

---

# Ultrasound Image Filtering by Anisotropic Diffusion with Numerical Simulation

---

*Clovis Tauber*<sup>1,\*</sup> and *Pierre Spiteri*<sup>2,†</sup>

<sup>1</sup>UMRS INSERM U930 - CNRS ERL3106,  
Université François Rabelais, Tours, France

<sup>2</sup>IRIT - ENSEEIHT, UMR CNRS 5505, Toulouse, France

**Keywords:** Anisotropic diffusion, nonlinear filtering, speckle, ultrasound imaging, iterative method, boundary value problem, large scale system, relaxation method, conjugate gradient method.

## 1. Introduction

Ultrasound imaging is a widely used method for medical applications. It is non-invasive, real-time and relatively inexpensive. However, compared to other medical imaging methods, it suffers from low signal to noise ratio, the presence of speckle, signal dropouts, and a low contrast. The speckle appears as a multiplicative, locally correlated noise caused by destructive interference of signals reflected from scatterers within one resolution cell. This phenomenon degrades the resolution of images and makes difficult the detection of their features. Hence, the development of effective noise reduction filters for echographic

---

\*E-mail address: clovis.tauber@univ-tours.fr

†E-mail address: pierre.spiteri@enseeiht.fr

images is of primary importance in order to i) enhance the image; and ii) facilitate higher level processing such as image segmentation, pattern recognition and motion tracking.

Early speckle reducing filters originated in the domain of synthetic aperture radar (SAR). Two main models of speckle filters have been developed, namely the minimum mean-square error (MMSE) filters which include those of Lee [20], Kuan et al. [18], and Frost et al. [12], and the maximum a posteriori Bayesian Gaussian filters [19, 22]. These filters use the coefficient of variation to differentiate homogeneous regions from boundaries. They effectively smooth homogeneous regions, but tend to leave sharp regions and the surroundings of edges unfiltered due to hard thresholding.

In ultrasound images, the speckle has long been treated as an additive noise. Taxt [34, 35] applied homomorphic filtering to transform the speckle into an additive noise [26]. Czerwinski et al. used sticks to enhance boundary detection in ultrasound images [9]. Their technique operates by applying a set of patterns (i.e. sticks) as a filter bank, and retaining the statistically largest output at each point. Perona et al. [27] introduced anisotropic diffusion to filter not only homogeneous regions but also edge neighborhoods. This technique has been widely studied for additive noise filtering [5, 6, 37]. An inherent limitation of these diffusion models is their tendency to produce a totally flat image after a high number of iterations. Later, Montagnat proposed a 4D anisotropic diffusion for ultrasound images [25]. And more recently, Krissian et al. proposed a speckle constrained anisotropic diffusion for ultrasound images [17] that leads to good speckle filtering while also being based on the intensity gradient.

Yu et al. [41, 42] were the first to replace the classical intensity gradient with the local coefficient of variation in the coefficient of diffusion, to stop the diffusion near the image contours. They established a parallel between the MMSE filters of Lee [20] and Kuan et al. [18] and their proposed filter. Their method leads to interesting results as the coefficient of variation allows a precise localization of the edges in images affected by speckle. However, their method has some limitations regarding the conservation of energy within image regions.

You et al. [40], then Black et al. [4] proposed robust versions of the anisotropic diffusion by using coefficients of diffusion based on M-estimators to effectively stop the diffusion above a predefined threshold. The consequence of this modification is the stopping of the diffusion process after a finite number of iterations, in contrast with the Perona et al. [27] or Yu et al. [41] models which reduce the diffusion process with time but do not effectively stop it. Black et al. [4] also proposed a robust estimator of the scale parameter of the anisotropic diffusion for additive noises. Their almost automatic method brought significant theoretical advances and lead to good results on optical images. However, their methods and robust estimator are not adapted to filter images affected by speckle. They lead to insufficient filtering of multiplicative noise.

To sum up, the methods proposed in the literature can significantly reduce the level of speckle, but they usually have one or several of the following limitations: i) insufficient smoothing of the homogeneous regions, ii) insufficient preservation of the semantically correct edges, iii) non preservation of the intra-regions energies, or iv) non automatic estimation of the scale parameter. The aim of the present chapter is then to give some answers to these described open problems.

In this chapter, we propose a robust anisotropic diffusion method for speckle filtering in

ultrasound images, called  $\gamma$ -diffusion, that addresses these limitations [32]. It combines the advantages of SAR filters, anisotropic diffusion and robust statistics. This chapter includes : i) an unbiased speckle filter that produces semantically correct scale spaces; preserves edges even after a high number of iterations; and conserves global and intra region energy; ii) a robust estimator to automatically compute the scale parameter of the coefficient of variation; and iii) a comparative study with eleven speckle filters from the litterature.

The proposed model is a nonlinear evolutive boundary value problem of diffusion with homogeneous Neumann boundary condition (see [1]). In the definition of this model, the coefficient of diffusion vanishes locally at some points of the domain  $\Omega$  where the partial differential equation is defined. This property implies the fact that the discretization matrix can be reducible. The solution of the evolutive boundary value problem can be achieved by various time marching schemes. The use of an explicit scheme can be problematic since such method suffers from stability limitations. Whereas, the use of a fully implicit time marching scheme is also problematic as it requires the use of the Newton method due to nonlinearity. In addition, the implementation of the Newton method is difficult in the considered model, since the computation of the derivative of the coefficient of diffusion is not easy. Furthermore, the convergence of the Newton method at each time step is not guaranteed in the general case and practical criteria of convergence are not obvious to obtain. This two facts constitute a strong limitation when an implicit time marching scheme is considered to solve the evolutive problem. Another difficulty consists in the fact that, in the general case, when an implicit or semi-implicit scheme is used, the linear systems to solve at each time step can be ill-conditioned, due to the Neumann boundary condition.

In order to solve numerically this anisotropic diffusion problem, and better take into account its nonlinear part, we propose in the second part of this chapter two time marching semi-implicit discretization schemes derived from the classical implicit and the Gear implicit schemes (see [10]). These semi-implicit schemes are compared with the classical explicit scheme. The consistency and stability are studied specifically for all the considered schemes.

In the case of semi-implicit and Gear semi-implicit schemes, it is necessary to solve at each time step a large linear system. Due to the size of these systems, iterative methods, such as point relaxation or conjugate gradient methods, are preferred to direct methods. An originality of the present study, specific to image processing, comes from the spatial discretization step size along each direction which is fixed implicitly to a constant value defined by the distance between two consecutive pixels. Commonly, in the literature, the value of the spatial discretization step size is fixed to one (see [27] and [38]). This particularity, inherent only to our knowledge to image processing, involves very particular properties of the numerical parameters, mainly the spectral radius and the condition number, allowing to estimate the asymptotic rate of convergence of point relaxation and conjugate gradient methods. However, in this kind of application, despite the Neumann boundary condition, these iterative algorithms have a very high convergence speed. Note that this behavior is ineffective in the physical context of heat conduction equation where the spatial discretization step size can be very small leading to problems hard to solve numerically by the same methods. Precisely, in such a case, the spectral radius of the relaxation matrix is close to one or the condition number leads to infinity.

The remainder of the chapter is organized as follows. In section 2, we study how edges

can be detected in images affected by multiplicative noise. In section 3, the mathematical model of the anisotropic diffusion problem is presented, and an automatic estimation of its parameters is proposed. In section 4, spatial discretization schemes, explicit and semi-implicit discretization time marching schemes are established and analyzed; particularly the consistency and the numerical stability are studied. Section 5 is devoted to the analysis of the behavior of the point relaxation and the conjugate gradient methods; in particular, due to the value of the spatial discretization step size fixed to one, we show why we obtain efficient rate of convergence and why preconditioning is not necessary when conjugate gradient method is used for the solution of each linear algebraic system to solve at each timestep. In section 6, we present the image processing results on both synthetic and real ultrasound images. Finally, in section 7, we present the results of numerical simulations concerning the performance of the iterative methods used for solving the linear systems.

## 2. The Coefficient of Variation as a Measure of Edges in Ultrasound Images

Imaging speckle is a phenomenon which occurs when a coherent source and an incoherent detector are used to interrogate a medium which is rough on the scale of the wavelength. The most widely used techniques to reduce speckle in synthetic aperture radar imaging include the filters of Lee, Frost, Kuan and the Gamma Maximum a Posteriori ([21, 22] and references therein). These methods all rely on the coefficient of variation to characterize the speckle. There are two types of coefficients of variation (CV): the global coefficient of variation, denoted  $\gamma_g$  and the local coefficient of variation, denoted  $\gamma$ .

The global CV is defined as:

$$\gamma_g^2 = \frac{\text{var}(u)}{\bar{u}^2} \quad (1)$$

where  $\text{var}(u)$  and  $\bar{u}$  are the variance and the mean of the intensity image  $u$ , calculated over an area that doesn't contain edges. The local coefficient of variation  $\gamma$  is a local version of the global CV, defined as:

$$\gamma^2(p) = \frac{1}{|\eta(p)|} \sum_{\eta(p)} \frac{(u(x,y,t) - \bar{u}_{\eta(p)})^2}{\bar{u}_{\eta(p)}^2} \quad (2)$$

where  $\eta(p)$  is the neighborhood of a pixel  $p$  and  $|\eta(p)|$  is the number of pixels composing the neighborhood.  $\bar{u}_{\eta(p)}$  is the mean intensity of  $u$  calculated over  $\eta(p)$ . In practice, one can use a  $3 \times 3$  pixels square window for  $\eta(p)$ .

**Proposition 1** *The global coefficient of variation characterizes the level of speckle in the image.*

**Proof.** Let  $R$  be the reflectivity of an observed scene,  $u$  the actual observation (image) and  $\vartheta$  a multiplicative noise:

$$u(i, j) = R(i, j) \cdot \vartheta(i, j) \quad (3)$$

The global CV is calculated in an  $N$ -pixel area  $\Xi_h$  of homogeneous reflectivity:

$$\forall (i, j) \in \Xi_h, R(i, j) = k, \quad (4)$$

where  $k$  is a constant equal to the value of the reflectivity. Thus the global CV can be written as:

$$\gamma_g(\Xi_h) = \frac{\frac{1}{N} \sum_{(i,j) \in \Xi_h} (k\vartheta(i, j) - k\bar{\vartheta})^2}{(k\bar{\vartheta})^2} = \frac{\text{var}(\vartheta)}{\bar{\vartheta}^2},$$

where  $\bar{\vartheta}$  is the mean of the noise over the considered region. This shows that the global CV does not depend on the reflectivity of the scene; it only depends on the multiplicative noise. Hence, it characterizes the speckle affecting the image. ■

**Proposition 2** *The coefficient of variation is an edge detector robust to the speckle.*

**Proof.** The variance of a variable  $A$  can be written as follows:

$$\text{var}(A) = E(A^2) - E(A)^2 \quad (5)$$

It is reasonable to consider that the real reflectivity of a scene and the speckle affecting it are independent. As a consequence, the equation of the local CV can be written:

$$\gamma = \frac{\text{var}(R\vartheta)}{E(R\vartheta)^2} = \frac{1}{E(R\vartheta)^2} [E((R\vartheta)^2) - E(R\vartheta)^2] \quad (6)$$

$$= \frac{1}{E(R\vartheta)^2} [( \text{var}(R) + E(R)^2 ) \cdot ( \text{var}(\vartheta) + E(\vartheta)^2 ) - E(R)^2 E(\vartheta)^2] \quad (7)$$

$$= \underbrace{\frac{\text{var}(\vartheta)E(R)^2}{E(R\vartheta)^2}}_{\text{global CV}} + \text{var}(R) \frac{(\text{var}(\vartheta) + E(\vartheta)^2)}{E(R\vartheta)^2} \quad (8)$$

The first term of equation(8) corresponds to the global CV. The second term is positive, increasing with the variance of the real reflectivity. This shows that  $\gamma$  is close to  $\gamma_g$  when the variance of  $R$  is small, which corresponds to homogeneous regions, and greater than  $\gamma_g$  when the variance of  $R$  is higher, which is the case on edges. Therefore, the coefficient of variation is an effective edge detector for images affected by speckle. ■

### 3. Anisotropic Diffusion for Speckle Filtering

To effectively reduce the speckle, without degrading the features and the structures of images, a filter must comply with the following requirements:

- to be highly effective with any speckle distribution model,
- to respect the three principles of semantically correct scale space [27],
- to preserve edges, even after an infinite number of iterations,
- to ensure the conservation of intensity globally and within regions,

- to require the fewest number of parameters.

In most anisotropic filters in the literature, at least one of these properties is not respected. This is either because they rely on the gradient of intensity to control the diffusion, which is not adequate for multiplicative noise, or because they tend toward a flat image after an infinite diffusion. In order to effectively filter speckle in ultrasound images, we present an original coefficient of diffusion that complies with all the previous requirements. In this section, we first establish the partial derivative equation (PDE) of our anisotropic diffusion, called  $\gamma$ -diffusion. We then derive the property of the associated coefficient of diffusion in order to ensure the conservation of energy. Based on this property, we construct the coefficient of the  $\gamma$ -diffusion and we review the way to calculate its parameters.

### 3.1. The $\gamma$ -diffusion Partial Derivative Equation

In section 2. the local and global versions of the coefficient of variation were presented. We established that comparing these measures leads to effective detection of edges in the presence of multiplicative noise. Therefore, the  $\gamma$ -diffusion model is based on these parameters.

In a bounded domain  $\Omega \subset \mathbb{R}^2$  the partial derivative equation (PDE) of the  $\gamma$ -diffusion is defined as follows:

$$\frac{\partial u(x,y,t)}{\partial t} - \operatorname{div}(c(x,y,t,u)\nabla u(x,y,t)) = 0, \text{ everywhere in } \Omega, 0 < t \leq T \quad (9)$$

where  $\nabla u$  stands for the gradient of the image intensity  $u$ . In order to prevent energy from escaping outside the domain  $\Omega$ , the Neumann boundary condition are set.

The diffusion is controlled by the coefficient of diffusion  $c$ . Before introducing the coefficient of diffusion, we study in the next section how the conservation of energy imposes a constraint that should be respected by the coefficient of diffusion.

### 3.2. Conservation of Intra Region Energy

Classically, the property of energy conservation of an anisotropic diffusion is described using the divergence theorem [4, 16, 42]. Here we consider this theorem from the perspective of the conservation of the energy within homogeneous regions. This approach leads to establish analytically an important property that should be verified by the coefficient of diffusion.

Let  $S$  be a 2D region of the image bounded by  $C = \partial S$ , and  $\vec{n}$  be the outward vector normal to  $C$ . The divergence theorem applied to the two dimensional case of the  $\gamma$ -diffusion leads to :

$$\iint_S \frac{\partial u}{\partial t} dS = \iint_S \operatorname{div}(c(x,y,t,u)\nabla u) dS = \oint_C c(x,y,t,u)\nabla u \cdot \vec{n} dC \quad (10)$$

Applying this result to an image leads to the following conclusions:

1. if we set  $S$  to the entire image domain, then the symmetrical border conditions ensure that  $\nabla u \cdot \vec{n} = 0$  on  $\partial S = C$ . There is thus no diffusion through  $\partial S$ . This implies that no

energy escapes the outer borders of the image. The global energy is therefore strictly conserved.

