} \sum_i c_{ii}\lambda_i^2 \quad \text{(by the Schwarz inequality)} \\ &= \operatorname{tr} BA^tB \cdot \operatorname{tr} BA^tBD^2 = \operatorname{tr} A \cdot \operatorname{tr} AU^2. \end{aligned}$$

Then the lemma is proved by taking $C = (\operatorname{tr} A)^{\frac{1}{2}}$. □

LEMMA 2.7 (gradient estimate). *Let u be a smooth solution of*

$$\frac{\partial u}{\partial t} - g(\omega * \nabla G) \sum_{i,j} a_{ij}(\nabla u)u_{ij} - f(u) = 0 \text{ in }]0, +\infty[\times \mathbb{R}^N, \tag{13}$$

where a_{ij} are smooth on $\mathbb{R}^N$ and $\omega \in L^\infty(]0, +\infty[\times\mathbb{R}^N)$. Then

$$\|\nabla u(t, \cdot)\|_{L^\infty(\mathbb{R}^N)} \le e^{Ct}\|\nabla u_0\|_{L^\infty(\mathbb{R}^N)}, \tag{14}$$

where C depends on $\sup|a_{ij}|$, $\sup_s |f'(s)|$, *and* $\sup_{|p|<R} |D^2g(p)|$, *with* $R = \|\omega\|_{L^\infty(\mathbb{R}^N)} \times \|\nabla G\|_{L^1(\mathbb{R}^N)}$.

Proof. To prove the inequality we use the classical Bernstein method [14]. To this end, we differentiate (13) with respect to x_k, and we find

$$\begin{aligned} \frac{\partial u_k}{\partial t} &= \partial_k(g(\omega * \nabla G)) \sum_{i,j} a_{ij}(\nabla u)u_{ij} \\ &+ g(\omega * \nabla G)\left[\sum_l \left(\sum_{i,j} \frac{\partial a_{ij}}{\partial x_l}\right) u_{lk}u_{ij} + \sum_{i,j} a_{ij}(\nabla u)u_{ijk}\right] + f'(u)u_k. \end{aligned}$$

Hence, we obtain by multiplying by $2u_k$

$$\begin{aligned} 2u_k\frac{\partial u_k}{\partial t} &= 2(\partial_k(g(\omega * \nabla G))) \sum_{i,j} a_{ij}(\nabla u)u_{ij}u_k \\ &+ 2g(\omega * \nabla G)\left[\sum_l \left(\sum_{i,j} \frac{\partial a_{ij}}{\partial x_l}\right) u_{lk}u_{ij} + \sum_{i,j} a_{ij}(\nabla u)u_{ijk}\right] u_k + 2f'(u)u_k^2 \end{aligned}$$

since

$$\begin{aligned} \partial_{ij}(u_k^2) &= \partial_i(2u_{kj}u_k) = 2u_{ijk}u_k + 2u_{ik}u_{jk}, \\ \sum_k u_{ikj}u_k &= \partial_{ij}|\nabla u|^2 - 2\sum_k u_{ik}u_{jk}, \\ 2\sum_l u_{kl}u_l &= \sum_l \partial_k(u_l)^2 = \partial_k|\nabla u|^2. \end{aligned}$$

Summing over k, we get

$$\begin{aligned} \frac{\partial}{\partial t}|\nabla u|^2 &= 2\sum_k \partial_k(2g(\omega * \nabla G)) \sum_{i,j} a_{ij}(\nabla u)u_{ij}u_k \\ &+ g(\omega * \nabla G) \sum_k \sum_{i,j} \frac{\partial a_{ij}}{\partial x_k}\partial_k|\nabla u|^2 u_{ij} + g(\omega * \nabla G) \sum_{i,j} a_{ij}(\nabla u)\partial_{ij}|\nabla u|^2 \\ &- 4g(\omega * \nabla G) \sum_k \sum_{i,j} a_{ij}(\nabla u)u_{ik}u_{jk} + 2f'(u)|\nabla u|^2. \end{aligned}$$

If we introduce

$$\ell(v) = g(\omega * \nabla G) \sum_{i,j} a_{ij}(\nabla u)\partial_{ij} v + g(\omega * \nabla G) \sum_k \sum_{i,j} \frac{\partial a_{ij}}{\partial x_k} \partial_k u_{ij} v$$

we can write

$$\left(\frac{\partial}{\partial t} - \ell\right)|\nabla u|^2 = 2 \sum_k \frac{\partial g(\omega * \nabla G)}{\partial x_l} \sum_{i,j} a_{ij}(\nabla u) u_{ij} u_k$$
$$- 4g(\omega * \nabla G) \sum_k \sum_{i,j} a_{ij}(\nabla u) u_{ik} u_{jk} + 2f'(u)|\nabla u|^2$$

as

$$|\omega * \partial_{lk} G| \leq C,$$
$$(15) \qquad \left|\frac{\partial g}{\partial x_k}(\omega * \nabla G)\right| \leq C(g(\omega * \nabla G))^{\frac{1}{2}}, \qquad (g(x))^{\frac{1}{2}} \in W^{1,\infty}(\mathbb{R}^N),$$
$$\left|u_k \frac{\partial g}{\partial x_k}\right| \leq C|\nabla u||g(x)|^{\frac{1}{2}},$$

where C depends only on $\sup|\omega|$ and g; then we have

$$\left(\frac{\partial}{\partial t} - \ell\right)|\nabla u|^2 \leq -4g(\omega * \nabla G) \sum_k \sum_{i,j} a_{ij}(\nabla u) u_{ik} u_{jk}$$
$$+ 2C \sum_k (g(\omega * \nabla G))^{\frac{1}{2}} \sum_{i,j} a_{ij}(\nabla u) u_{ij} u_k + 2f'(u)|\nabla u|^2.$$

