Denoising textured images via regularized anisotropic diffusion

Denoising Textured Images via Regularized Anisotropic Diffusion

Wallace Casaca, Marcos Proença de Almeida
Institute of Mathematics and Computer Science
University of São Paulo, Brazil

Maurílio Boaventura
Institute of Biosciences, Humanities and Exact Sciences
São Paulo State University, Brazil

1 Introduction

A quite relevant subject in the image processing field is *image denoising*. The fundamental idea of denoising is to recover an image damaged by noise, which was originally degraded by means of some interference or deterioration process. In this context, the notion of noise may be understood as a modification of color intensity of the pixels which are scattered throughout the image according to a specific mathematical formation law. The noise modifies the image appearance, reducing or even eliminating parts of the visual information. Applications involving denoising such as *image transmission*, *image acquisition*, and *image information processing* have been increasingly employed in practical problems encompassing photographs, realistic images, forensic and medical images, security recording systems, etc. In practice, we can say that the ideal situation of a noiseless image rarely occurs.

The major causes of image noise contamination usually occur during the transmission and acquisition steps, since almost all image manipulation devices are subject to environmental or technical limitations. The noise due to effects of magnetic fields when a telescope in space captures images and sends them to Earth and the contamination produced during the acquisition of a nanotechnology image are classical examples of environmental and technical constraints, respectively.

From a mathematical point of view, it is possible to model almost all types of image noise degradations by employing a simple additive model of degeneration as suggested by the image of Figure 1. Due to the existence of different classes of noise, the function $n(x)$ is numerically defined as a random variable, similar to approaches found in statistical applications. Thus, we can use *probability density functions* to describe the behavior of noise on image quality. For instance, the classical Gaussian (or Normal) probability distribution is employed to model the effects of *Gaussian noise* while the Bernoulli distribution is applied to estimate the *impulsive (or "salt and pepper") noise*. Gaussian noise is common in images acquired by cameras and telescopes and it modifies all the pixels of the image while the impulsive noise appears in applications comprehending data transmission and scanning errors. In general, these two noise types are the most studied in the literature on account of the frequency with which they occur in practical situations.

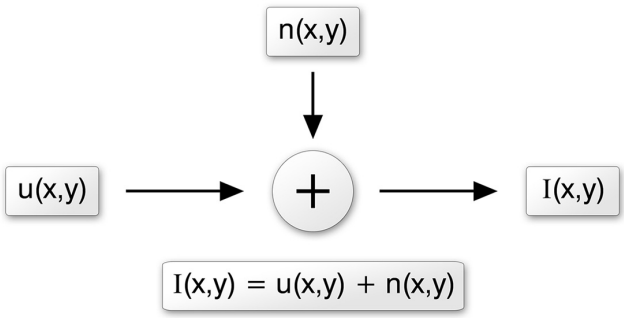


Figure 1: Degradation mathematical model for additive noise. The image $u(x,y)$ represents the original image, $n(x,y)$ the noise function and $I(x,y)$ the target signal to be restored.

Gaussian and impulsive noise impacts in a real image can be seen in Figure 2. While Figure 2(c) presents all the pixels damaged by the Gaussian noise (see the respective plot in RGB space), only a few pixels in Figure 2(e) were affected by the impulsive noise. The distinction between these two classes of noise leads to the use of specific approaches for denoising an image. Therefore, this chapter brings the

noise removal problem from the *perspective of Gaussian noise*, since it appears more frequently in all kinds of applications encompassing digital images. Moreover, according to the *classical probability theory*, the Gaussian distribution is often used to model unknown noise, by reason of the *Central Limit Theorem* (Ross, 2009), which ensures that the distribution of the sum of a large number of independent random variables (noise) tends to a Normal distribution.