2. when  $S$  represents a homogeneous region of the image, the intra region energy is preserved if the coefficient of diffusion is zero on  $\partial S$ .

We can notice that classical anisotropic filters with coefficient of diffusion that never reach zero cannot preserve intra region energy. The following section presents our coefficient of diffusion, which do reach zero value on the borders of image regions.

### 3.3. The Coefficient of $\gamma$ -diffusion

Robust statistics and anisotropic diffusion are strongly related [4, 40]. In fact, anisotropic diffusion can be considered as the estimation of a piecewise homogeneous image from noisy data.

The relationship established by You et al. [40] and Black et al. [4] permits to use any weight function of an M-estimator as a basis for a coefficient of diffusion. Table 1 shows the M-estimators used in some relevant anisotropic diffusion models.

**Table 1. Relationship between coefficients of diffusion and M-estimators**

Coefficient	M-estimator
Perona et al. [27]	Cauchy and Welsch
Black et al. [4]	Tukey
Charbonnier et al. [5]	$L_1 - L_2$

The Cauchy M-estimator is known to not guarantee a unique solution, due to its non convexity. To adress this problem in the continuous domain, Catte et al. [6] introduced a gaussian pre-filtering step. The  $L_1 - L_2$  M-estimator has the advantage of reducing the influence of large errors and of being convex; however it never entirely stops the influence of outliers.

Our coefficient of diffusion is derived from the weight function of Tukey [36], applied to the local coefficient of variation, as was proposed in [4]. This function assigns zero weights to outliers having a magnitude above a certain threshold. This allows stopping completely the diffusion in the directions corresponding to high values of the local coefficient of variation. As such, the intra region energy is conserved if the local CV is higher enough that the global CV on the borders of the region. The coefficient of variation acts here as an edge detector more robust to speckle than the gradient of the intensity, leading to a more precise filtering.

Let  $\gamma(x, y, t, u)$  be the local coefficient of variation at  $(x, y)$ , and  $\gamma_s$  a scale parameter estimated in the forthcoming equation (16). The coefficient of  $\gamma$ -diffusion is

$$c(x, y, t, u) = \begin{cases} \left[ 1 - \frac{\gamma^2(x, y, t, u)}{\gamma_s^2(u, t)} \right]^2 & \text{if } \gamma(x, y, t, u) \leq \gamma_s(u, t), \\ 0 & \text{elsewhere.} \end{cases} \quad (11)$$

### 3.4. Parameters Estimation

#### 3.4.1. Robust Estimation of the Coefficient of Variation

Controlling the process of anisotropic diffusion requires a precise edge detection. For this purpose, we exploit property 2 of the coefficient of variation. We measure the degree to which a pixel belongs to an edge by comparing the local and the global coefficients of variation. Obviously, this necessitates the calculation of the latter at each iteration. However, as stated earlier, this coefficient must be calculated in a homogeneous area. The selection of such an area requires an undesirable interaction. In this section we present an automatic estimation of the global coefficient that avoids any interaction and allows a fine assessment of the level of speckle at each iteration.

From property 2, we know that the local CV is greater than the global CV on edges. Therefore, estimating the global CV consists in determining the value of the local CV beyond which pixels are considered as edges.

The estimated global coefficient of variation  $\gamma_e$  is, therefore, a scale parameter from a set of data with unknown probability density function. It is reasonable to consider that there are far less pixels associated with edges than pixels in homogeneous regions. The values of the local CV corresponding to edges can thus be considered as outliers. In order to diagnose these outliers, we first robustly normalize the observations [29]:

$$v_{i,j} = \frac{\gamma(i,j) - \underset{p \in \Omega}{\text{med}}(\gamma(p))}{1.4826 \underset{q \in \Omega}{\text{med}}(|\gamma(q) - \underset{r \in \Omega}{\text{med}}(\gamma(r))|)} \quad (12)$$

where  $\underset{p \in \Omega}{\text{med}}(\gamma(p))$  is the median of  $\gamma$  at the current iteration.

$\gamma_e$  is the value of  $\gamma$  corresponding to the rejection rule  $v_{i,j} = 1$ , established by Rousseeuw [29]. Thus:

$$\gamma_e = 1.4826 \underset{p \in \Omega}{\text{med}}(|\gamma(p) - \underset{q \in \Omega}{\text{med}}(\gamma(q))|) + \underset{r \in \Omega}{\text{med}}(\gamma(r)). \quad (13)$$

#### 3.4.2. Scale Parameter Estimation

The diffusion should be strong when  $\gamma < \gamma_e$ , since this condition indicates a homogeneous region. The process should progressively decrease thereafter, until it becomes totally nil. We have, then, two parameters:

- $\gamma_e$ : the threshold above which a pixel is considered to be on a contour;
- $\gamma_s$ : the threshold above which the diffusion is stopped in the corresponding direction.

$\gamma_s$  is calculated from  $\gamma_e$ . Let  $|\Psi(x,y,t,u)| = c(x,y,t,u)|\nabla u(x,y,t)|$  be the diffusion flow. We should have  $\Psi'(x,y,t,u) > 0$  when the pixel  $(x,y)$  is in an homogeneous area and  $\Psi'(x,y,t,u) < 0$  when it is on a contour [27]. Thus, when  $\gamma = \gamma_e$ , the function  $\Psi'$  should be zero and  $|\Psi|$  should be maximal. Since  $|\Psi|$  depends on the gradient of intensity, we cannot determine its maximum by only considering the value of  $\gamma$ . To solve this problem and determine  $\gamma_s$ , we consider that  $|\Psi|$  reaches its maximum when  $\gamma c(x,y,t,u)$  is maximal:

$$\operatorname{argmax}_{\gamma}(\gamma c(x, y, t, u)) = \gamma_e. \quad (14)$$

Writing  $\Gamma = \frac{\gamma}{\gamma_s}$ , the maximum can be found as:

$$[\Gamma(1 - \Gamma^2)^2]' = 0 \Leftrightarrow \gamma = \frac{\gamma_s}{\sqrt{5}} \quad (15)$$

We finally obtain

$$\gamma_s = \sqrt{5}\gamma_e. \quad (16)$$

## 4. Discretization Schemes of Anisotropic Diffusion

Consider the following model of anisotropic diffusion of speckled images

$$\begin{cases} \frac{\partial u(x, y, t)}{\partial t} - \operatorname{div}(c \nabla u(x, y, t)) = 0, \text{ everywhere in } \Omega, 0 < t \leq T, \\ \frac{\partial u(x, y, t)}{\partial n} |_{\partial \Omega} = 0, \quad \forall t \in [0, T] \text{ (Boundary Condition),} \\ u(x, y, 0) = u_0(x, y), \text{ (Initial Condition),} \end{cases} \quad (17)$$

where  $\Omega = [0, d] \times [0, b]$  is a rectangular domain,  $\partial \Omega$  is the boundary of the domain  $\Omega$ ,  $u = u(x, y, t)$  is the intensity,  $T$  is a strictly positive real number,  $n$  is the outward vector normal to  $\partial \Omega$  and the nonnegative coefficient of diffusion  $c = c(x, y, t, u)$  depends of the intensity  $u$ . Note that, for this particular application, we have

$$0 \leq c(x, y, t, u) \leq 1, \text{ everywhere in } \Omega, 0 < t \leq T. \quad (18)$$

Note also that the boundary value problem (17) is nonlinear. In order to solve numerically the problem (17), we will construct and analyze in the sequel appropriate discretization schemes [33].

### 4.1. Spatial Discretization Scheme

For the spatial discretization of the operator arising in (17), it is necessary to take into account that the diffusion coefficient is not constant in the domain  $\Omega$ . Then, the discretization scheme is carried out by taking the mean of two intermediate schemes, called in the sequel forward-backward scheme and backward-forward scheme.

#### 4.1.1. Forward-backward Scheme

Consider the approximation of  $-\frac{\partial}{\partial x}(c(x_i, y_j) \frac{\partial u(x_i, y_j)}{\partial x})$ , in the case where the spatial discretization step, denoted by  $h$  in the sequel, is uniform and  $y_j$  is fixed. Then, if appropriate assumption of regularity concerning  $c$  and  $u$  are verified, the forward-backward scheme

leads to

$$\begin{aligned} \frac{\partial}{\partial x}(c(x_i, y_j) \frac{\partial u(x_i, y_j)}{\partial x}) &= \frac{1}{h}(c(x_{i+1}, y_j) \frac{u(x_{i+1}, y_j) - u(x_i, y_j)}{h}) \\ &\quad - \frac{1}{h}(c(x_i, y_j) \frac{u(x_i, y_j) - u(x_{i-1}, y_j)}{h}) + \mathcal{E}_x^f(x_i, y_j), \end{aligned}$$

where  $\mathcal{E}_x^f(x_i, y_j)$  is the local truncation error given by

$$\mathcal{E}_x^f(x_i, y_j) = -\frac{h}{2} \left( \frac{\partial c(x_i, y_j)}{\partial x} \frac{\partial^2 u(x_i, y_j)}{\partial x^2} + \frac{\partial^2 c(x_i, y_j)}{\partial x^2} \frac{\partial u(x_i, y_j)}{\partial x} \right) + O(h^2).$$

#### 4.1.2. Backward-forward Scheme

Similarly, when the same properties of  $c$  and  $u$  hold, the backward-forward scheme leads to

$$\begin{aligned} \frac{\partial}{\partial x}(c(x_i, y_j) \frac{\partial u(x_i, y_j)}{\partial x}) &= \frac{1}{h}(c(x_i, y_j) \frac{u(x_{i+1}, y_j) - u(x_i, y_j)}{h}) \\ &\quad - \frac{1}{h}(c(x_{i-1}, y_j) \frac{u(x_i, y_j) - u(x_{i-1}, y_j)}{h}) + \mathcal{E}_x^b(x_i, y_j), \end{aligned}$$

where  $\mathcal{E}_x^b(x_i, y_j)$  is the corresponding local truncation error given by

$$\mathcal{E}_x^b(x_i, y_j) = \frac{h}{2} \left( \frac{\partial c(x_i, y_j)}{\partial x} \frac{\partial^2 u(x_i, y_j)}{\partial x^2} + \frac{\partial^2 c(x_i, y_j)}{\partial x^2} \frac{\partial u(x_i, y_j)}{\partial x} \right) + O(h^2).$$

#### 4.1.3. Final Discretization Scheme

Then, the final discretization scheme of the considered operator is obtained by taking the mean of the two previous schemes; so we obtain

$$\begin{aligned} \frac{\partial}{\partial x}(c(x_i, y_j) \frac{\partial u(x_i, y_j)}{\partial x}) &= \frac{1}{2h^2} ((c(x_{i+1}, y_j) + c(x_i, y_j))u(x_{i+1}, y_j) \\ &\quad - (c(x_{i+1}, y_j) + 2c(x_i, y_j) + c(x_{i-1}, y_j))u(x_i, y_j) \\ &\quad + (c(x_i, y_j) + c(x_{i-1}, y_j))u(x_{i-1}, y_j)) + O(h^2). \end{aligned}$$

Similarly, the second partial derivative with respect to  $y$  is obtained by using the same procedure, when  $x_i$  is fixed; then we obtain

$$\begin{aligned} \frac{\partial}{\partial y}(c(x_i, y_j) \frac{\partial u(x_i, y_j)}{\partial y}) &= \frac{1}{2h^2} ((c(x_i, y_{j+1}) + c(x_i, y_j))u(x_i, y_{j+1}) \\ &\quad - (c(x_i, y_{j+1}) + 2c(x_i, y_j) + c(x_i, y_{j-1}))u(x_i, y_j) \\ &\quad + (c(x_i, y_j) + c(x_i, y_{j-1}))u(x_{i-1}, y_{j-1})) + O(h^2). \end{aligned}$$

The boundary value problem (17) is usually solved using an implicit or a semi-implicit time marching numerical scheme. To take into account such temporal discretization, we

have to perturb the stationary boundary value problem with Neumann boundary condition, associated with problem (17) by  $\delta u$ , where  $\delta$  is a positive real number. In fact, according to the numerical time marching scheme used,  $\delta$  is proportional to the inverse of the time step. Then, according to the previous remark, for the two-dimensional elliptic problem, we have to discretize the following stationary problem

$$\begin{cases} -\operatorname{div}(c(x,y)\nabla u(x,y)) + \delta u(x,y) = 0, \text{ everywhere in } \Omega, \\ \frac{\partial u(x,y)}{\partial n} \Big|_{\partial\Omega} = 0, \text{ (Boundary Condition),} \end{cases}$$

instead of the differential operator  $-\operatorname{div}(c(x,y)\nabla u(x,y))$ . Thus, for the two-dimensional elliptic problem,

$$\begin{cases} -\frac{\partial}{\partial x}(c(x,y)\frac{\partial u}{\partial x}) - \frac{\partial}{\partial y}(c(x,y)\frac{\partial u}{\partial y}) + \delta u(x,y) = f(x,y), \text{ in } ]0,a[ \times ]0,b[, \delta > 0, \\ \frac{\partial u(x,y)}{\partial n} \Big|_{\partial\Omega} = 0, \end{cases} \quad (19)$$

we can derive a global scheme from the two previous discretization approximations of the two considered operators. For sake of simplicity, we also denote by  $c_{i,j}$  for all  $i$  and  $j$  the value of  $c(x_i, y_j)$ . According to the size of the image to process, we also denote by  $n+2$  and  $m+2$ , respectively the number of pixels on  $[0, d]$  and  $[0, b]$ , respectively. Assume for example, that the grid points are numbered line by line in a natural ordering from left to right; assume also that the inner grid points of coordinates  $x_i = ih, 1 \leq i \leq n$ , and  $y_j = jh, 1 \leq j \leq m$ , are labelled by an index  $l$  defined according to the considered ordering, by  $l = (i-1)n + j$ . Then, by combining the two previous final schemes obtained in each direction, we can set the following scheme associated to the inner grid points

$$-\bar{a}_{l,l-n}u_{l-n} - \bar{a}_{l,l-1}u_{l-1} + (\bar{a}_{l,l} + \delta)u_l - \bar{a}_{l,l+1}u_{l+1} - \bar{a}_{l,l+n}u_{l+n} = f_l, 1 \leq l \leq nm,$$

where for  $1 \leq i \leq n$  and  $1 \leq j \leq m$ ,  $u_l \equiv u_{i,j}$ ; then the absolute value of the off-diagonal entries of the matrix are given by

$$\bar{a}_{l,l-n} = \frac{(c_{i,j} + c_{i-1,j})}{2h^2} \text{ and } \bar{a}_{l,l+n} = \frac{(c_{i,j} + c_{i+1,j})}{2h^2},$$

$$\bar{a}_{l,l-1} = \frac{(c_{i,j} + c_{i,j-1})}{2h^2} \text{ and } \bar{a}_{l,l+1} = \frac{(c_{i,j} + c_{i,j+1})}{2h^2},$$

and the diagonal ones are defined by

$$\bar{a}_{l,l} = \frac{(c_{i-1,j} + 2c_{i,j} + c_{i+1,j})}{2h^2} + \frac{(c_{i,j-1} + 2c_{i,j} + c_{i,j+1})}{2h^2}.$$

Note also that, due to the Neumann boundary condition, for the grid points near the boundary  $\partial\Omega$ , the above scheme can be obviously simplified in a straightforward way by taking into account that the value of the intensity at this grid point is equal to the value of the intensity at the next grid nodes which belongs to  $\partial\Omega$ . Let us denote by  $\bar{A}$  the discretization matrix with entries  $\bar{a}_{l,j}$ . Then, it can be verified that the discretization matrix  $\bar{A}$  is symmetric with strictly positive diagonal entries and nonpositive off-diagonal entries. Thus, the matrix

$A$ , defined by  $A = 2h^2\bar{A}$ , has the same properties. Moreover, the matrix  $(\delta I + \frac{1}{2h^2}A)$ , where  $I$  is the identity matrix, is symmetric and strictly diagonally dominant, with strictly positive diagonal entries. Consequently, by applying a classical result [2],  $(\delta I + \frac{1}{2h^2}A)$  is symmetric positive definite and, obviously is a Stieltjes matrix (or symmetric M-matrix). Moreover the local truncation error for the approximation is given by  $O(h^2)$  and the resulting spatial discretization scheme is of order two and is consequently consistent.