Taking into account (15) and Lemma 2.6 we obtain

$$\left(\frac{\partial}{\partial t} - \ell\right)|\nabla u|^2 \leq -4g(\omega * \nabla G) \sum_k \sum_{i,j} a_{ij}(\nabla u) u_{ik} u_{jk}$$
$$+ 2C \sum_k (g(\omega * \nabla G))^{\frac{1}{2}} \left(\sum_{i,j} a_{ij}(\nabla u) u_{ki} u_{kj}\right)^{\frac{1}{2}} u_k + 2f'(u)|\nabla u|^2$$
$$\leq -4g(\omega * \nabla G) \sum_k \sum_{i,j} a_{ij}(\nabla u) u_{ik} u_{jk} + g(\omega * \nabla G) \sum_k \sum_{i,j} a_{ij}(\nabla u) u_{ik} u_{jk}$$
$$+ C^2|\nabla u|^2 + 2f'(u)|\nabla u|^2.$$

This last inequality is due to $2ab \leq a^2 + b^2$.

Due to $\sum_k \sum_{i,j} a_{ij}(\nabla u) u_{ik} u_{jk} \geq 0$, and taking into account that $f'(u)$ is bounded, we have

$$(16) \qquad \left(\frac{\partial}{\partial t} - \ell\right)|\nabla u|^2 \leq C|\nabla u|^2.$$

Substituting v for $|\nabla u|^2$ in (16) we can write

$$\left(\frac{\partial}{\partial t} - \ell\right)v \leq Cv.$$

By putting $\tilde{v} = e^{-Ct}v$, we obtain $(\frac{\partial}{\partial t} - \ell)\tilde{v} \le 0$, and recalling that ℓ is an elliptic operator, by applying the maximum principle, we deduce (14). □

LEMMA 2.8. *Let u^ε be a smooth solution of*

$$\frac{\partial u^\varepsilon}{\partial t} - g_\varepsilon(u^\varepsilon * \nabla G)\sum_{i,j} a_{ij}^\varepsilon(\nabla u^\varepsilon)\partial_{ij}u^\varepsilon + f(u^\varepsilon) = 0.$$

Then for all $\delta > 0$ there exists $u_\delta^\varepsilon \in W^{2,\infty}(\mathbb{R}^N)$ such that for $0 < s \le t \le T$ we have

(1) $\mathrm{Sup}_{x\in\mathbb{R}^N} |u^\varepsilon(s,x) - u_\delta^\varepsilon(x)| \le C_T\delta$,

(2) $\mathrm{Sup}_{x\in\mathbb{R}^N} |D^2 u_\delta^\varepsilon(x)| \le \frac{C_T}{\delta}$,

(3) $\|u^\varepsilon(t,\cdot) - u_\delta^\varepsilon(\cdot)\|_{L^\infty} \le (M + \frac{C_T}{\delta})(t-s) + C_T\delta$,

(4) $u^\varepsilon(t,x)$ *is equicontinuous with respect to x and t,*

where C_T is a constant which does not depend on δ.

Proof. To prove (1) and (2), we use the mollifier, proposed to us by Paiva [17], to define $u_\delta^\varepsilon = J_\delta u^\varepsilon(s,\cdot)$, where J_δ is

$$(J_\delta u^\varepsilon)(x) = \frac{1}{\delta^N}\int_{|x-y|<\delta} \rho\left(\frac{|x-y|}{\delta}\right)u^\varepsilon(s,y)dy,$$

where ρ is a C^∞ function which verifies the following properties:

$$\rho(x) \ge 0 \qquad \text{if } |x| \le 1,$$
$$\rho(x) = 0 \qquad \text{if } |x| > 1,$$
$$\int_{|x|\le 1} \rho(x)dx = 1.$$

We easily obtain

$$D^2 u_\delta^\varepsilon(x) = \frac{1}{\delta^N}\int_{|x-y|<\delta} D\rho\left(\frac{|x-y|}{\delta}\right)Du^\varepsilon(t,y)dy$$

and thus a constant $C_T \ge \sup|Du^\varepsilon| \int_{|x|<1} |D\rho(x)|$ exists such that

$$\mathrm{Sup}_{x\in\mathbb{R}^N} |D^2 u_\delta^\varepsilon(x)| \le \frac{C_T}{\delta}.$$

Besides, we can write

$$|u^\varepsilon(s,x) - u_\delta^\varepsilon(x)| \le \frac{1}{\delta^N}\int_{|x-y|<\delta} \rho\left(\frac{|x-y|}{\delta}\right)|u^\varepsilon(s,y) - u^\varepsilon(s,x)|dy$$

and since in Lemma 2.7

$$|u^\varepsilon(s,y) - u^\varepsilon(s,x)| \le C_T|x-y|$$

we obtain

$$|u^\varepsilon(s,x) - u_\delta^\varepsilon(x)| < C_T\delta.$$

To prove that $u^\varepsilon(t,x)$ is equicontinuous it is enough to prove

$$|u^\varepsilon(t,y) - u^\varepsilon(t,x)| \le C_T|x-y|,$$
$$|u^\varepsilon(s,x) - u^\varepsilon(t,x)| \le M|t-s| + 2C_T|t-s|^{\frac{1}{2}}.$$

The first inequality has already been proved, and only the second inequality remains to be proved, for which we do

$$v(x,t) = u^\varepsilon(t,x) - u^\varepsilon_\delta(x).$$

Since the derivative in t is $v_t = u^\varepsilon_t$ we can write

$$v_t - g_\varepsilon(u^\varepsilon * \nabla G)\sum_{i,j} a_{ij}(\nabla u^\varepsilon)\partial_{ij}v = f(u^\varepsilon_\delta) + g_\varepsilon(u^\varepsilon * \nabla G)\sum_{i,j} a_{ij}(\nabla u^\varepsilon)\partial_{ij}u^\varepsilon_\delta$$

$$\Rightarrow v_t - g_\varepsilon(u^\varepsilon * \nabla G)\sum_{i,j} a_{ij}(\nabla u^\varepsilon)\partial_{ij}v \leq M + \frac{C_T}{\delta} = M'_T.$$

If we now denote

$$\ell_1(v) = g_\varepsilon(u^\varepsilon * \nabla G)\sum_{i,j} a_{ij}(\nabla u^\varepsilon)\partial_{ij}v$$

and make the substitution

$$\tilde{v}^1(x,t) = v(x,t) - M'_T(t-s)$$

then $\tilde{v}^1_t - \ell_1(\tilde{v}^1) \leq 0$, and, since ℓ_1 is an elliptic operator, applying the maximum principle we obtain that $\sup_{\mathbb{R}^N\times(s,t)} \tilde{v}^1(x,t)$ is attained at $t = s \Longrightarrow v(x,t) \leq M'_T(t-s) + \sup_x v(x,s) \leq M'_T(t-s) + C_T\delta$.