Mathematically, many of the noise removal methods assume that the input image is a function of bounded variation which has no texture, while the desired image is sectionally smooth. On the other hand, the noise is usually characterized by a high-frequency oscillation function. Hence, in this case, a good denoising model should be capable of repairing a noisy image and preserving its main features like edges, boundaries and contrast.

To deal with this issue, interesting approaches based on nonlinear modeling have been proposed in the literature such as *anisotropic diffusion* (Perona & Malik, 1990; Nordstrom, 1991; Alvarez *et al*., 1992), *variational methods* (Rudin *et al*., 1992; Chan *et al*., 2001), *iterative regularization* (Osher *et al*., 2005) and *nonlinear inverse scale methods* (Burger *et al*., 2006). Nevertheless, in many cases these approaches may not be appropriate for treating highly textured images because those images have a strong presence of contrast information such as textures, fine details and complex patterns that are mixed with noise and must be distinguished during the recovery stage. Broadly speaking, images of these categories are characterized by incorporating distinct types of information: they are simultaneously defined by geometric shapes (edges, corners and homogenous structures) and oscillatory pieces (textures, high-frequency details and, obviously, noise). Figure 3 displays an example of this kind of image and its corrupted version by adding noise.

In order to treat each group of images with specific features in a unified way, there is a growing tendency of studies that consider a combination of already mentioned approaches allied to recent techniques that rely on mathematical theory of harmonic analysis which include *curvelets* (Starck *et al*., 2002) and *wave atoms* (Demanet & Ying, 2007a; Demanet & Ying, 2007b). Over the last years, attractive methods from the branch of denoising have been proposed with a view to integrating anisotropic diffusion and variational modeling with curvelets and wave atoms aiming to produce regularized filters, as it can be seen in (Ma & Plonka, 2007; Plonka & Ma, 2008; Liu *et al*., 2008; Liu, 2011). However, as these methods are directly based on those transformations, in practical applications they tend to substantially blur the image, besides reproducing the classical *Gibbs phenomenon* (the problem of oscillations close to discontinuity points).

To minimize the mentioned drawbacks, in this chapter we describe a recent *Partial Differential Equation* (PDE) based on nonlinear regularized diffusion which was initially proposed in (Casaca & Boaventura, 2009). The evaluated PDE is capable of restoring a noisy image with a high concentration of textures and fine patterns without the necessity of excessively smoothing important features of the image, i.e., edges, intrinsic details and natural scene parts. It combines the mathematical ideas described in (Nordstrom, 1991; Alvarez *et al*., 1992; Barcelos *et al*., 2003) to recent harmonic analysis works (Demanet & Ying, 2007b; Demanet & Ying, 2007a) about wave atoms so as to define a hybrid method that is regularized by a diffusivity term sensitive to texture and edges. It is worth mentioning that this book chapter is an extended version of (Casaca & Boaventura, 2009).

This chapter is organized as follows: in section 2, we present some existing models which aim noise removal from textured images; In section 3, we describe our regularized PDE-based filter; Numeric implementation is discussed in section 4 while the experimental results and a performance analysis of the evaluated method to other denoising techniques are presented in section 5; In section 6, we present our conclusions about the current filter.