#### 4.2. Explicit Discretization Scheme for the Anisotropic Diffusion Problem

In this subsection, we draw up an appropriate explicit discretization scheme for the nonlinear boundary value problem (17), by taking into account the scheme obtained in subsection 4.1.. Let us write now the problem (17) for  $t_k = k\tau, k \in \mathbb{N}$ , where  $\tau$  denotes the time step; then, we can write

$$\frac{\partial u(x, y, t_k)}{\partial t} - \frac{\partial}{\partial x} (c(x, y, t_k, u^{(k)}) \frac{\partial u(x, y, t_k)}{\partial x}) - \frac{\partial}{\partial y} (c(x, y, t_k, u^{(k)}) \frac{\partial u(x, y, t_k)}{\partial y}) = 0,$$

where  $u^{(k)}$  corresponds to the value of  $u$  when  $t = t_k$ . Let  $\alpha = \frac{\tau}{2h^2}$ ; note that, for the particular application concerning image processing, the spatial discretization step size  $h$  is here mentioned for memory; in fact in image processing, the spatial discretization step size along each direction  $h$  is defined implicitly by the distance between two pixels and can be fixed so that  $h = h_x = h_y$ ; practically,  $h$  is fixed to the value one (see [27] and [37]). In the sequel, we will see that the value of  $h$  has no effect with respect to the behaviour of the studied algorithms since the spatial step size is non-dimensional. Denote also by  $U^{(k)}$  the vector of components  $u_l^{(k)}, 1 \leq l \leq nm$ , approximation of  $u(x_i, y_j, t_k)$ . In the sequel  $U^{(0)}$  is given by the components  $u_0(x_i, y_j)$  of the initial condition arising in (17). Then, classically, the explicit scheme is given by

$$U^{(k+1)} = (I - \alpha A^{(k)})U^{(k)}, k = 0, 1, \dots \quad (20)$$

Entries of the matrix  $A^{(k)}$  are defined in a similar way than the entries of the matrix  $A$  in subsection 4.1. ; in fact this matrix  $A^{(k)}$  represents the spatial discretization matrix defined at step  $k$  of the problem (17) in which the coefficients  $c_{i,j}$  are replaced by  $c_{i,j}^{(k)}$  with  $c_{i,j}^{(k)} = c(x_i, y_j, t_k, u_{i,j}^{(k)})$ . Then, we have immediately the following results.

**Proposition 3** *Assume that the diffusion coefficient  $c$  and the intensity  $u$  are continuous and sufficiently differentiable. Then, the explicit scheme (20) is consistent and the local truncation error of the scheme denoted by  $E_{i,j}^{(k)}$  verifies  $|E_{i,j}^{(k)}| \leq C(\tau + h^2)$ , where  $C$  is a constant.*

**Proof.** The result is obviously an immediate consequence of the fact that the explicit scheme (20) is of order one with respect to the time and of order two with respect to the space. ■

**Remark 1** *Note that, except for edges where the diffusion coefficient vanishes, images are mostly constituted of homogeneous regions where the coefficient of variation is relatively smooth. So, diffusion coefficient is also smooth and the previous discretization scheme is based on realistic framework.*

**Proposition 4** Denote by  $\lambda_l^{(k)}$ ,  $l = 1, \dots, nm$  the eigenvalues of the spatial discretization matrix at each time step  $k$ . Denote also by  $\lambda_M^{(k)}$  the greatest eigenvalue of the matrix  $A^{(k)}$  at each step  $k$ . If the following condition  $\alpha < \frac{2}{\lambda_M^{(k)}}$  is verified, then the explicit scheme (20) is stable.

**Proof.** Note that, at each time step  $k$ , the discretization matrix  $A^{(k)}$  is symmetric. Then  $A^{(k)}$  is diagonalizable and the eigenvalues of the matrix  $(I - \alpha A^{(k)})$  are equal to  $\mu_l^{(k)} = 1 - \alpha \lambda_l^{(k)}$ . Thus obviously, the explicit scheme (20) is stable if  $-1 < 1 - \alpha \lambda_l^{(k)} < 1$ ; so the expected condition leads by taking the more restrictive bound, which achieves the proof. ■

**Corollary 1** Denote by  $\|A^{(k)}\|$  every norm of the matrix  $A^{(k)}$ . If the time step size verifies for each  $k \in \mathbb{N}$

$$\tau < \frac{4h^2}{\|A^{(k)}\|},$$

then the explicit scheme (20) is stable. Particularly, if we consider the case where the norm of the matrix  $A^{(k)}$  is induced by the uniform norm or by the  $l_1$ -norm, respectively, and if we denote by  $\|A^{(k)}\|_\infty$  or by  $\|A^{(k)}\|_1$ , respectively, the corresponding particular matrix norm, then the previous condition can be written as follows

$$\tau < \frac{4h^2}{\max_{1 \leq i \leq nm} \left( \sum_{j=1}^{nm} |a_{i,j}^{(k)}| \right)} = \frac{4h^2}{\max_{1 \leq j \leq nm} \left( \sum_{i=1}^{nm} |a_{i,j}^{(k)}| \right)}. \tag{21}$$

**Proof.** Recall that, for every norm we have  $\lambda_M^{(k)} \leq \|A^{(k)}\|, \forall k \in \mathbb{N}$ . Then, when  $\lambda_M^{(k)} = \|A^{(k)}\|$  the first condition follows from a straightforward way from Proposition 4. Since the discretization matrix is symmetric

$$\|A^{(k)}\|_\infty = \max_{1 \leq i \leq nm} \left( \sum_{j=1}^{nm} |a_{i,j}^{(k)}| \right) \text{ and } \|A^{(k)}\|_1 = \max_{1 \leq j \leq nm} \left( \sum_{i=1}^{nm} |a_{i,j}^{(k)}| \right);$$

then  $\|A^{(k)}\|_\infty = \|A^{(k)}\|_1$ , which achieves the proof. ■

**Remark 2** In the case where the differential operator is the classical Laplacian multiplied by  $c$ , where  $c$  denotes here the constant diffusion coefficient, then the discretization matrix  $A$  is such that  $\sum_{j=1}^{nm} |a_{i,j}| = 8$ , for all  $i \in \{1, \dots, nm\}$ ; then the previous condition (21) is  $\tau < \frac{h^2}{4c}$ .

This previous bound is consistent with both bounds for this classical problem and with the result of Proposition 4.

### 4.3. Semi-implicit Discretization Scheme for the Anisotropic Diffusion Problem

Due to the stability condition, and particularly the fact that this condition implies that the time step  $\tau$  must be small, one likes better use fully-implicit time marching scheme; in our case, such fully-implicit scheme is defined by

$$(I + \alpha A^{(k+1)})U^{(k+1)} = U^{(k)}, k = 0, 1, \dots$$

Due to the nonlinearity, particularly the definition of the diffusion coefficient, the previous algebraic system can be hard to solve. If the Newton method is used, the derivative of the diffusion coefficient with respect to the components of the intensity, can be hard to obtain, due to the definition of the diffusion coefficient. Moreover, the solution of the fully-implicit scheme can be more complicated to solve and the iterative methods, like the Newton method, do not necessarily converge. It is the reason why, for image processing, the following semi-implicit time marching scheme

$$(\theta I + \alpha A^{(k)})U^{(k+1)} = U^{(k)}, k = 0, 1, \dots,$$

is preferred. In this scheme,  $\theta$  can take the two values 1 or  $\frac{3}{2}$  corresponding to the following situations

$$\theta = \begin{cases} 1 & \text{for the use of the semi-implicit scheme} \\ \frac{3}{2} & \text{for the use of the Gear semi-implicit scheme.} \end{cases} \quad (22)$$

More precisely the semi-implicit scheme is defined as follows

$$(I + \alpha A^{(k)})U^{(k+1)} = U^{(k)}, k = 0, 1, \dots, \quad (23)$$

and corresponds to a time marching scheme with two levels  $k$  and  $k + 1$ , while the Gear semi-implicit scheme corresponds to a scheme with three levels  $k - 1$ ,  $k$  and  $k + 1$  and satisfies

$$\left(\frac{3}{2}I + \alpha A^{(k)}\right)U^{(k+1)} = 2U^{(k)} - \frac{1}{2}U^{(k-1)}, k = 1, \dots; \quad (24)$$

note that it is necessary to initialize the Gear semi-implicit scheme, for example, with the semi-implicit scheme or the explicit scheme. Concerning the numerical analysis of the behavior of the two semi-implicit time marching schemes, we have the following result.

**Proposition 5** *Assume that the diffusion coefficient  $c$  and the intensity  $u$  are continuous and sufficiently differentiable. Then, the semi-implicit schemes (23) and (24) are unconditionally stable. Moreover,*

- 1) *the semi-implicit scheme (23) is consistent and the local truncation error of the scheme denoted by  $(E_{i,j}^{si})^{(k)}$  verifies  $|(E_{i,j}^{si})^{(k)}| \leq C^{si}(\tau + h^2)$ , where  $C^{si}$  is a constant,*
- 2) *the semi-implicit Gear scheme (24) is consistent and the local truncation error of the scheme denoted by  $(E_{i,j}^{ge})^{(k)}$  verifies  $|(E_{i,j}^{ge})^{(k)}| \leq C^{ge}(\tau^2 + h^2)$ , where  $C^{ge}$  is a constant.*

**Proof.** The proof follows from a straightforward adaptation from the one considered to prove similar result in the case of the classical heat conduction equation (see for example [10]). ■

**Remark 3** *Clearly, it follows from the result of the Proposition 5 that unlike the explicit scheme (20), the semi-implicit schemes (23) and (24) do not suffer from any time step size restriction and can be fully adapted to the desired accuracy without the need to choose small time steps for stability reasons. Moreover, note that Remark 1 holds again here.*

## 5. Sequential Iterative Methods for Image Processing

In the previous subsection 4.3., we have shown that the semi-linearization and the discretization of the anisotropic diffusion problem leads to the solution of a large scale linear algebraic system. The real number  $\theta$  being defined by (22), the matrix  $\mathcal{A}^{(k)}$  is a regular matrix, since at each time step  $k \in \mathbb{N}$  it was verified that the matrix  $\mathcal{A}^{(k)} = (\theta I + \alpha A^{(k)})$ , arising in the linear system to solve, is symmetric positive definite. Moreover, since the algebraic system to solve is a large scale system, iterative methods are numerically more appropriate. From a general point of view concerning the iterative sequential method we have the following result.

**Proposition 6** *The matrix defined by  $\mathcal{A}^{(k)} = (\theta I + \alpha A^{(k)})$ ,  $\forall k \in \mathbb{N}$ , is an M-matrix.*

**Proof.** Since  $\theta$  is positive, the matrix  $\mathcal{A}^{(k)}$  is strictly diagonally dominant; furthermore the diagonal entries of  $\mathcal{A}^{(k)}$  are strictly positive and the off-diagonal entries are non positive. Then  $\mathcal{A}^{(k)}$  is an M-matrix ■

**Remark 4** *Note that due to the value of the diffusion coefficient, all the coefficients of the same row (or possibly of several rows) of the matrix  $A^{(k)}$  can be equal to zero. Then, the matrix  $A^{(k)}$  and consequently the matrix  $\mathcal{A}^{(k)}$  is not necessarily irreducible. Nevertheless, due to the strict diagonal dominance,  $\mathcal{A}^{(k)}$  is invertible.*

The result of Proposition 6 is interesting for the study of the behaviour of the iterative method, particularly regarding the convergence of such methods. Recall the following result

**Proposition 7** *Let  $\mathcal{A}^{(k)}$  be a nonsingular M-matrix where  $\mathcal{A}^{(k)} = C^{(k)} - \mathcal{R}^{(k)}$  is a splitting of  $\mathcal{A}^{(k)}$  such that  $\mathcal{T}^{(k)} = (C^{(k)})^{-1}\mathcal{R}^{(k)}$  is a nonnegative matrix. Then  $\rho(\mathcal{T}^{(k)}) < 1$ , where  $\rho(\mathcal{T}^{(k)})$  is the spectral radius of the matrix  $\mathcal{T}^{(k)}$  i.e.  $\mathcal{A}^{(k)} = C^{(k)} - \mathcal{R}^{(k)}$  is a convergent splitting.*

The reader is referred to the litterature, in particular to [2] for a proof of this classical result. Nevertheless, from a practical point of view, it is interesting to obtain sharp estimates of the spectral radius of the iteration matrix, previously denoted  $\mathcal{T}^{(k)}$  in Proposition 7. When possible, the obtainment of such estimates can help the user to choose efficiently sequential iterative methods for image processing.