We write

$$\tilde{v}^2(x,t) = -v(x,t) - M'_T(t-s).$$

We have that $\tilde{v}^2_t - \ell_1(\tilde{v}^2) \leq 0$, and for the same reason we obtain that $-v(x,t) \leq M'_T(t-s) + \sup_x v(x,s) \leq M'_T(t-s) + C_T\delta$.

In fact, it has been seen that $|v(x,t)| < M'_T(t-s) + C_T\delta$. Then we have that

$$|u^\varepsilon(s,x) - u^\varepsilon(t,x)| \leq |u^\varepsilon(t,x) - u^\varepsilon_\delta(x)| + |u^\varepsilon(s,x) - u^\varepsilon_\delta(x)|$$

$$\leq |u^\varepsilon(t,x) - u^\varepsilon_\delta(x)| + \|u^\varepsilon(s,\cdot) - u^\varepsilon_\delta\|_{L^\infty} \leq \left(M + \frac{C_t}{\delta}\right)(t-s) + C_T\delta + C_T\delta.$$

Substituting $\delta = |t-s|^{\frac{1}{2}}$ we have

$$|u^\varepsilon(s,x) - u^\varepsilon(t,x)| \leq M(t-s) + 2C_T|t-s|^{\frac{1}{2}}$$

and we deduce easily that $u^\varepsilon(t,x)$ is equicontinuous. □

Proof of existence of solutions. Using the general theory of quasilinear uniformly parabolic equations [5], [11], [14], one checks easily that there exists a function u^ε which is smooth on $[0,+\infty[\times\mathbb{R}^N$ and a solution of

$$\frac{\partial u^\varepsilon}{\partial t} - g_\varepsilon(u^\varepsilon * \nabla G)\sum_{i,j} a^\varepsilon_{ij}(\nabla u^\varepsilon)\partial_{ij}u^\varepsilon + f(u^\varepsilon) = 0 \quad \text{in }]0,+\infty[\times\mathbb{R}^N.$$

Taking into account the consistency and stability properties of viscosity solutions, it is enough to prove that u^ε (or a subsequence of u^ε) converges uniformly in $[0,T]\times\mathbb{R}^N$ to a function $u \in C([0,T]\times\mathbb{R}^N) \cap L^\infty(0,T;W^{1,\infty}(\mathbb{R}^N))$ for all $T < \infty$.

Since the functions u^ε are equicontinuous, by using the Ascoli–Arzola theorem, we conclude the proof.

3. Numerical analysis.

3.1. Optimal quantizer. In order to compute our model the first problem to solve is the choice of the quantizer Q_S. Usually, the Lloyd functional (1) has been used to choose such a quantizer. However, in applications to image processing it is desirable that the quantizer levels are not too close each other. To avoid this, we add two terms in the Lloyd functional, so we work with the functional (2)

$$E(Q_S) = \sum_{k=1}^{S}\left(\int_{t_k}^{t_{k+1}} (s-u_k)^2 dH(s) + \frac{L}{t_{k+1}-t_k} + C\left(\frac{t_k+t_{k+1}}{2} - u_k\right)^2\right),$$

where L and C are positive constants. We notice that when L and C tend to infinity, the minimum of the above functional tends to the uniformly distributed quantizer; i.e., the distance between the quantizer levels is constant. Moreover, using functional (2) we remove some nonuniqueness problems that can appear when the Lloyd functional is used. Indeed, in the case of the histogram of the image being 0 in a region, then the Lloyd functional can have several global minima. But in the case of the functional (2) this nonuniqueness problem is removed because it tends to distribute uniformly the quantizer levels.

The first way we tried to compute a global minimum of functional (2) was the descent gradient procedure. We have seen that this procedure does not work because it depends strongly on the initial condition that we choose. The problem is due to the fact that functional (2), as in the case of Lloyd functional, has in general a lot of local minima that can be far away from the global minimum. To solve this problem we use an exhaustive search among all possible choices of quantizer Q_S. We will assume that the codewords $\{u_k\}$ are integer numbers in the interval $[0, 255]$, which are the usual bounds of the gray level images. To avoid the codewords $\{u_k\}$ and the separators $\{t_k\}$ having the same values, we will also assume that the separators $\{t_k\}$ are float numbers in the set $\{-0.5, 0.5, 1.5, 2.5, \dots, 255.5\}$. In general, the exhaustive search of the global minimum of functional (2) has a complexity of $\mathcal{O}(K^{S+1})$, where K represents the dimension of the search space (in our case 257) and S represents the number of quantizer levels. However, using dynamic programming techniques we can lower the complexity of the search to $\mathcal{O}(K^2 S)$.