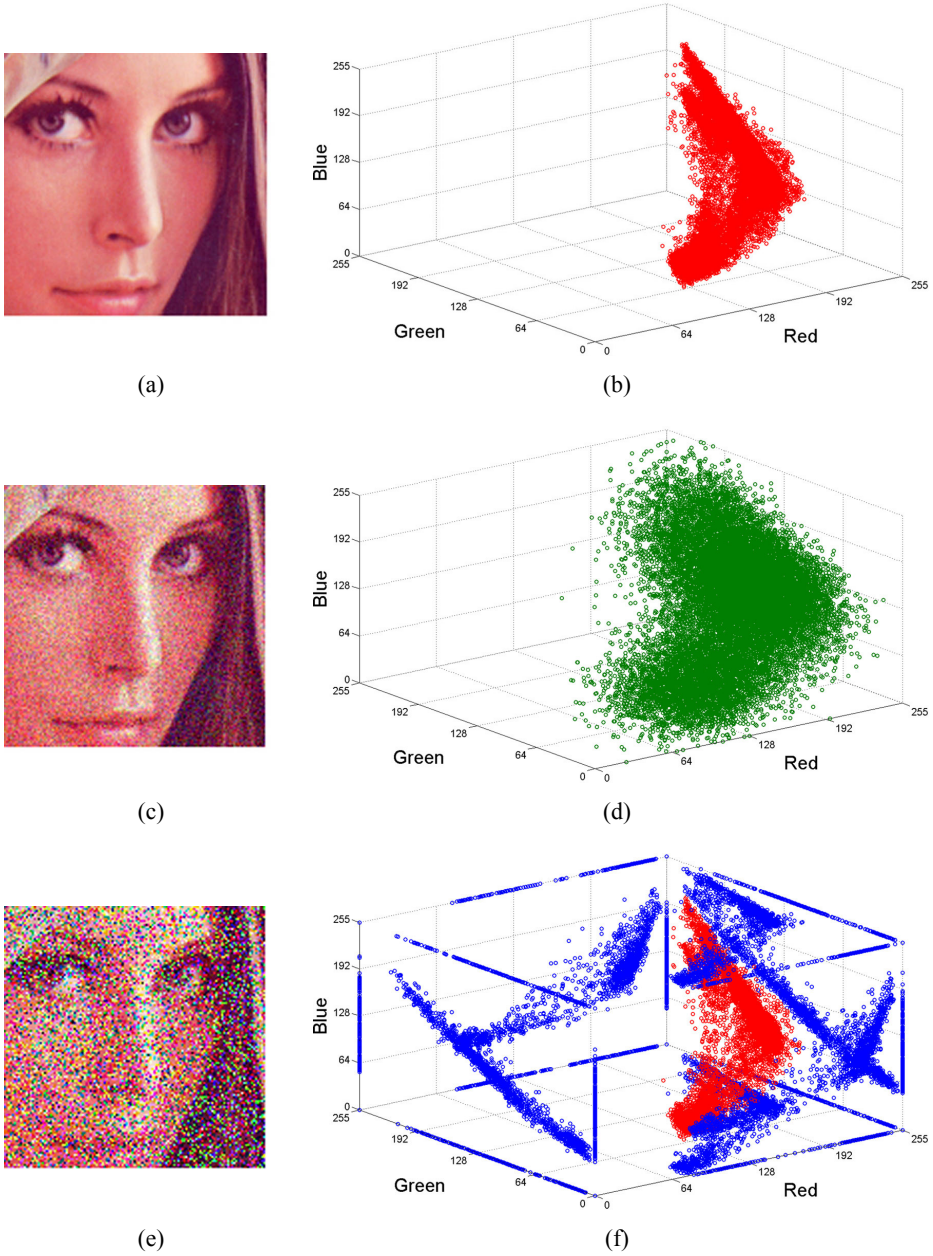
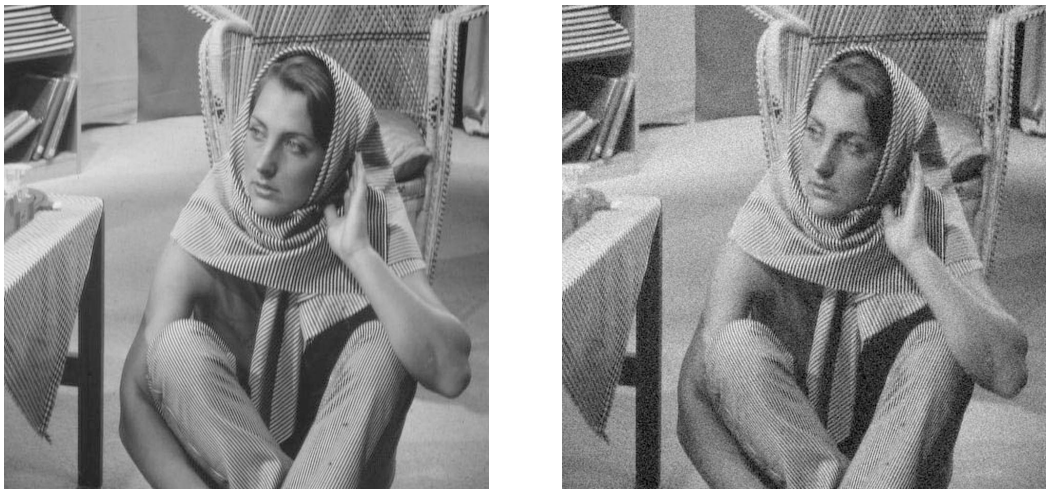