We first consider the point relaxation method and we begin by the point Jacobi method. According to the previous notations, at each time step  $k \in \mathbb{N}$  we can write  $A^{(k)} = D^{(k)} - (L^{(k)} + (L^{(k)})^T)$ , where  $D^{(k)}$  is the diagonal part of  $A^{(k)}$  and  $L^{(k)}$  is the strictly lower triangular part of the matrix  $A^{(k)}$ . Denote by  $\mathcal{D}^{(k)} = \alpha D^{(k)} + \theta I$  and  $\mathcal{J}^{(k)}$  the Jacobi's matrix associated to the matrix  $\mathcal{A}^{(k)} = (\theta I + \alpha A^{(k)})$ . Then  $\mathcal{J}^{(k)} = \alpha(\mathcal{D}^{(k)})^{-1}(L^{(k)} + (L^{(k)})^T)$ . The spectral radius  $\rho^{(k)}$  of the matrix  $\mathcal{J}^{(k)}$  verifies  $\rho^{(k)} \leq \|\mathcal{J}^{(k)}\|_{\infty}$ . Since

$$\sum_{j \neq l} a_{l,j}^{(k)} \leq a_{l,l}^{(k)}, \forall l \in [1, nm] \text{ and } \forall k \in \mathbb{N},$$

then the matrix  $(\alpha D^{(k)} + \theta I)^{-1} D^{(k)}$  being a diagonal matrix, the spectral radius  $\rho^{(k)} = \rho(\mathcal{J}^{(k)})$  verifies

$$\rho^{(k)} = \rho(\alpha(\mathcal{D}^{(k)})^{-1}(L^{(k)} + (L^{(k)})^T)) \leq \max_{1 \leq l \leq nm} \left( \frac{\alpha \sum_{j \neq l} a_{l,j}^{(k)}}{\alpha a_{l,l}^{(k)} + \theta} \right) \leq \max_{1 \leq l \leq nm} \left( \frac{\alpha a_{l,l}^{(k)}}{\alpha a_{l,l}^{(k)} + \theta} \right).$$

Note that  $\rho^{(k)} < 1$ . Consider the mapping  $\phi : s \rightarrow \phi(s) = \frac{\alpha s}{\alpha s + \theta}$ . The derivative of the mapping  $\phi$  with respect to the variable  $s$  is given by  $\phi'(s) = \frac{\alpha\theta}{(\alpha s + \theta)^2} > 0$ . So the mapping  $\phi$  is strictly isotone and

$$\rho^{(k)} \leq \frac{\alpha \max_{1 \leq l \leq nm} (a_{l,l}^{(k)})}{\alpha \max_{1 \leq l \leq nm} (a_{l,l}^{(k)}) + \theta} < 1.$$

Due to (18), the diffusion coefficient is such that  $0 \leq c \leq 1$ , and then  $\max_{1 \leq l \leq nm} (a_{l,l}^{(k)}) \leq 8$ ; thus

$$\rho^{(k)} \leq \hat{\rho} = \frac{8\alpha}{8\alpha + \theta} = \frac{4\tau}{4\tau + \theta} < 1. \tag{25}$$

Since the derivative with respect to  $\tau$  of the mapping  $\psi : \tau \rightarrow \psi(\tau) = \frac{4\tau}{4\tau + \theta}$  is equal to  $\psi'(\tau) = (\frac{4\tau}{4\tau + \theta})' = \frac{4\theta}{(\theta + 4\tau)^2}$ , then this derivative is strictly positive and  $\rho^{(k)}$  is a strictly isotone mapping with respect to  $\tau$ .

**Remark 5** *The previous estimate  $\hat{\rho}$  of the spectral radius  $\rho^{(k)}$  of the Jacobi matrix is global at each time step  $k$ . Note that the matrix  $A^{(k)}$  changes during the time marching process. Consequently, the corresponding value of  $\rho^{(k)}$ , may also change at each time step  $k \in \mathbb{N}$ . Then a sharper bound of  $\rho^{(k)}$  can be obtained by using the classical inequality*

$$\rho^{(k)} \leq \|j^{(k)}\|, \forall k \in \mathbb{N},$$

where  $\|j^{(k)}\|$  denotes the matrix norm of the Jacobi matrix at each step  $k$ . In order to obtain an accurate estimation of  $\rho^{(k)}$  with the minimum computational cost, it is recommended to use for the norm of  $\|j^{(k)}\|$  the one induced by the uniform vector norm or by the  $l_1$  vector norm. Then, using the uniform vector norm, since  $j^{(k)}$  is a nonnegative matrix, we obtain

$$\rho^{(k)} \leq \max_{1 \leq i \leq nm} \left( \sum_{j=1}^{nm} j_{i,j}^{(k)} \right).$$

Note that we obtain a similar result when using the  $l_1$  vector norm ; consequently the estimate of  $\rho^{(k)}$  is very easy.

**Remark 6** *For sake of clarity, consider that the matrix  $A^{(k)}$  is the sum of two matrices  $A_x^{(k)}$  and  $A_y^{(k)}$ , where each previous matrix corresponds to the discretization of the differential operators with respect to  $x$  and  $y$ , respectively. Consider also that the spatial step size  $h_x$  and  $h_y$  are different in each direction  $x$  and  $y$  and are defined respectively by  $h_x = \frac{d}{n+1}$  and  $h_y = \frac{b}{m+1}$ , where  $d$  and  $b$ , respectively, are the length and the width, respectively, of the domain  $\Omega$ . Denote by  $\alpha_x = \frac{\tau}{2} \frac{(n+1)^2}{d^2} = \frac{\tau}{2} \frac{1}{h_x^2}$ , and  $\alpha_y = \frac{\tau}{2} \frac{(m+1)^2}{b^2} = \frac{\tau}{2} \frac{1}{h_y^2}$ . With the previous notations, we can write*

$$j^{(k)} = (\theta I + \alpha_x D_x^{(k)} + \alpha_y D_y^{(k)})^{-1} (\alpha_x (L_x^{(k)} + (L_x^{(k)})^T) + \alpha_y (L_y^{(k)} + (L_y^{(k)})^T)),$$

where  $A_x^{(k)} = D_x^{(k)} - L_x^{(k)} - (L_x^{(k)})^T$  and similarly for  $A_y^{(k)}$ . Since the sum of the absolute values of the off-diagonal entries of both matrices  $A_x^{(k)}$  and  $A_y^{(k)}$  are less or equal to the

corresponding diagonal entries  $(a_x)_{l,l}$  and  $(a_y)_{l,l}$ , then the spectral radius  $\rho^{(k)}$  of the Jacobi matrix  $\mathcal{J}^{(k)}$  satisfies

$$\rho^{(k)} \leq \frac{\alpha_x(a_x)_{l,l} + \alpha_y(a_y)_{l,l}}{\theta + \alpha_x(a_x)_{l,l} + \alpha_y(a_y)_{l,l}}.$$

Let  $X = \frac{(a_x)_{l,l}}{h_x^2}$  and  $Y = \frac{(a_y)_{l,l}}{h_y^2}$ ; then

$$\rho^{(k)} \leq \frac{\tau}{2} \frac{X + Y}{\theta + \frac{\tau}{2}(X + Y)}.$$

Moreover, since the diffusion coefficient is bounded by one, we have obviously  $(a_x)_{l,l} \leq 4$  and  $(a_y)_{l,l} \leq 4$ ; then

$$\rho^{(k)} \leq \hat{\rho} = 2\tau \frac{\frac{(n+1)^2}{d^2} + \frac{(m+1)^2}{b^2}}{\theta + 2\tau(\frac{(n+1)^2}{d^2} + \frac{(m+1)^2}{b^2})} < 1 - \varepsilon,$$

where  $\varepsilon$  is a small positive real number. In fact, the number  $n$  and  $m$  being in the general case sufficiently large, from a numerical point of view only, we can see that  $\rho^{(k)}$  has a value close to one. Nevertheless, in image processing, recall that one considers that the parameters are non-dimensional such that  $n + 1 \simeq d$  and  $m + 1 \simeq b$ , and, consequently the spatial discretization step size satisfies  $h_x = h_y \approx 1$  (see [27] and [37]); so the upper bound  $\hat{\rho}$  of the number  $\rho^{(k)}$  is well given by (25) and, since  $\tau$  is small, the upper bound of  $\rho^{(k)}$  is small.

Classically, note that we can deduce from the estimate of the spectral radius of the Jacobi matrix, the asymptotic rate of convergence of the point Jacobi's method, denoted by  $R_\infty(\mathcal{J}^{(k)})$  and defined by

$$R_\infty(\mathcal{J}^{(k)}) = -Ln(\rho^{(k)}).$$

Consider now the use of the point Gauss-Seidel method for the iterative solution of the algebraic system derived from the discretization of the anisotropic diffusion problem by the semi-implicit time marching scheme. The study of the sequential Gauss-Seidel algorithm is easier thanks to the analysis of the behaviour of the sequential Jacobi method. Indeed, if  $\rho(\mathcal{G}S^{(k)})$  denotes the spectral radius of the iteration matrix associated to the Gauss-Seidel method, then, classically, we have  $\rho(\mathcal{G}S^{(k)}) = (\rho(\mathcal{J}^{(k)}))^2$ ; using the above obtained estimates we have

$$\rho(\mathcal{G}S^{(k)}) \leq \left( \frac{\alpha \max_{1 \leq l \leq nm} (a_{l,l}^{(k)})}{\alpha \max_{1 \leq l \leq nm} (a_{l,l}^{(k)}) + \theta} \right)^2 < 1,$$

and the corresponding asymptotic rate of convergence of the point Gauss-Seidel's method, denoted by  $R_\infty(\mathcal{G}S^{(k)})$  obviously verifies  $R_\infty(\mathcal{G}S^{(k)}) = 2R_\infty(\mathcal{J}^{(k)})$ ; so the Gauss-Seidel method is twice faster than the Jacobi method. We summarize hereafter in Table 2 some values of the spectral radius of the point Jacobi matrix and of the point Gauss-Seidel matrix, for different values of  $\tau$ , in both cases of semi-implicit and Gear semi-implicit time marching schemes.

**Table 2. Bounds of spectral radius of the Jacobi and Gauss-Seidel matrices**

$\tau$	Semi-implicit method		Semi-implicit Gear method	
	Jacobi	Gauss-Seidel	Jacobi	Gauss-Seidel
$\frac{1}{5}$	0.44444	0.19753	0.34782	0.12098
$\frac{1}{7}$	0.36363	0.13223	0.27586	0.07609
$\frac{1}{10}$	0.28571	0.08163	0.21052	0.04432
$\frac{1}{20}$	0.16666	0.02777	0.11764	0.01384

Concerning the behavior of the point successive over relaxation method, classically, by the Frankel and Young theory, we know that the optimal relaxation parameter is given by

$$\omega_{opt}^{(k)} = \frac{2}{1 + \sqrt{1 - (\rho(j^{(k)}))^2}}. \quad (26)$$

Using the previous estimate  $\hat{\rho}$  of  $\rho^{(k)}$ , let us estimate an appropriate value  $\hat{\omega}_{opt}$  of  $\omega_{opt}^{(k)}$ . Since  $\rho^{(k)} \simeq \hat{\rho} = \frac{4\tau}{4\tau + \theta}$ , then  $1 - (\rho^{(k)})^2 \simeq \frac{\theta(\theta + 8\tau)}{(\theta + 4\tau)^2}$ . Thus, as  $\sqrt{\theta(\theta + 8\tau)} = \theta(1 + \frac{8\tau}{\theta})^{\frac{1}{2}}$ , then  $\sqrt{\theta(\theta + 8\tau)} \simeq \theta(1 + \frac{4\tau}{\theta}) = (\theta + 4\tau)$ . Thus, an estimate  $\hat{\omega}_{opt}$  of  $\omega_{opt}^{(k)}$  is given by  $\hat{\omega}_{opt} \simeq 1$ . In fact, the previous estimate of  $\omega_{opt}^{(k)}$  can be slightly improved by considering few additional terms in the Taylor's formula of  $\sqrt{\theta(\theta + 8\tau)}$ . We can write

$$(1 + \frac{8\tau}{\theta})^{\frac{1}{2}} \simeq 1 + \frac{4\tau}{\theta} - \frac{8\tau^2}{\theta^2},$$

and then

$$\omega_{opt}^{(k)} \simeq \hat{\omega}_{opt} = \frac{2}{2 - \frac{8\tau^2}{\theta(\theta + 4\tau)}} > 1.$$

Note that since  $\hat{\rho} \geq \rho^{(k)}$ , then  $1 + \sqrt{1 - (\rho^{(k)})^2} \geq 1 + \sqrt{1 - \hat{\rho}^2}$  and finally  $\hat{\omega}_{opt}$  is an over estimation of  $\omega_{opt}^{(k)}$ .

**Remark 7** Among the iterative relaxation methods, the block relaxation methods such as the block Jacobi method or the block Gauss-Seidel method and the block successive overrelaxation method are suitable for the solution of the linear system derived from the anisotropic diffusion problem at each time step  $k \in \mathbb{N}$ . Generally, this kind of block iterative relaxation method has also a good asymptotic rate of convergence, particularly the block successive overrelaxation associated to the optimal relaxation parameter  $\bar{\omega}_{opt}$ . However, contrary to the classical case, i.e. for the solution of the heat conduction equation, where for example the diagonal entries of the spatial discretization matrix are equal to 4 and the off-diagonal entries are equal to -1 or zero, in the considered application of image processing, note that all the entries of the matrix  $\mathcal{A}^{(k)} = (\theta I + \alpha A^{(k)})$  are different. Then, for all blocks and at each time step of the block relaxation method, it is necessary to perform a Gaussian elimination or, in the symmetric definite positive case, a Cholesky elimination, which needs a  $\mathcal{L}\mathcal{U}$  factorization of each block-diagonal matrices. Since all block-diagonal

and all right hand sides of each linear subsystems are different, each factorization and each forward and backward substitution are time consuming. Then, even if the block relaxation method converges faster than the point relaxation method, it does not necessarily need less time of computation. Then the use of block relaxation methods for solving by a numerical way the anisotropic diffusion problem seems to be less interesting.

**Remark 8** In the previous linear algebraic systems, we have shown that the matrices  $(\theta I + \alpha A^{(k)})$  have the property to be  $M$ -matrices. In fact this property is very important, since it ensures the convergence of parallel synchronous and asynchronous Schwarz alternating method corresponding to a subdomain method with overlapping [8] and also the convergence of the subdomain method without overlapping [13]. This kind of algorithm can be very useful when a 3D speckled image is treated.