The dynamic programming techniques have already been used to compute the global minimum of the Lloyd functional (see, for instance, [6] or more recently [21], where the authors show that in some cases the complexity of the search can be reduced to $\mathcal{O}(KS)$). We can apply dynamic programming techniques in the case of the energy functional being separable in the following sense: if $\hat{Q}_S$ is a global minimum of the energy functional in $[t_1, t_{S+1}]$ with $t_1 < u_1 < t_2 < u_2 < \cdots < u_S < t_{S+1}$, then the quantizer $\hat{Q}_{S-1}$ defined by $t_2 < u_2 < \cdots < u_S < t_{S+1}$ is a global minimum of the energy in the interval $[t_2, t_{S+1}]$. Of course, the functional (2) satisfies this property. So we compute the global minimum of the functional (2) in two steps. In the first step we compute the energy and the associated codeword $u_{t,t'}$ for each interval $[t, t']$ included in $[-0.5, 255.5]$. We notice that the associated codeword $u_{t,t'}$ can be computed using the derivative of the functional (2) with respect to u_k and we obtain

$$u_{t,t'} = \frac{\int_t^{t'} sh(s)ds + C\frac{t+t'}{2}}{\int_t^{t'} h(s)ds + C}.$$

In the second step, we compute the global minimum $\hat{Q}_S$ of the functional (2) recursively by computing $\hat{Q}_2^{t'}, \hat{Q}_3^{t'}, \dots$, taking into account that $\hat{Q}_k^{t'} = \{0, u_{0,t'}\} \cup \hat{Q}_{k-1}^{t'}$, where $\hat{Q}_k^{t'}$ represents the global minimum of the functional (2) in the interval $[t', 255.5]$ using k quantization levels.

Of course, in practice, the integrals have to be replaced by additions. Moreover, we can lower the number of operations to compute the energy, taking into account that

$$\int_t^{t'} (s-u_{t,t'})^2 h(s)ds = \int_t^{t'} s^2 h(s)ds - u_{t,t'} \int_t^{t'} 2sh(s)ds + u_{t,t'}^2 \int_t^{t'} h(s)ds.$$

So, we can compute very rapidly this energy and the codewords $u_{t,t'}$ if we have computed previously the integrals $\int_0^{t'} h(s)ds$, $\int_0^{t'} sh(s)ds$, and $\int_0^{t'} s^2 h(s)ds$ for any $t' = -0.5, 0.5, 1.5, \dots, 255.5$.

Once the optimal quantizer $\hat{Q}_S$ is computed, we can address the problem of solving the PDE (3) which generates our quantization model. The asymptotic state of the solution of the differential equation represents a smooth quantization of the original image using the optimal quantizer $\hat{Q}_S$.

As we have explained in the introduction, we associate with each quantizer Q_S a function $f(.)$ satisfying for any k, $f(t_k) = f(u_k) = 0$, $\frac{\partial f}{\partial s}(t_k) > 0$, and $\frac{\partial f}{\partial s}(u_k) < 0$. We fix the values $t_1 = -0.5$ and $t_{S+1} = 255.5$. The values of the function $f(.)$ outside the interval $[-0.5, 255.5]$ are not really important because the image distribution will be in general in the interval $[0, 255]$. In the interval $[-0.5, 255.5]$ we define $f(.)$ in the following way:

$$f(s) = -\frac{1}{u_k - t_k}(s-t_k)(s-u_k) \quad \text{if} \quad s \in [t_k, u_k),$$

$$f(s) = \frac{1}{t_{k+1} - u_k}(s-t_{k+1})(s-u_k) \quad \text{if} \quad s \in [u_k, t_{k+1}).$$

This choice of function $f(.)$ fits the value of the derivative of f in the level quantizers to -1 in the case of u_k and to 1 in t_k. Outside the interval $[-0.5, 255.5]$ we define the function f following some stability criteria for equation (3) developed in the next sections.

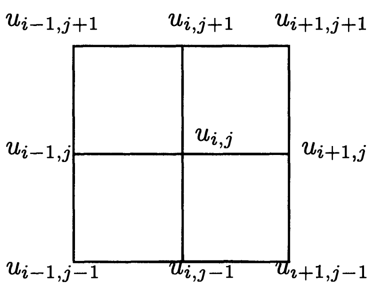
3.2. Differential operator discretization in $\mathbb{R}^2$. Denoting by ξ the direction orthogonal to ∇u, we have

$$u_{\xi\xi} = \|\nabla u\| \operatorname{div}\left(\frac{\nabla u}{\|\nabla u\|}\right).$$

Let θ be such that $\xi = (-\sin\theta, \cos\theta)$. We can write

$$u_{\xi\xi} = \ell(u) = \sin^2\theta u_{xx} - 2\sin\theta\cos\theta u_{xy} + \cos^2\theta u_{yy}. \tag{17}$$

We are going to write $u_{\xi\xi}$ as a linear combination of values of u on a 3×3 fixed matrix.



$$\ell(u) = -\frac{4\lambda_0}{h^2}(u_{i,j}) + \lambda_1(u_{i+1,j} + u_{i-1,j}) + \lambda_2(u_{i,j+1} + u_{i,j-1}) + \lambda_3(u_{i+1,j+1} + u_{i-1,j-1}) + \lambda_4(u_{i+1,j-1} + u_{i-1,j+1}), \tag{18}$$

where h is the increment value which we consider constant and equal to both variables.

The discretization by this method depends on a free parameter λ_0 which we shall choose by a geometrical criterion.

By applying Taylor's formula and equating (17) with (18) we obtain the following values for λ_i:

$$\lambda_1 = \frac{2\lambda_0 - \sin^2\theta}{h^2},$$
$$\lambda_2 = \frac{2\lambda_0 - \cos^2\theta}{h^2},$$
$$\lambda_3 = \frac{-\lambda_0 + \frac{1}{2}(1 + \sin\theta\cos\theta)}{h^2},$$
$$\lambda_4 = \frac{-\lambda_0 + \frac{1}{2}(1 - \sin\theta\cos\theta)}{h^2}.$$

Now, we search for a trigonometrical polynomial on θ for the value of $\lambda_0(\theta)$ in such a way that when there is diffusion in some principal directions, in the other directions the diffusion is zero.

If we put $\lambda_0(\theta) = a + b\cos\theta + c\sin\theta + d\cos^2\theta + e\sin\theta + f\sin\theta\cos\theta + g\sin^2\theta\cos^2\theta$, then, taking into account the values of λ_i already calculated and recalling that (17) is an approximation of (18),

$$-4\lambda_0 = -2 + 4\sin^2\theta\cos^2\theta,$$
$$\lambda_1 = \cos^2\theta(\sin^2\theta - \cos^2\theta),$$
$$\lambda_2 = \sin^2\theta(\sin^2\theta - \cos^2\theta),$$
$$\lambda_3 = \sin^2\theta\cos^2\theta + \frac{1}{2}\sin\theta\cos\theta,$$
$$\lambda_4 = \sin^2\theta\cos^2\theta - \frac{1}{2}\sin\theta\cos\theta$$

are the final values obtained.

3.3. Algorithm and stability. The algorithm is divided into the following steps:

(1) Computation u_x and u_y:

$$u_x(i) = \frac{1}{4h}\left(u_{i+1,j} - u_{i-1,j} + \frac{1}{2}((u_{i+1,j+1} - u_{i-1,j-1}) + (u_{i+1,j-1} - u_{i-1,j+1}))\right),$$
$$u_y(i) = \frac{1}{4h}\left(u_{i,j+1} - u_{i,j-1} + \frac{1}{2}((u_{i+1,j+1} - u_{i-1,j-1}) + (u_{i+1,j-1} - u_{i-1,j+1}))\right).$$

(2) Computation $\cos\theta$ and $\sin\theta$:

$$\cos\theta = \frac{u_x}{\sqrt{u_x^2 + u_y^2}}, \quad \sin\theta = \frac{u_y}{\sqrt{u_x^2 + u_y^2}}.$$

(3) Computation of the function g. First, we compute the convolution of the image with a Gaussian kernel with a given standard deviation σ, and then we compute the norm of the gradient using the algorithm presented in step 1. We use a threshold parameter to determine when the diffusion is lowered.

(4) Iterative scheme:

$$\frac{u_{i,j}^{n+1} - u_{i,j}^{n}}{\triangle\, t} = g(u_{i,j}^{n})\ell(u_{i,j}^{n}) + f(u_{i,j}^{n}).$$

To simplify the stability analysis of the algorithm, taking into account that equation (4) satisfies invariance by translations, we have taken the function $f(x)$ centered with respect to the origin.

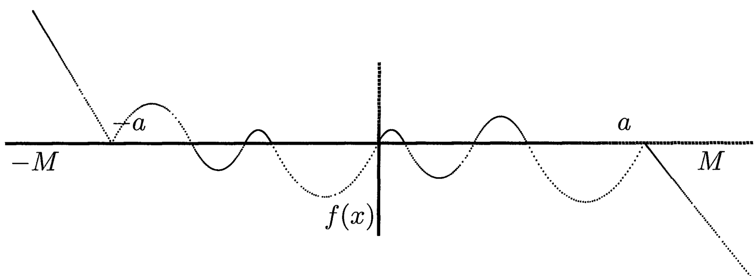
The numerical scheme stability is established in the following theorem.

THEOREM 3.1. *Let*

$$u_{i,j}^{n+1} = u_{i,j}^{n}(1 - 4g(u_{i,j}^{n})C_d\lambda_0 \triangle\, t) + g(u_{i,j}^{n})C_d \triangle\, t \sum_{k,l} \lambda_{k,j}\; u_{i+k,j+l}^{n} + \triangle\, t\; f(u_{i,j}^{n}),$$

where $f(x)$ satisfies

$$\begin{aligned} &|f(u)| \le C_r && \text{if } u \in [-a,a],\\ &f(u) = k(-a-u) && \text{if } u < -a,\\ &f(u) = k(a-u) && \text{if } u > a. \end{aligned}$$



If $0.65C_d - 2k < 0$ *is verified then we have that* $\forall M > a,\ \exists \delta > 0$ *in such a way that when* $0 < \triangle\, t < \delta$

$$\text{if } \forall i,j\ |u_{i,j}^{n}| < M \Longrightarrow |u_{i,j}^{n+1}| \le M.$$

Proof. We recall that $0 < g(s) \le 1$. Let us divide the problem into three cases.

Case 1. $u_{i,j}^{n} \in [-M,-a]$. We have $u_{i,j}^{n+1} = u_{i,j}^{n}(1 - 4g(u_{i,j}^{n})C_d\lambda_0 \triangle\, t - k \triangle\, t) + C_d g(u_{i,j}^{n}) \triangle\, t \sum_{k,l} \lambda_{k,j}\; u_{i+k,j+l}^{n} - \triangle\, t\, ka$.

If we impose that $(1 - 4g(u_{i,j}^{n})C_d\lambda_0 \triangle\, t - k \triangle\, t) \ge 0$, we obtain

(i) $u_{i,j}^{n+1} \le C_d g(u_{i,j}^{n}) \triangle\, t \sum_{k,l} |\lambda_{k,j}|\, M \le M \Longleftrightarrow C_d g(u_{i,j}^{n}) \triangle\, t \sum_{k,l} |\lambda_{k,j}| \le 1$.

(ii) $u_{i,j}^{n+1} \ge (1 - 4C_d g(u_{i,j}^{n})\lambda_0 \triangle\, t - k \triangle\, t)(-M) + C_d g(u_{i,j}^{n}) \triangle\, t \sum_{k,l} |\lambda_{k,j}|(-M) - \triangle\, t\, ka$
$\ge (1 - 4g(u_{i,j}^{n})C_d\lambda_0 \triangle\, t - k \triangle\, t)(-M) + C_d g(u_{i,j}^{n}) \triangle\, t \sum_{k,l} |\lambda_{k,j}|(-M) - \triangle\, t\, kM$
$\ge (1 - 4g(u_{i,j}^{n})C_d\lambda_0 \triangle\, t - 2k \triangle\, t + C_d g(u_{i,j}^{n}) \triangle\, t \sum_{k,l} |\lambda_{k,j}|)(-M) > -M$

whenever

$$\left|\left(1-\left(\left(4\lambda_0 \bigtriangleup t-\Delta t\sum_{k,l}|\lambda_{k,j}|\right)g(u^n_{i,j})g(u^n_{i,j})C_d+2k\right)\bigtriangleup t\right)\right|<1 \tag{19}$$

and we obtain a condition on $\bigtriangleup t$: $((4\lambda_0 \bigtriangleup t-\Delta t\sum_{k,l}|\lambda_{k,j}|)g(u^n_{i,j})C_d+2k)\bigtriangleup t<2$ must be verified.