On another side, for solving at each time step  $k \in \mathbb{N}$  the linear system derived from the anisotropic diffusion problem, we can also use the classical conjugate gradient method. Classically, we know that the asymptotic rate of convergence of the classical conjugate gradient method is related to the condition number  $\kappa^{(k)}$  of the matrix  $\mathcal{A}^{(k)} = (\theta I + \alpha A^{(k)})$  as follows

$$R_\infty(CG^{(k)}) \leq 4 \left( \frac{\sqrt{\kappa^{(k)}} - 1}{\sqrt{\kappa^{(k)}} + 1} \right)^2.$$

In order to estimate the efficiency of the classical conjugate gradient method for solving at each time step  $k \in \mathbb{N}$  the linear system, it is necessary to estimate the condition number of the matrix  $(\theta I + \alpha A^{(k)})$ . Recall that when the matrix  $\mathcal{A}^{(k)}$  is symmetric,  $\kappa^{(k)}$  is given by

$$\kappa^{(k)} = \frac{\theta + \alpha \lambda_{max}^{(k)}}{\theta + \alpha \lambda_{min}^{(k)}},$$

where  $\lambda_{max}^{(k)}$  and  $\lambda_{min}^{(k)}$  are the maximal and minimal eigenvalues. Since  $\frac{1}{\theta + \alpha \lambda_{min}^{(k)}} < 1$ , then  $\kappa^{(k)} \leq \kappa_1^{(k)} = \theta + \alpha \lambda_{max}^{(k)}$  and we can estimate the condition number by the value  $\kappa_1^{(k)}$ , which, since  $\alpha$  is small, is very close to  $\theta$ ; in such a case, the preconditioning is not necessary. Note also that, since  $\lambda_{max}^{(k)} \leq \max_{1 \leq i \leq nm} \left( \sum_{j=1}^{nm} |a_{ij}^{(k)}| \right) \leq 16$ , then,  $h$  being fixed to one, we finally obtain

$$\kappa^{(k)} \leq \kappa_1^{(k)} = \theta + \alpha \lambda_{max}^{(k)} \leq \hat{\kappa}_1 = \theta + 16\alpha = \theta + 8\tau.$$

Note also that we can obtain a sharper bound of the condition number; indeed since  $\frac{1}{\theta + \alpha \lambda_{min}^{(k)}} < \frac{1}{\theta}$ , and  $\lambda_{max}^{(k)} \leq 16$  then we obtain

$$\kappa^{(k)} \leq \kappa_2^{(k)} = 1 + \frac{\alpha \lambda_{max}^{(k)}}{\theta} \leq \hat{\kappa}_2 = 1 + \frac{8\tau}{\theta} \approx 1, \tag{27}$$

when  $\tau$  is small. Note, that  $\hat{\kappa}_1 = \hat{\kappa}_2$  if  $\theta = 1$ ; if  $\theta = \frac{3}{2}$ , corresponding to the Gear scheme, then  $\hat{\kappa}_2 < \hat{\kappa}_1$ .

We summarize hereafter in Table 3 some upper bounds of  $\kappa^{(k)}$ ,  $k \in \mathbb{N}$  for different values of  $\tau$ .

**Table 3. Upper bounds of the condition number for semi-implicit and semi-implicit Gear methods**

$\tau$	Semi-implicit	Semi-implicit Gear $\hat{\kappa}_1$	Semi-implicit Gear $\hat{\kappa}_2$
$\frac{1}{5}$	2.60	3.10	2.06
$\frac{1}{7}$	2.14	2.64	1.76
$\frac{1}{10}$	1.80	2.30	1.53
$\frac{1}{20}$	1.40	1.90	1.26

**Remark 9** For comparison, consider the numerical solution of the heat diffusion equation with Dirichlet boundary condition; then, the diffusion coefficient  $c$  being constant, we have to solve

$$\begin{cases} \frac{\partial u(x,y,t)}{\partial t} - c\Delta u(x,y,t) = f(x,y,t) \text{ everywhere in } \Omega \times [0, T], \\ u(x,y,t)|_{\partial\Omega} = 0, \quad \forall t \in [0, T] \text{ (Boundary condition),} \\ u(x,y,0) = u_0(x,y), \text{ (Initial condition),} \end{cases} \quad (28)$$

where, in this classical case the spatial discretization step size  $h$  is not fixed to one. If we consider the solution of problem (28) by an implicit scheme, then the condition number of each linear system to solve at each time step  $k$ , is obviously given by

$$\kappa^{(k)} = \frac{h^2 + 4c\tau(\sin(\frac{nm\pi h}{2}))^2}{h^2 + 4c\tau(\sin(\frac{\pi h}{2}))^2} \leq 1.0 + \frac{4c\tau}{h^2}(\sin(\frac{nm\pi h}{2}))^2;$$

then

$$\kappa^{(k)} \leq 1.0 + \frac{4c\tau}{h^2} \rightarrow \infty \text{ if } h \rightarrow 0,$$

and in this case the condition number increases, on the contrary of the behavior arising in image processing where the spatial discretization step size  $h$  is fixed to one.

**Remark 10** According to the notations of Remark 6, consider that  $A^{(k)} = A_x^{(k)} + A_y^{(k)}$ . Consider also that the spatial step size  $h_x$  and  $h_y$  are different in each direction  $x$  and  $y$  and are defined as in the previous Remark 6.  $\alpha_x$  and  $\alpha_y$  being defined in a similar way, then we can write

$$\kappa^{(k)} \leq \frac{\theta + \alpha_x(\lambda_x)_{Max} + \alpha_y(\lambda_y)_{Max}}{\theta + \alpha_x(\lambda_x)_{min} + \alpha_y(\lambda_y)_{min}},$$

where  $\lambda_x$  is one eigenvalue of  $A_x^{(k)}$  and similarly for  $A_y^{(k)}$ . Then,  $\frac{1}{\theta}$  being an upper bound of the inverse of the denominator, we obtain

$$\kappa^{(k)} \leq 1 + \frac{\alpha_x(\lambda_x)_{Max}}{\theta} + \frac{\alpha_y(\lambda_y)_{Max}}{\theta},$$

and since the diffusion coefficient is bounded by one, then

$$(\lambda_x)_{Max} \leq \|A_x^{(k)}\| \leq \max_{1 \leq i \leq nm} \left( \sum_{j=1}^{nm} |(a_x)_{i,j}| \right) \leq 2(a_x)_{i,i} \leq 8,$$

and similarly for  $(\lambda_y)_{Max} \leq 8$ ; then

$$\kappa^{(k)} \leq 1 + \frac{4\tau}{\theta} \left( \frac{(n+1)^2}{d^2} + \frac{(m+1)^2}{b^2} \right) = 1 + \frac{4\tau}{\theta} \left( \frac{1}{h_x^2} + \frac{1}{h_y^2} \right).$$

In fact, the numbers  $n$  and  $m$  are in the general case sufficiently large. So, from a numerical point of view only, we can see that  $h_x$  and  $h_y$  have small values. Nevertheless, in image processing, recall again that one considers that the parameters are non-dimensional such that  $n+1 \simeq d$  and  $m+1 \simeq b$ , and, consequently the spatial discretization step size satisfies  $h_x = h_y \approx 1$  (see [27] and [37]) and the upper bound of the condition number  $\kappa^{(k)}, k \in \mathbb{N}$  is well given by (27) and, since  $\tau$  is small, the upper bound of  $\kappa^{(k)}, k \in \mathbb{N}$  is close to one.

## 6. Image Processing Results

To study the effectiveness of our method, we present experimentations on synthetic and real images. Quantitative measures have been used to compare the  $\gamma$ -diffusion results with eleven filters widely known in the literature.

### 6.1. Stopping Criterion

The coefficient of the  $\gamma$ -diffusion has the property of reaching zero when  $\gamma \geq \gamma_s$ , where  $\gamma_s$  is defined in equation (16); while most other anisotropic diffusion models only tend to zero and never stop the smoothing process. In addition, the parameter  $\gamma_s$  is reevaluated at each iteration. It constantly decreases since the global CV becomes lower due to the diminution of the speckle. This allows our diffusion to tend to a stable state, where the image becomes almost piecewise constant.

Figure 1 shows the result of the  $\gamma$ -diffusion after 80000 iterations. The image is almost piecewise constant where significant edges and structures have been preserved. Comparative results for a smaller number of iterations can be found in section 6.4..

Due to this property, the  $\gamma$ -diffusion never completely diffuse the image into one constant intensity. Thus, the stopping criterion is not critical for our model. We simply use a criterion similar to the entropy criterion [31]. The image is considered sufficiently diffused when the estimated global coefficient of variation  $\gamma_e$ , defined in equation (13) becomes lower than a threshold  $\varepsilon_\gamma$ . To avoid cases where the threshold is difficult to achieve, we stop the diffusion after a maximal period  $k_{max}$ . As such, the diffusion stops after the number of time steps  $k_S$ , which is given by:

$$k_S = \min\{k_{max}; k | \gamma_e(U^{(k)}) < \varepsilon_\gamma\}. \quad (29)$$

For our experimentations, we empirically verified that  $\varepsilon_\gamma = 0.1$  is a realistic threshold for ultrasound images.

### 6.2. Ultrasound Synthetic Image Generation

There are numerous methods in the literature to simulate ultrasound images. Bamber and Dickinson consider the radiofrequency image as the convolution between a point spread

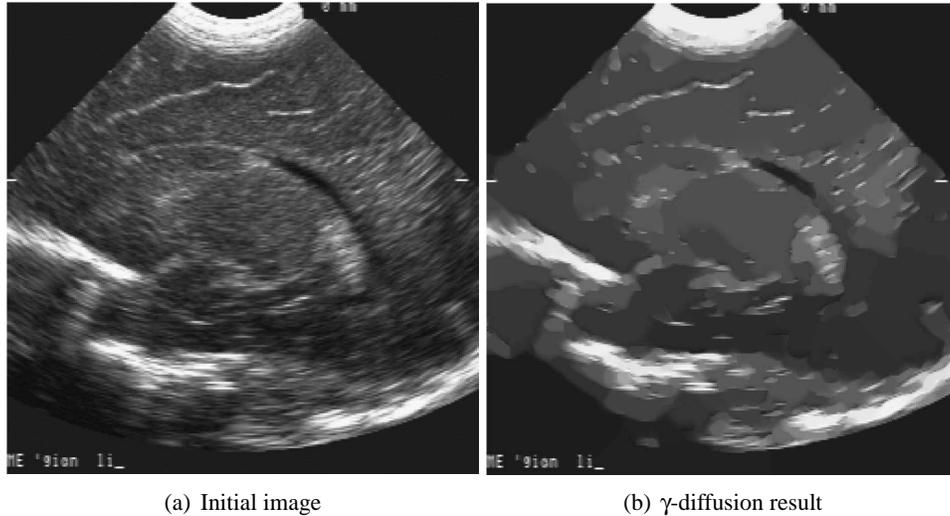


Figure 1. Convergence of  $\gamma$ -diffusion after 80000 iterations.

function (PSF) and the response of the tissue. The PSF is supposed linear and spatially invariant [3]. Meunier and Bertrand propose a similar method, but simplify the impulsional response of the tissue [23]. Dias et al. use directly the Rayleigh distribution to represent pixel intensity statistics in the radiofrequency image [11]. Jensen proposes FIELD II, a software for realistic ultrasound image simulation [15]. Jain, then Zong et al. use a multiplicative speckle model. The synthetic reflectivity is multiplied by a random variable, typically Rayleigh distributed [14, 43].

Most filters in the literature consider the distribution of Rayleigh as the speckle model. We adopted the same distribution to compare the  $\gamma$ -diffusion with other filters. The method of Dias has been used to simulate an ultrasound image affected by a speckle generated by the Rayleigh distribution (Fig. 2).

### 6.3. Criteria for Quantitative Evaluation

In order to evaluate comparatively the performance of the  $\gamma$ -diffusion, we calculated five quantitative indicators. The bias, total variance and total contrast are calculated for both types of images. A low total variance coupled with a relatively high total contrast tends to indicate a piecewise image. In addition, we calculated the Pratt Figure Of Merit (FOM) for synthetic images, where ground truth is available. Whereas, for real images we study the fluctuations of the global entropy.

#### *Pratt's Figure Of Merit*

This criterion evaluates the quality of edge preservation [28]. It naturally needs contour detection. For this purpose, we apply two Sobel filters on the original image and on the

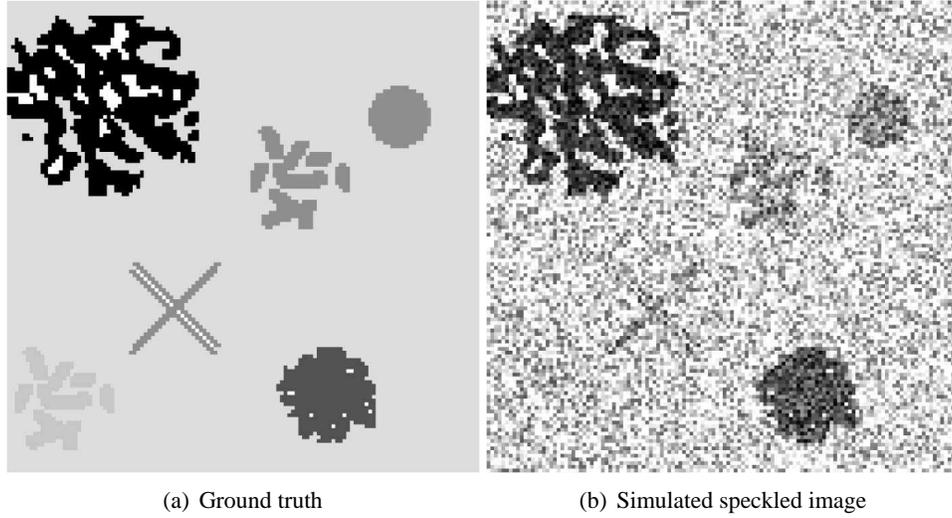


Figure 2. Ground truth and simulated speckled image.

result of each filter. The FOM can be written as

$$\text{FOM} = \frac{1}{\max(\hat{N}, N_{\text{theo}})} \sum_{i=1}^{\hat{N}} \frac{1}{1 + d_i^2 \beta} \quad (30)$$

with the following notations:

- $\hat{N}$  is the number of pixels of the detected contours ,
- $N_{\text{theo}}$  is the number of pixels of the contours of the original (noise free) image,
- $d_i$  is the euclidian distance between a contour pixel  $i \in \hat{N}$  and the nearest theoretical contour,
- $\beta$  is a parameter typically set to  $\frac{1}{9}$ .

The values of the FOM range from 0 (the lowest score) to 1 (the highest score). A value of 1 indicates that all contours have been preserved.

### Entropy

There are several interpretations of the entropy in the literature [31]. In our case, we consider that the image intensities represents the realisation of the random projections of a quantum of light over the image domain  $\Omega$ . The random event is, therefore, the location of the pixel, not its gray level as it is sometime considered. Based on this assumption, the probability for a quantum of light to reach a pixel can be written as follows:

$$p(\xi = (x_i, y_j)) = \frac{u(x_i, y_j)}{\sum_{(x_l, y_k) \in \Omega} u(x_l, y_k)} \quad (31)$$

where  $u(x_i, y_j)$  is the intensity of the pixel  $(x_i, y_j)$ . The entropy is high when the probabilities of all events have similar values; that is when the image is close to a constant intensity. The reader should be aware that this approach is the opposite of considering the probabilities of gray level. In this last case, a high entropy would denote a heterogeneous image.

#### 6.4. Comparative Results on a Synthetic Image

##### *Parameters of the Filters*

Table 4 summarizes the filters considered for the evaluation and the values used for their parameters. From the SAR literature, we chose the classical Wiener, Lee, Frost and Kuan filters. From the anisotropic diffusion literature, we chose the filters of Gauss (isotropic diffusion), Perona, Catte, Charbonnier, Weickert, Black and Yu. The homogeneous area used to calculate the global CV for the filters of Lee, Frost and Kuan is  $[[60, 91]] \times [[90, 121]]$ .