By elementary calculations, one easily obtains

$$-0.65\le\left(4\lambda_0 \bigtriangleup t-\Delta t\sum_{k,l}|\lambda_{k,j}|\right)g(u^n_{i,j})C_d\le 0$$

and therefore in order to verify (19), it is sufficient to take $\bigtriangleup t<\frac{2}{k}$.

Case 2. $u^n_{i,j}\in[-a,a]$. We have $|u^{n+1}_{i,j}|\le(1-4g(u^n_{i,j})C_d\lambda_0 \bigtriangleup t)a+C_dg(u^n_{i,j})\Delta t\sum_{k,l}|\lambda_{k,j}|\,M+\bigtriangleup t\,C_r$. Let us suppose that this expression is $\le M$. In that case, we obtain $(1-4g(u^n_{i,j})C_d\lambda_0 \bigtriangleup t)a+\bigtriangleup t\ C_r\le M(1-C_dg(u^n_{i,j})\bigtriangleup t\sum_{k,l}|\lambda_{k,j}|)$ and dividing everything by $(1-C_dg(u^n_{i,j})\bigtriangleup t\sum_{k,l}|\lambda_{k,j}|)$ we have

$$\frac{(1-4g(u^n_{i,j})C_d\lambda_0 \bigtriangleup t)a+\bigtriangleup t\ C_r}{(1-C_dg(u^n_{i,j})\bigtriangleup t\sum_{k,l}|\lambda_{k,j}|)}\le M. \tag{20}$$

When $\bigtriangleup t\to 0$, the left-hand side of (20) therefore has a as limit, a being less than M; therefore, if Δt is small enough for (20) to be true, and so, $|u^{n+1}_{i,j}|\le M$.

Case 3. $u^n_{i,j}\in[a,M]$. It is proved in the same way as in Case 1. □

4. Experimental results. The numerical experiences have been computed in two steps. In the first one we compute the optimal quantizer $\hat{Q}_S$ with respect to the energy (2). In this step we have to choose three parameters:

(1) the number of levels S to determine the quantizer $\hat{Q}_S$,
(2) the constant L in the energy functional (2),
(3) the constant C in the energy functional (2).

In the second step we solve numerically the differential equation (3). In this step we must choose five parameters:

(1) The standard deviation σ of the Gaussian function G_σ to compute $G_\sigma * u$.
(2) The threshold parameter T_g to determine the shape of the function $g(.)$. It means that the diffusion is lowered for $|\nabla G_\sigma * u|>T_g$.
(3) To balance the influence of the diffusion and reaction terms, we use a constant C_f. So we use $C_f f(u)$ as reaction term, where f is defined above.
(4) The discretization step Δt.
(5) The number of iterations.

The first experience that we present in Figure 1 shows the difference between the Lloyd functional (1) and the functional (2) proposed in this paper. We take a synthetic picture which represents a square in a dark background (Figure 1a). The normalized histogram of the picture, $h(s)$, contains only three nonzero values, which are $h(3)=0.497864$, $h(6)=0.480103$, and $h(140)=0.022034$. If we compute the optimal quantizer $\tilde{Q}_2$ using the Lloyd functional with 2 codewords, we obtain that $t_1=-0.5$, $u_1=3$, $t_2=3.5$, $u_2=12$, and $t_3=255.5$. So the gray level 3 remains unchanged and the gray levels 6 and 141 go to the same level 12. Therefore, as we can see in Figure 1b, the square is lost. However, if we compute the optimal quantizer $\hat{Q}_2$ associated to the functional (2) with $L=20$, and $C=0.0025$, we obtain $t_1=-0.5$,

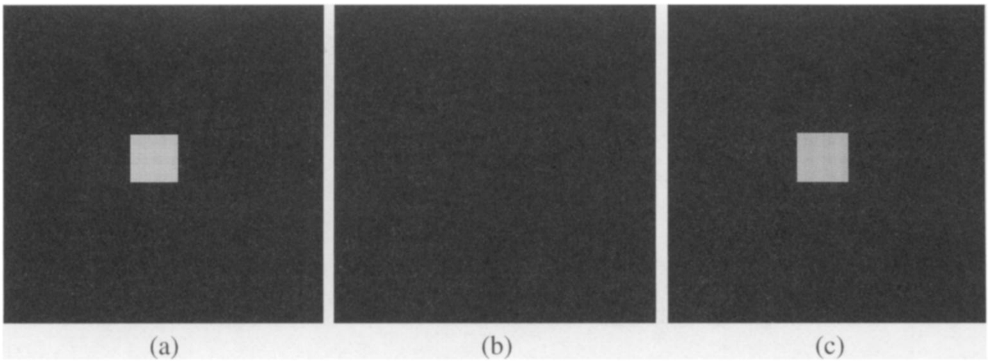


FIG. 1. *Figures follow from left to right*: (a) *original image,* (b) 2-*level quantization using Lloyd energy,* (c) 2-*level quantization using energy* (2).

$u_1 = 5$, $t_2 = 41.5$, $u_2 = 141$, and $t_3 = 255.5$. So the background is normalized to the gray-level value 5 and the square remains in the picture as we can see in Figure 1c.