**Table 4. Filters parameters used for experimentations on synthetic images**

Method	$\tau$	Scale parameter	number of iterations
Gauss	-	$\sigma = \sqrt{50}$	1
Wiener [39]	-	-	1
Lee [20]	-	-	1
Frost [12]	-	-	1
Kuan [19]	-	-	1
Perona [27]	0.05	K=24	500
Catte [6]	0.05	K=5	500
Charb. [7]	0.05	K=8	500
Weickert [38]	0.05	K=24	500
Black [4]	0.05	K=auto	500
Yu [41]	0.05	-	500
$\gamma$ -diff.	0.05	$\gamma_s$ =auto	500

##### *Evaluation Results*

Looking at the results (fig.3), we can make the following observations:

- all filters smooth the light stains on the bottom left,
- the Gaussian filter as expected over-smooths the image,
- the classical SAR filters stop filtering when the local CV is above a threshold. Homogeneous regions are smoothed but pixels near boundaries are left unfiltered,
- the results of the classical anisotropic diffusion filters are dependent on the scale parameter: when it is too low, the noise is not filtered; and when it is too high, the region boundaries are over-smoothed,

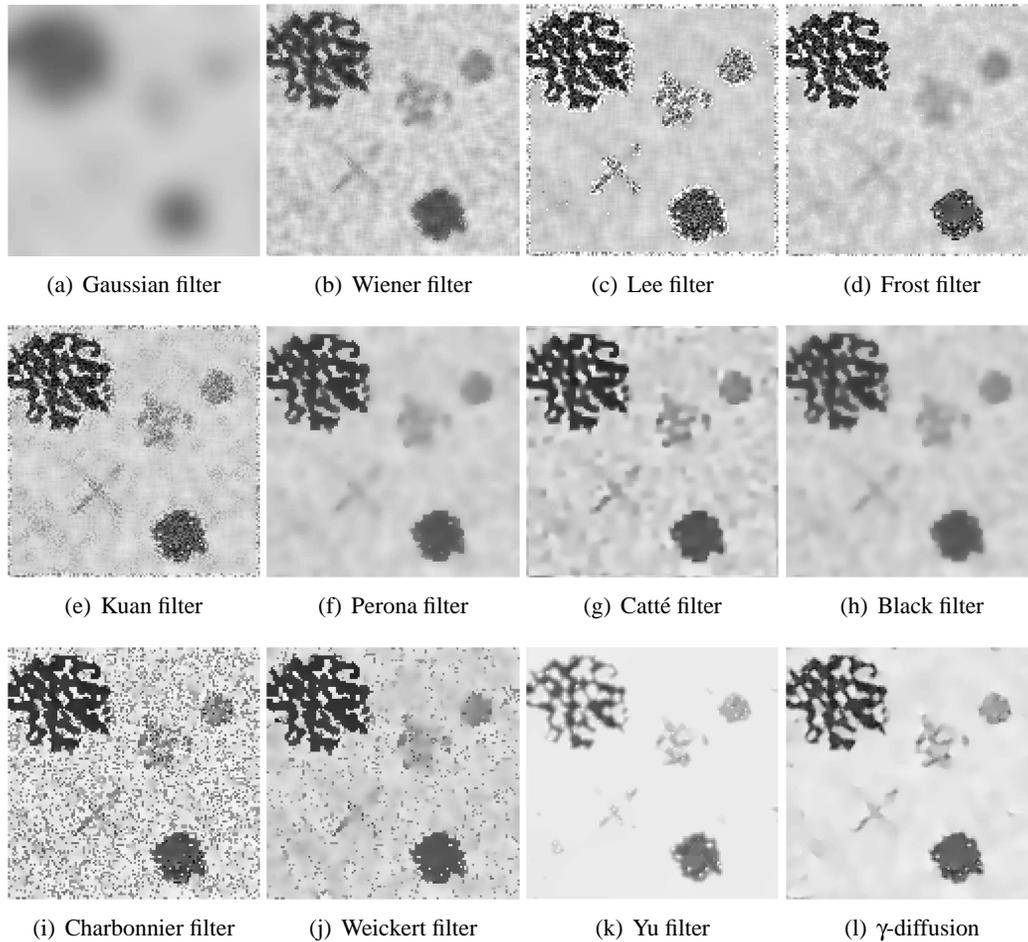


Figure 3. Results of the twelve filters on the synthetic image.

- the filters of Black et al. and Weickert et al. leave too much noise unfiltered. This is due to the fast decrease of their coefficient of diffusion function, and to the fact that the gradient is not adapted to filtering the speckle,
- the filter of Yu et al. gives better results, notably in homogeneous areas. However, the boundaries are smoothed because the coefficient of diffusion never reaches zero. The size of dark features also shrinks slightly,
- the  $\gamma$ -diffusion preserves the features of the image. The cross shaped feature is partly recovered, and all the other features are correctly filtered. The size of the dark and light features are both unchanged.

Table 5 shows the quantitative results. The bias is generally low except for the filter of Yu et al. [41], because it does not respect the maximum-minimum principle. This is especially the case when the local CV is lower than the global CV. The coefficient of diffusion would reach high values and biases the result. Our filter ensures global energy conservation.

**Table 5. Quantitative results: Bias, Total Variation ( $T_V$ ), Total Contrast ( $T_C$ ), and Figure Of Merit (FOM)**

Method	Bias	$T_V (\times 10^3)$	$T_C^2 (\times 10^6)$	FOM
Noisy image	0	4.7156	52.717	-
Gauss	0.4171	0.8490	1.0244	0.0031
Wiener	0.1068	1.8745	5.6486	0.6272
Lee	0.4217	2.9839	24.274	0.7172
Frost	0.3792	2.5352	16.072	0.6729
Kuan	0.9102	2.4155	15.184	0.6459
Perona	0.0097	1.8529	5.5846	0.7412
Catte	0.0085	2.1587	7.7872	0.6996
Charbonnier	0.0093	1.6549	2.8440	0.8470
Weickert	0.0093	2.5553	16.061	0.5330
Black	0.0057	3.8705	38.051	0.4233
Yu	3.1379	1.0426	3.4374	0.8776
$\gamma$ -diffusion	0.3870	2.0124	7.0718	0.9126

The total variance is low for the filter of Yu et al., that gives smoother homogeneous regions. This filter largely diffuses the image, but over-smooth the boundaries. This is confirmed by the low total contrast obtained by the filters of Yu, Gauss, and Charbonnier [7]. The filter of Black et al. [4] gets the highest contrast because it stops the diffusion early due to a low scale parameter. The general goal of an image restoration algorithm should be to reduce the variance in homogeneous regions while preserving the edges (i.e. maintaining some contrast). The  $\gamma$ -diffusion obtains a low variance and a relatively high contrast. Moreover, it gets the highest FOM, indicating better edge detection and preservation.

Figure 4 shows normalized data extracted from column 80 of the original and the filtered synthetic images. The curve corresponding to the Gaussian filter is totally different from the original image. The filters of Lee [20], Frost [12] and Kuan [19] recover approximatively the homogeneous regions. Lee and Kuan filters leave the two wells unchanged. The filter of Frost smooths the first well. Wiener smoothing of homogeneous regions is weak, although it smooths some peaks. Perona [27], Catte [6] and Charbonnier [7] filters over-smooth the data, even though they preserve boundaries. The filter of Weickert [38] leads to a good result in the second well, but generates outlier peaks. The filter of Black et al. [4] underestimates the scale parameter and detects too many peaks. The result of the filter of Yu et al. [41] is good, but a peak in the second well has been smoothed. The  $\gamma$ -diffusion leads to a curve very similar to the original. The peaks and the homogeneous parts of the second well are all detected.

## 6.5. Comparative Results on Real Ultrasound Images

In this section, the  $\gamma$ -diffusion is applied to two real ultrasound images and compared with a set of other filters (Fig. 5, 6), namely Lee from SAR domain, Perona et al. as

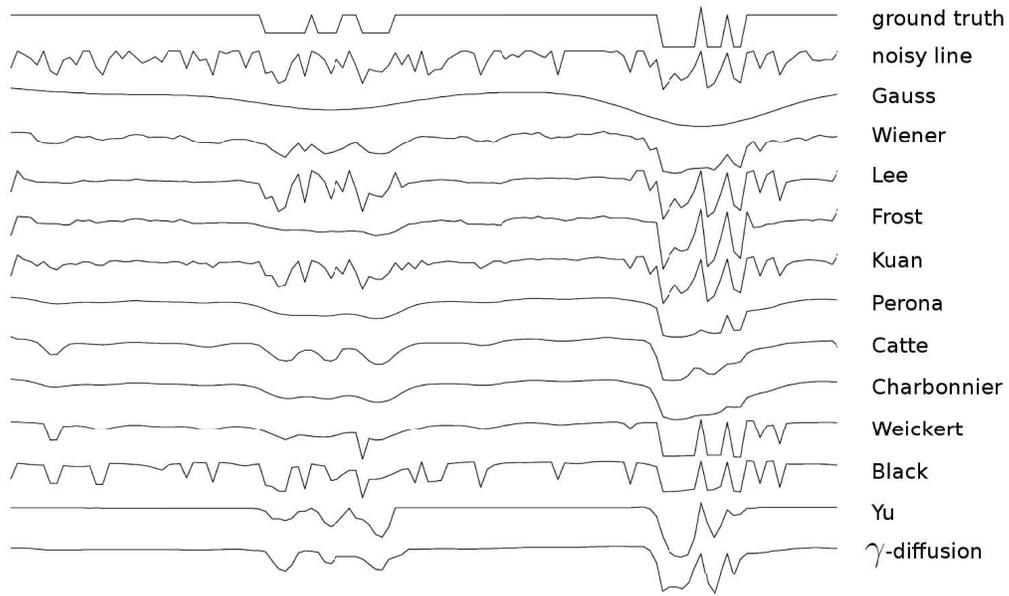


Figure 4. Normalized results of various filters on column 80 of the synthetic image.

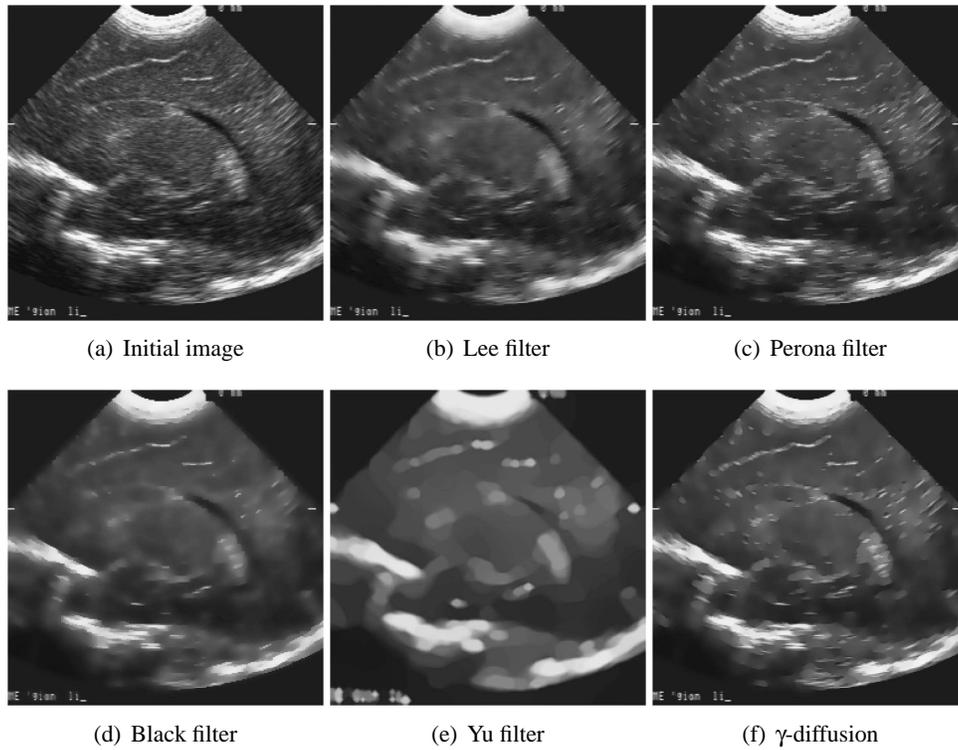


Figure 5. Results of Lee, Perona, Black, Yu and  $\gamma$ -diffusion filters on US image.

a classical anisotropic diffusion, Black et al. for its robustness and automatic parameter estimation, and Yu et al. as it is tailored to filter the speckle. The first image (Fig. 5) is a neonatal ultrasound image, while the second (Fig. 6) is an echocardiographic image featuring the right auricle. Results are similar on both images. The filters of Lee and Perona et al. smooth homogeneous areas but fail to preserve edges. The filter of Black et al. doesn't smooth contours and leave a large part of the noise. The filter of Yu et al. does result in a smooth image, but some important contours have disappeared or widened. The  $\gamma$ -diffusion leads to an almost piecewise constant image, preserving important contours of the initial image. To compare the results, we extracted the edges from the images of figure 6. The corresponding contours are presented in figure 7. The Lee and Black et al. filters preserve most initial contours, but they also conserve most of the noise. Perona smooths the homogeneous regions, but smooths the edges as well. The filter of Yu et al. leads to a shrunk right auricle, while high intensity regions have widened. The result of the  $\gamma$ -diffusion is sharp on the edges and smooth in homogeneous areas; the size of the regions has not changed. Table 6 shows the quantitative results of the compared filters. The values of the bias are similar to those found with the synthetic image. The values of the entropy confirm the visual observation: the  $\gamma$ -diffusion has the lowest entropy followed by the filter of Black et al.. The smoothest result corresponds to Perona filter. The total variation is biased because of large black regions present in the initial image.

**Table 6. Quantitative measures of filters results on US image**

Method	Bias	$T_V (\times 10^3)$	$T_C^2 (\times 10^6)$	$\Delta$ Entropy
Initial im.	0	2.1596	19.308	10.8902
Lee	0.2335	1.8799	9.3782	2,10
Perona	0.0226	1.8741	8.6711	2,67
Black	0.0286	2.0805	15.932	0,95
Yu	2.4367	2.4638	5.7362	2,37
$\gamma$ -diffusion	0.2066	2.1527	13.059	1,61

## 7. Numerical Experiments

### 7.1. Stability Validation

Sequential numerical experiments have been performed on two kinds of ultrasound medical images. The first one called in the sequel image A has  $116 \times 154$  pixels and the size of the system to solve is 17328. It is an ultrasound image of the left ventricle and auricle of a human heart (see figure 8a). The second image is a synthetic ultrasound image called image B, that has  $1024 \times 1024$  pixels, and leads to a system of size 1 044 484.

The global objectives enumerated in section 1. were reached. In particular the images presented in Figure 8b, in Figure 8c and in Figure 8d show that the processing reduces the variance in homogeneous regions and increases the contrast. Figure 8b corresponds to a

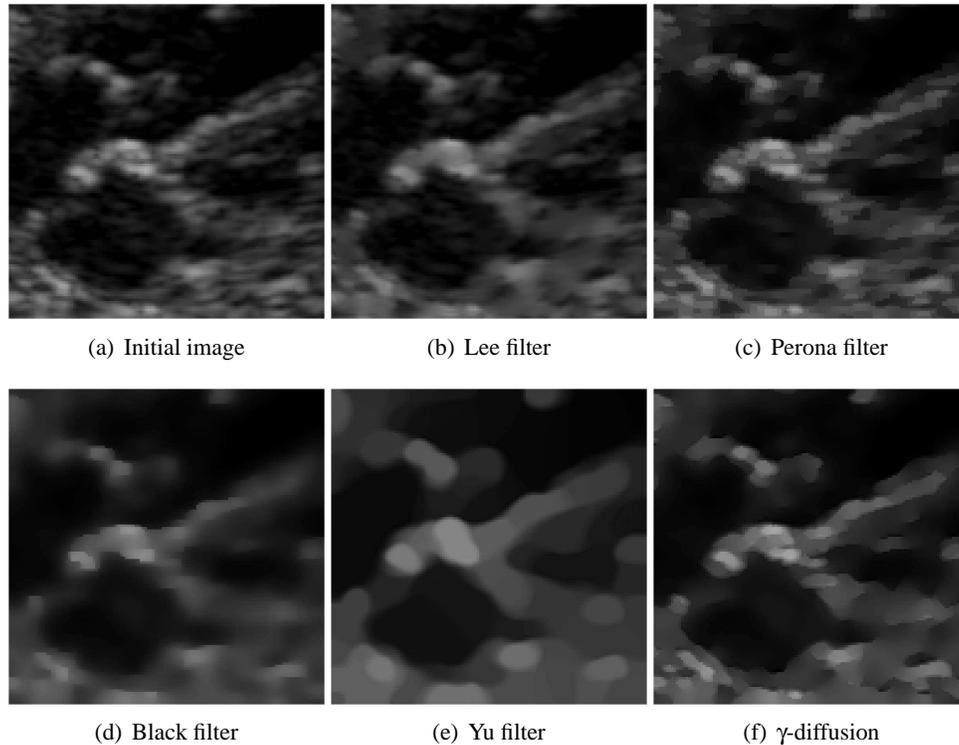


Figure 6. Results of Lee, Perona, Black, Yu and  $\gamma$ -diffusion filters on cardiac US image of a 12 weeks old fetus.

diffusion for one second with fifty time steps while figures 8c and 8d present the result after a diffusion of twenty five seconds with respectively twenty five and one thousand time steps. More precisely, we can note that the speckle is progressively reduced in homogeneous regions, while the boundaries between the different parts of the heart are more perceivable. Moreover, due to the modelization of the diffusion coefficient, infinite diffusion does not give uniform images, but instead tends to images similar to the ones obtained in Figure 8c and in Figure 8d.

Finally, Figure 8e corresponds to the same image processing filter using an instable explicit time marching scheme. It can be noted from Figure 8c and from Figure 8d that the semi-implicit scheme remains stable for time step  $\tau = 1$  second, which is not the case for the explicit scheme (see Figure 8e).

## 7.2. Results of Sequential Numerical Experiments

Sequential simulations have been performed on a 2.4 GHz Apple IMac with processor Intel core 2 duo. The results of sequential experiments are summarized in Figures 9 to 14 and in Tables 7 to 10.

For the two considered images, Tables 7 and 8 give the user CPU elapsed time for explicit scheme, semi-implicit scheme and Gear semi-implicit scheme, when the linear sys-

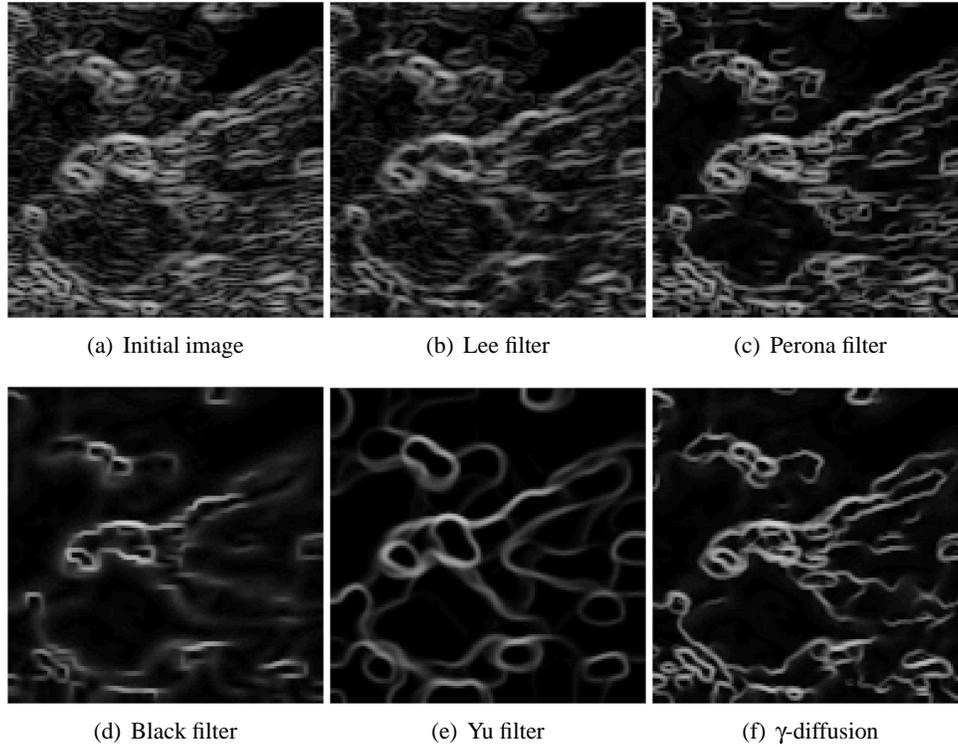


Figure 7. Contour detection from figure 6.

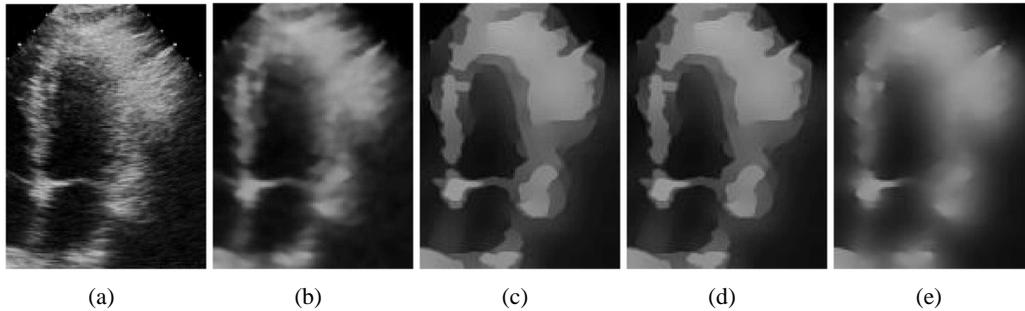


Figure 8. Image processing results for several schemes and parameters.

tems are solved by relaxation algorithms and conjugate gradient algorithm. In Figure 9, note that the conjugate gradient method has the best rate of convergence; as a consequence this last method allows to obtain the lowest elapsed time for the iterative solution of the linear system at each time step  $k \in \mathbb{N}$ .

Nevertheless, the difference of performance between the compared iterative methods is not considerable. This is due to the fact that the cost of computation of the diffusion coefficient and of the entries of the matrix  $A^{(k)}$  at each time step is important. When the conjugate gradient method is used, we can see that at each time step, the same time is

necessary for solving the linear system, and for computing the diffusion coefficient and the entries of the matrix  $A^{(k)}$ . Moreover, in the case of using the semi-implicit scheme, and when the linear systems are solved by the optimized successive over relaxation method, the total difference of elapsed time is about 12.80% greater than when we use the conjugate gradient algorithm. If we used the Gear semi-implicit scheme, the difference with the total of elapsed time is less (about 11,23%); in this latter situation, the diagonal dominance of the matrix to inverse, i.e.  $(\theta I + \alpha A^{(k)})$ , is more important.

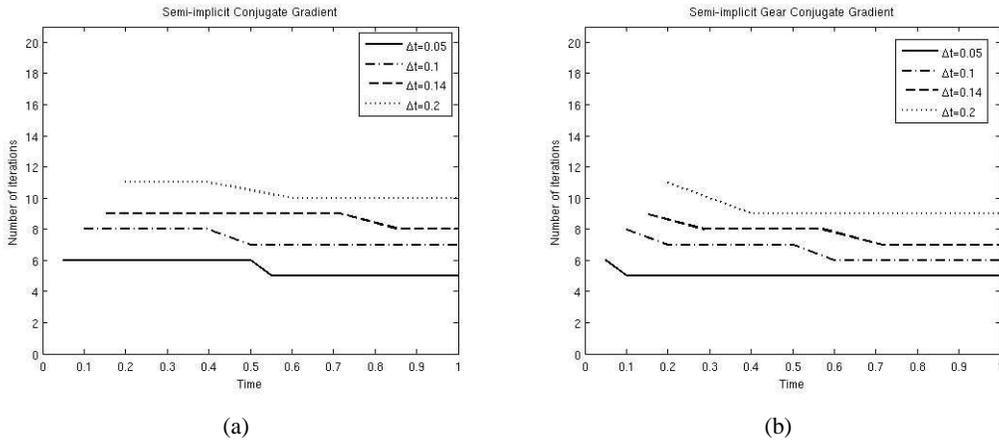


Figure 9. Number of iterations with gradient conjugate method. (a) Semi-implicit scheme (b) Gear semi-implicit scheme.

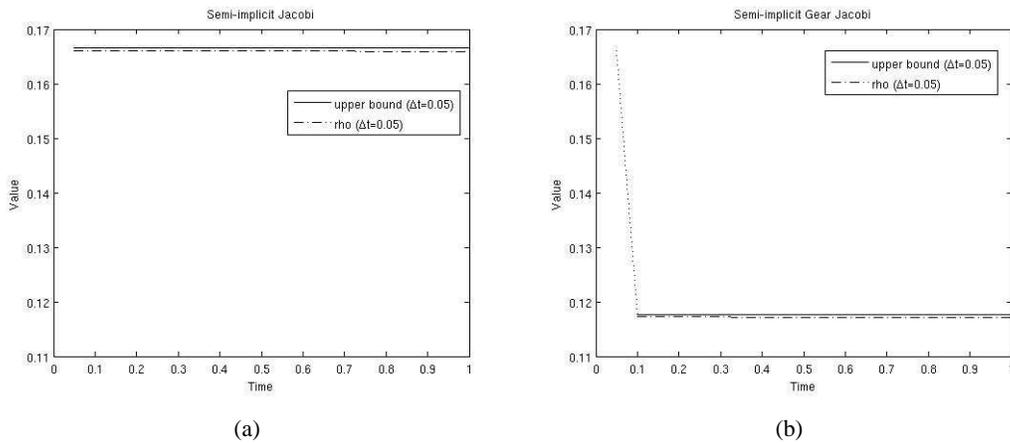


Figure 10. Spectral radius of Jacobi matrix. (a) Semi-implicit scheme (b) Gear semi-implicit scheme.

For all used iterative methods, i.e. Jacobi, optimized S.O.R. and conjugate gradient methods, and for the two sizes of images considered, the computational times are low. Table 9, shows the difference in computational time and in percentage of the total computational time between the optimized S.O.R. and conjugate gradient methods.

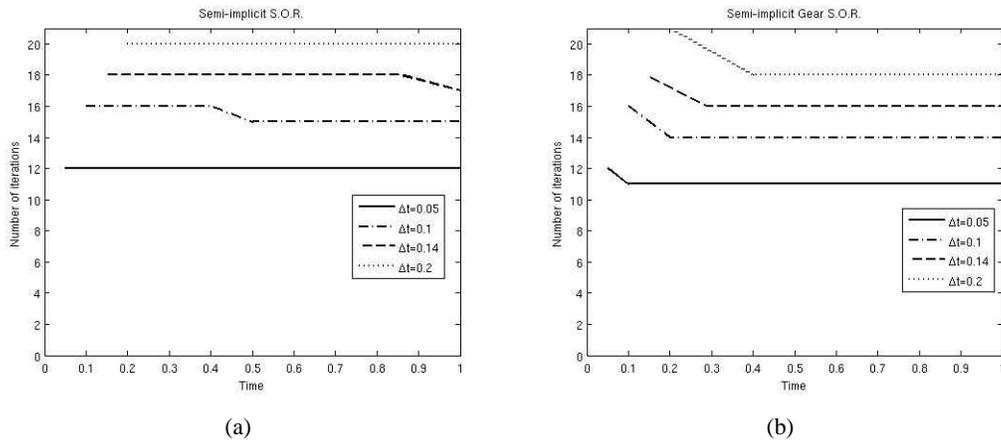


Figure 11. Number of iterations with S.O.R. method. (a) Semi-implicit scheme (b) Gear semi-implicit scheme.

For the explicit scheme, we obtain the best elapsed time; however, this scheme is unstable if  $\tau$  is too large.

**Table 7. C.P.U. time (sec.) for computation of image A**

$\tau$	Explicit	Semi-implicit			Gear semi-implicit		
		Jacobi	S.O.R.	Conjugate	Jacobi	S.O.R.	Conjugate
$\frac{1}{5}$	0.139	0.246	0.200	0.187	0.231	0.192	0.183
$\frac{1}{7}$	0.157	0.281	0.281	0.217	0.263	0.224	0.211
$\frac{1}{10}$	0.186	0.331	0.280	0.257	0.309	0.268	0.251
$\frac{1}{20}$	0.283	0.488	0.425	0.390	0.458	0.412	0.384

**Table 8. C.P.U. time (sec.) for computation of image B**

$\tau$	Explicit	Semi-implicit			Gear semi-implicit		
		Jacobi	S.O.R.	Conjugate	Jacobi	S.O.R.	Conjugate
$\frac{1}{5}$	11.181	17.500	13.754	12.419	16.235	13.290	12.222
$\frac{1}{7}$	13.560	20.193	16.483	14.612	18.623	15.819	14.328
$\frac{1}{10}$	17.219	24.429	20.089	17.858	22.696	19.313	17.564
$\frac{1}{20}$	29.144	37.194	31.721	28.162	34.483	30.973	27.845

Table 10 shows in the case of semi-implicit scheme and for the image B, the number of zero rows in the matrix  $A^{(k)}$ , during the evolution process. At the first time step, 0.32% of the rows of the matrix  $A^{(1)}$  are zero; at the last time step, when  $k = 10$ , 2.17% of the rows of the matrix  $A^{(10)}$  are zero. Then, Table 10 shows clearly that the matrix  $A^{(k)}, \forall k \in \mathbb{N}$ , are

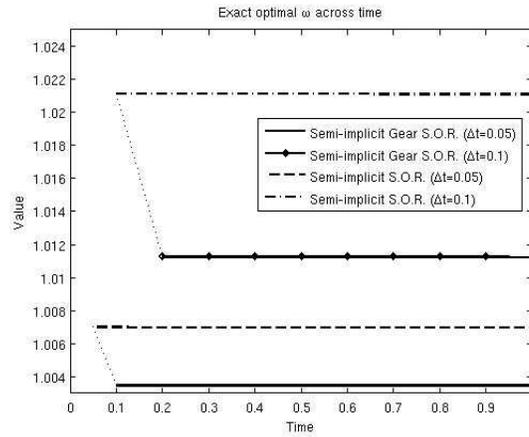


Figure 12. Evolution of optimal relaxation value for Semi-implicit scheme and Gear semi-implicit scheme.