In the second experience we present a medical image provided by the service of vascular interventional radiology of the hospital Nuestra Señora del Pino de Las Palmas de Gran Canaria (Figure 2). In the top, from left to right, we present (a) the

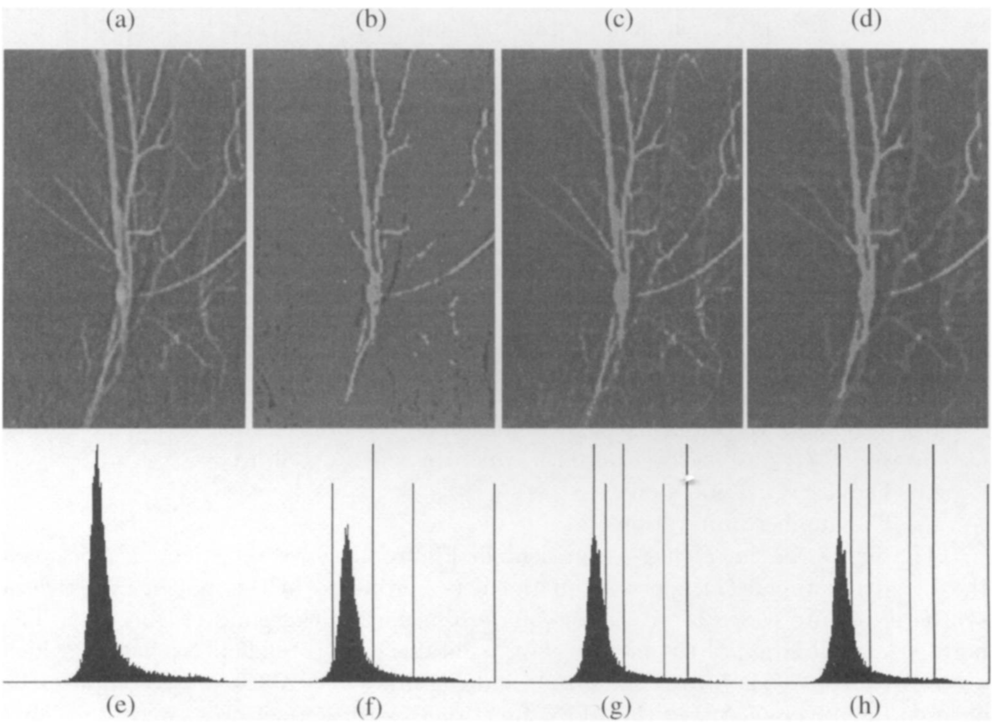


FIG. 2. *Figures follow from left to right and from top to bottom*: (a) *original image,* (b) 3-*level quantization using uniform distributed quantizer,* (c) 3-*level quantization using energy* (2), (d) 3-*level quantization using equation* (3), (e) *histogram of the original image,* (f) *location of uniform distributed quantizer,* (g) *location of the optimal quantizer using the energy* (2), (h) = (g).