**Table 9. Difference and corresponding percentage of elapsed time for optimized S.O.R. and conjugate gradient methods for image B**

$\tau$	Semi-implicit method		Semi-implicit Gear method	
	Difference	Percentage	Difference	Percentage
$\frac{1}{5}$	1.335	10.75 %	1.068	8.74 %
$\frac{1}{7}$	1.871	12.80 %	1.491	10.40 %
$\frac{1}{10}$	2.231	12.49 %	1.749	9.95 %
$\frac{1}{20}$	3.559	12.63 %	3.128	11.23 %

reducible, this reducibility increases during the computational process.

**Table 10. Percentage of zero rows for the matrix  $A^{(k)}$  for image B during 10 time steps**

Time step n	Number of zero rows	Percentage of zero rows
1	3 433	0.32 %
2	4 750	0.45 %
3	6 140	0.58 %
4	7 790	0.74 %
5	9 670	0.92 %
6	11 932	1.14 %
7	14 482	1.38 %
8	17 229	1.65 %
9	19 972	1.91 %
10	22 729	2.17 %

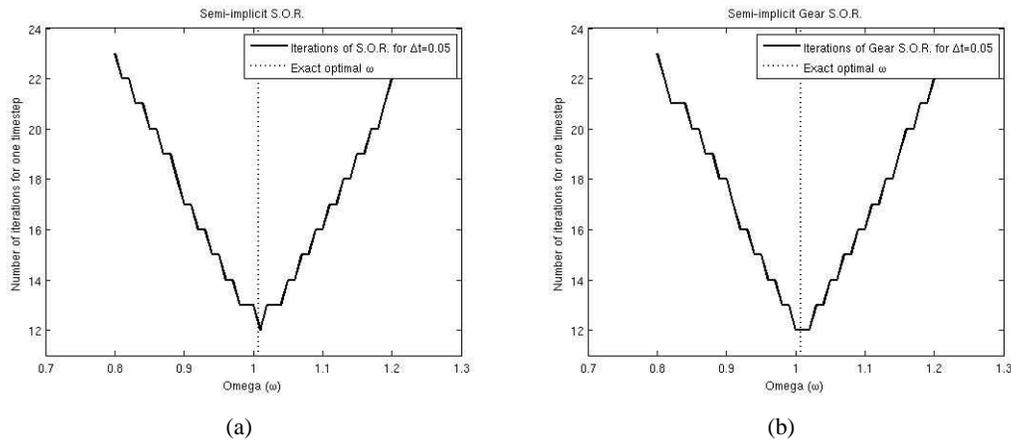


Figure 13. Experimental value of optimal relaxation (a) Semi-implicit scheme (b) Gear semi-implicit scheme.

Figures 10a and 10b show the evolution of the spectral radius  $\rho^{(k)}$  of the Jacobi matrix for both semi-implicit methods; the dashed line represents the exact value of  $\rho^{(k)}$  computed during the evolution process using the power method while the solid line represents the evolution of an estimate of  $\|g^{(k)}\|_\infty$ . It can be noted that, during the evolution process,  $\rho^{(k)}$  is practically constant. Moreover the estimate of  $\rho^{(k)}$  by  $\|g^{(k)}\|_\infty$  and the global upper bound (25) of  $\rho^{(k)}$  are very accurate (see Table 2). Furthermore, since  $\rho^{(k)}$  is slightly constant at each time step  $k$ , the same number of iterations of the Jacobi method is necessary for solving the linear systems at each time step; this has been observed by numerical experiments. As a consequence, since the spectral radius of the point Gauss-Seidel matrix is the square of the spectral radius of the Jacobi matrix, and since the optimal value of the optimal relaxation parameter is subordinated to the value of  $\rho^{(k)}$ , then at each time step, the same number of iterations of the point Gauss-Seidel method and of the point S.O.R method is necessary for solving the linear systems. The same remark holds when using the conjugate gradient method to solve the linear systems in both semi-implicit schemes (see Figure 11).

When the Gear semi-implicit scheme is used, it can be noted in Figures 9b, 10b and 11b, that between the first time step and the second one, a discontinuity appears in the number of iterations during the evolution process. This discontinuity is due to the fact that for the first step of the Gear semi-implicit scheme, the semi-implicit scheme is used.

Figures 13a and 13b show the influence of the value of the relaxation parameter on the number of iterations of the optimized S.O.R method at the first time step. We observe that we have the same behavior of the S.O.R optimized method for all time steps. It can be noted that the optimal value of the relaxation parameter is near one according to the estimate (27). Furthermore, during the time marching algorithm, Figure 12 shows that, in accordance with the fact that during the evolution process the spectral radius of the Jacobi matrix is practically constant, the optimal value of the relaxation parameter is constant at each time step for both semi-implicit scheme and Gear semi-implicit scheme.

Finally, Figure 14 shows that, in this particular application of image processing, the SSOR preconditioned conjugate gradient method does not improve the asymptotic rate of

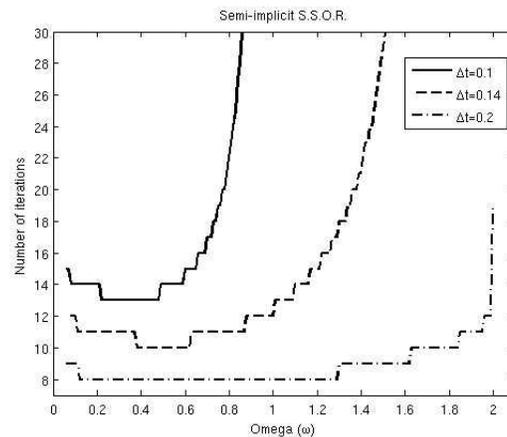


Figure 14. Experimental values of optimal relaxation for S.S.O.R. preconditioning for Semi-implicit scheme.

convergence of the conjugate gradient method (see [24] and [30]).

## 8. Conclusion

We presented a novel robust anisotropic diffusion model adapted to speckle filtering in ultrasound images. Traditional filters are based on the gradient and well-suited for additive noise, but not for multiplicative noise such as the speckle. Our model is based on the local coefficient of variation and the Tukey error norm. It is therefore inherently tailored to multiplicative noise. We also developed a robust estimator to compute the scale parameter of the model. This parameter is reevaluated at each iteration allowing a precise control of the diffusion. The classification of pixels in homogeneous areas or edges is refined iteratively.

The experimentations showed that our filter performs well compared to other methods from the literature, and that stability of semi-implicit scheme is ensured, given a proper parameterization of the time step. The solution of the linear algebraic numerical systems with conjugate gradient method was robust, as confirmed by the experimentations; note also that the use of the point relaxation methods is also suitable for our application.

## References

- [1] G. Aubert and P. Kornprobst, Mathematical problems in image processing - partial differential equations and the calculus of variations, *Applied Mathematical Sciences* 147 (2006).
- [2] O. Axelsson, Iterative solution methods, *Cambridge university press* (1996).
- [3] J.C. Bamber and R.J. Dickinson, Ultrasonic b-scanning: a computer simulation, *Phys. Med. Biol* **25** (1980), 463–479.

- 
- [4] M.J. Black, G. Sapiro, D.H. Marimont, and D. Heeger, Robust anisotropic diffusion, *IEEE Trans. on Image Proc.* (1998), 421–432.
  - [5] L. Blanc-Feraud, P. Charbonnier, Aubert G., and M. Barlaud, Nonlinear image processing: modeling and fast algorithm for regularization with edge detection, *Image Processing, Proceedings., International Conference on* **1** (1995), 474–477.
  - [6] F. Catte, P.L. Lions, J.M. Morel, and T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM Journal on Numerical Analysis* **29** (1992), 182–193.
  - [7] P. Charbonnier, L. Blanc-Feraud, Aubert G., and M. Barlaud, Deterministic edge-preserving regularization in computed imaging, *IEEE Transactions on Image Processing* **6** (1997), 298–311.
  - [8] M. Chau, C. Tauber, and P. Spiteri, Parallel schwarz alternating methods for anisotropic diffusion of speckled medical images, *Numerical Algorithms* **51** (2009), 85–114.
  - [9] R.N. Czerwinski, D.L. Jones, and W.D. O’Brien, Detection of lines and boundaries in speckle images-application to medical ultrasound, *IEEE Trans. on Medical Imaging* **18** (1999), 126–136.
  - [10] R. Dautray and J.L. Lions, Analyse mathématique et calcul numérique pour les sciences et les techniques, *Masson* **9** (1988).
  - [11] J.M. Dias and J.M. Leitaó, Wall position and thickness estimation from sequences of echocardiographic images, *Medical Imaging, IEEE Transactions on* **15** (1996), 25–38.
  - [12] V.S. Frost, J.A. Stiles, K.S. Shanmugan, and J.C. Holtzman, A model for radar images and its application to adaptive digital filtering of multiplicative noise, *IEEE Trans. Pattern Anal. Machine Intell.* **4** (1982), 157–166.
  - [13] L. Giraud and P. Spiteri, Resolution parallele de problemes aux limites non-lineaires(in french), *Modelisation Mathematique et Analyse Numerique* **25** (1991), 579–606.
  - [14] A.K. Jain, Fundamentals of digital image processing, *Englewood Cliffs, NJ: Prentice-Hall* (1989).
  - [15] J.A. Jensen, Field: A program for simulating ultrasound systems, *Medical and Biological Engineering and Computing* **34** (1996), 351–353.
  - [16] K. Krissian, Multiscale analysis: Applications to medical imaging and 3d vessel detection, *PhD Thesis, University of Nice Sophia-Antipolis* (2000).
  - [17] Karl Krissian, Kirby Vosburgh, Ron Kikinis, and Carl-Fredrik Westin, Speckle-constrained anisotropic diffusion for ultrasound images, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition ('05)* (San Diego CA, USA), June 2005, pp. 547 – 552.

- 
- [18] D.T. Kuan, A.A. Sawchuk, T.C. Strand, and P. Chavel, Adaptive noise smoothing filter for images with signal dependant noise, *IEEE Trans. Pattern Analysis and Machine Intelligence* **7** (1985), 165–177.
- [19] ———, Adaptive restoration of images with speckle, *IEEE Trans. Acoust., Speech, Signal Processing* **35** (1987), 373–383.
- [20] J.S. Lee, Digital image enhancement and noise filtering by use of local statistics, *IEEE Trans. Pattern Analysis and Machine Intelligence* **2** (1980), 165–168.
- [21] A. Lopes, E. Nezry, R. Touzi, and H. Laur, Structure detection and statistical adaptive speckle filtering in sar images, *Int. J. Remote Sensing* **14** (1993), 1735–1758.
- [22] A. Lopes, R. Touzi, and E. Nezry, Adaptive speckle filters and scene heterogeneity, *IEEE Trans. Geoscience and Remote Sensing* **28** (1990), 992–1000.
- [23] J. Meunier and M. Bertrand, Ultrasonic texture motion analysis: theory and simulation, *Medical Imaging, IEEE Transactions on* **14** (1995), 293–300.
- [24] J.C. Miellou and P. Spiteri, Optimization of the relaxation parameter for s.s.o.r. and a.d.i. preconditioning, *Numerical Algorithm* **29** (2002), 155–195.
- [25] J. Montagnat, M. Sermesant, H. Delingette, G. Malandain, and N. Ayache, Anisotropic filtering for model-based segmentation of 4d cylindrical echocardiographic images, *Pattern Recognition Letters - Special Issue on Ultrasonic Image Processing and Analysis* **24** (2003), 815–828.
- [26] A.V. Oppenheim and R.W. Schaffer, Discrete-time signal processing, *Englewood Cliffs, NJ: Prentice Hall* (1989).
- [27] P. Perona and J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.* **12** (1990), 629–639.
- [28] W.K. Pratt, Digital image processing, *New York: Wiley* (1977).
- [29] P.J. Rousseeuw and A.M. Leroy, Regression and outlier detection, *New York: Wiley* (1987).
- [30] P. Spiteri, J.C. Miellou, and J. Bahi, Evaluations of parameters for the optimization of s.s.o.r. and a.d.i. preconditioning, *Numerical Algorithm* **29** (2002), 249–265.
- [31] J. Sporring and J. Weickert, Information measures in scale-spaces, *IEEE Transactions on Information Theory* **45** (1999), 1051–1058.
- [32] C. Tauber, Filtrage anisotrope robuste et segmentation par b-spline snake: application aux images échographiques, *PhD Thesis, Institut National Polytechnique de Toulouse* (2005).
- [33] C. Tauber, P. Spiteri, and H. Batatia, Iterative methods for anisotropic diffusion of speckled medical images, *Applied Numerical Mathematics* **60** (2010), 1115–1130.

- 
- [34] T. Taxt, Restoration of medical ultrasound images using two-dimensional homomorphic deconvolution, *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on* **42** (1995), 543–554.
- [35] ———, Three-dimensional blind deconvolution of ultrasound images, *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on* **48** (2001), 867–871.
- [36] J.W. Tukey, Exploratory data analysis, *Addison-Wesley, Reading, MA* (1977).
- [37] J. Weickert, Anisotropic diffusion in image processing, *Ph.D. thesis, Dept. of Mathematics, University of Kaiserslautern, Germany* (1996).
- [38] J. Weickert, B.M. Romeny, and M.A. Viergever, Efficient and reliable schemes for nonlinear diffusion filtering, *IEEE Transactions on Image Processing* **7** (1998), 398–410.
- [39] N. Wiener, Extrapolation, interpolation, and smoothing of stationary time series., *New York: Wiley. ISBN 0-262-73005-7* (1949).
- [40] Y.L. You, W. Xu, A. Tannenbaum, and M. Kaveh, Behavioral analysis of anisotropic diffusion in image processing, *IEEE Trans. On Image Proc.* **5** (1996), 1539–1553.
- [41] Y. Yu and S.T. Acton, Speckle reducing anisotropic diffusion, *IEEE Trans. on Image Proc.* **11** (2002), 1260–1270.
- [42] ———, Edge detection in ultrasound imagery using the instantaneous coefficient of variation, *IEEE Trans. on Image Proc.* **13** (2004), 1640–1655.
- [43] E. Zong, A.F. Laine, and E.A. Geiser, Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing, *IEEE Trans. on Medical Imaging* **17** (1998), 532–540.