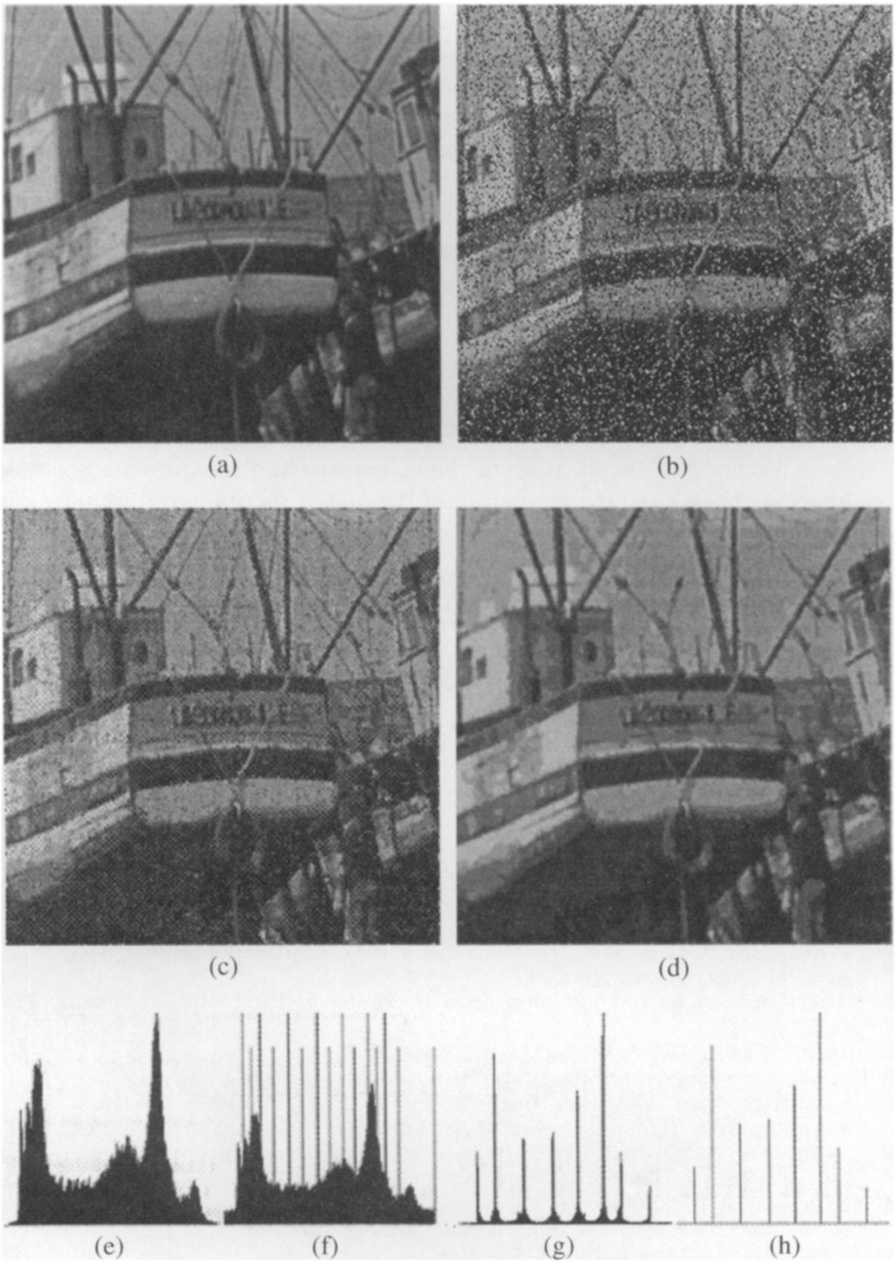


FIG. 3. *Figures follow from left to right and from top to bottom:* (a) *original image,* (b) *original image with noise,* (c) 8*-level quantization of the noised image using equation* (3) *with* 20 *iterations,* (d) *result after* 200 *iterations,* (e) *histogram of the original image,* (f) *location of optimal quantizer,* (g) *histogram of image* (c), (h) *histogram of picture* (d).

original image; (b) a 3-gray level quantized image using a uniform distribution of codewords and separators (without PDE); (c) a 3-gray level quantized image using as quantizer $\hat{Q}_3$ a global minimum of the functional (2) with $L = 1$ and $C = 0.0025$ (without PDE); (d) a 3-gray level quantized image using the above quantizer $\hat{Q}_3$ and the PDE (3) with $\Delta t = 0.2$, $C_f = 4.0$, $\sigma = 2$, $T_g = 25$, and 100 iterations.

At the bottom of the picture from left to right we present (e) the histogram of the original picture, (f) the quantizer levels which correspond to a uniform distribution, and (g) the quantizer levels which correspond to minimize the functional (2). (h) The quantizer $\hat{Q}_3$ obtained above is used to compute the PDE (3), so (g) = (h). In (f), the long vertical lines represent the location of the codewords and the short vertical lines the location of the separators.

In the third experiment, we present an application of our model to quantization and denoising. In Figure 3a we present the original image. In (b) we have introduced a noise in the picture. We take 25% of the picture in a random way and we change its value randomly. In (c) we compute the solution of the PDE (3) using as initial datum the noised picture with the following parameters. We choose an 8-gray level optimal quantizer $\hat{Q}_8$ with $L = 10$ and $C = 0.0025$. We solve numerically the PDE with $\Delta t = 0.1$, $C_f = 10.0$, $\sigma = 3$, $T_g = 45$, and 20 iterations. We present the result after 200 iterations in (d). In (e) we see a histogram of the original picture. The histogram of the noised picture and the location of the codewords and separators associated to $\hat{Q}_8$ are shown in (f). We see the histogram of the noised picture after 20 iterations of our algorithm after 200 iterations in (g) and (h).

Acknowledgments. The authors would like to thank the support of CEREMADE laboratory of the University Paris IX (Dauphine), where the authors stayed during the elaboration of part of this paper.

We would like to thank the anonymous reviewers for the valuable remarks; they contributed to an improvement of the paper.

REFERENCES

[1] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel, *Axioms and fundamental equation of image processing*, Arch. Rational Mech., 132 (1993), pp. 199–257.

[2] L. Alvarez, P. L. Lions, and J. M. Morel, *Image selective smoothing and edge detection by non linear diffusion.* II, SIAM J. Numer. Anal., 29 (1992), pp. 845–866.

[3] L. Alvarez and J. M. Morel, *Formalization and computational aspects of image analysis*, Acta Numerica, (1994), pp. 1–61.

[4] G. Barles, *Remark on a Flame Propagation Model*, Report 464, INRIA, Le Chesnay, France, 1985.

[5] H. Brezis, *Analyse Fonctionnelle*, Masson, Paris, 1987.

[6] J. D. Bruce, *Optime Quantization*, Sc.D. thesis, M.I.T., Cambridge, MA, 1964.

[7] G. H. Cottet and L. Germain, *Image Processing Through Reaction Combined with Nonlinear Diffusion*, Technical Report RT 78, University of Grenoble, France, 1992.

[8] Y. G. Chen, Y. Giga, and S. Goto, *Uniqueness and existence of generalized mean curvature flow equations*, J. Differential Geom., 33 (1991), pp. 749–786.

[9] M. C. Crandall, H. Ishii, and P. L. Lions, *User's guide to viscosity solution of second order partial differential equation*, Bull. Amer. Math. Soc., 27 (1992), pp. 1–67.

[10] M. C. Crandall and P. L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277 (1983), pp. 1–42.

[11] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, NJ, 1964.

[12] A. K. Jan, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[13] J. J. Koenderink, *The structure of images*, Biol. Cybernet., 50 (1984), pp. 363–370.

[14] O. A. Ladyzhenskaja, V. A. Solonnikov, and N. N. Ural'tseva, *Linear and Quasilinear Equations of Parabolic Type*, American Mathematical Society, Providence, RI, 1968.

[15] S. P. Lloyd, *Least squares quantization in PCM*, IEEE Trans. Inform. Theory, IT-28 (1982), pp. 129–137.

[16] D. Marr and H. Hildreth, *Theory of edge detection*, Proc. Roy. Soc. London Ser. B., 207 (1980), pp. 187–210.

[17] F. Paiva, *Personal communication*, 1994.

[18] P. PERONA AND J. MALIK, *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Machine Intell., 12 (1990), pp. 429–439.
[19] M. SONER, *Motion of a Set by the Curvature of Its Mean Boundary*, Research Report N: 90-82-NAMS2, Department of Mathematics, Carnegie Mellon University, 1991, preprint.
[20] A. P. WITKIN, *Scale-space filtering*, Prov. IJCAI, Karlsruhe, (1983), pp. 1019–1021.
[21] X. WU AND K. ZHANG, *Quantizer monotonicities and globally optimal scalar quantizer design*, IEEE Trans. Inform. Theory, 39 (1993), pp. 1049–1053.