Figure 2: Impact of gaussian and impulsive noise. Samples with red, green and blue circles denote pixels without noise, with gaussian noise, and impulsive noise, respectively. (a) Original image and (b) its plot in RGB space, (c) Image with gaussian noise and (d) its RGB plot, and (e) with impulsive noise and (f) its RGB plot.



(a) (b)

Figure 3: Example of an image containing both geometric and texture information. (a) Original image and (b) contaminated by moderate gaussian noise.

2 PDE-based methods for denoising

In the last twenty years many PDE-based techniques have been proposed aiming to smooth images contaminated by noise. In most cases, those techniques are directly originated through two ways: by introducing a diffusive evolutionary equation (e.g., the Perona-Malik PDE (Perona & Malik, 1990)) or by solving a minimization problem of energy functional where, in general, a PDE is accrued (e.g., all the variants of the classical total variation model from Rudin-Osher-Fatemi (Rudin *et al*., 1992)). In both fronts, a constant concern has always been recovering the image without losing its edges and boundaries. Since the PDE-based filtering can be seen from a mathematical perspective of space scale representation, the major goal is to avoid the monotonic decreasing of basic image structures when the scale increases.

2.1 Texture non-sensitive denoising

Aiming to preserve the elementary image characteristics and simultaneously remove noise, an important contribution in the sense of theoretical and paradigm-shift in denoising was given in (Perona & Malik, 1990) by means of the nonlinear PDE

$$\frac{\partial u}{\partial t} = div(g(|\nabla u|)\nabla u), \tag{1}$$

where $g = g(s)$ is a nonnegative decreasing function satisfying $g(s) \to 0$ when $s \to +\infty$ and $g(0) = 1$. A classical example for this function is $g(s) = \exp(-\tau^2 s)$, with τ being a constant. The central idea of Perona-Malik equation (1) is to employ an oriented diffusion during the noise removal process by exchanging the constant conductivity coefficient of the well-known heat equation (that is, taking $g(s) = 1$ in (1)) by a function $g(|\nabla u|)$ containing the properties before mentioned.

This PDE can be seen as a selective counterbalance between diffusion in the orthogonal and tangential

directions with regard to image edges; in short words:

$$\frac{\partial u}{\partial t} = g(|\nabla u|)\frac{\partial^2 u}{\partial \vec{\xi}^2} + b(|\nabla u|)\frac{\partial^2 u}{\partial \vec{\eta}^2}, \tag{2}$$

with $b(|\nabla u|) = g(|\nabla u|) + 2|\nabla u|^2 \frac{d}{ds} g(|\nabla u|)$ and $\vec{\xi}$ and $\vec{\eta}$ representing the tangential and orthogonal vectors with respect to the edges of u, respectively (see Figure 4). Since we can also write the heat equation in terms of second order derivatives with unitary coefficients

$$\frac{\partial u}{\partial t} = div(\nabla u) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial \vec{\xi}^2} + \frac{\partial^2 u}{\partial \vec{\eta}^2},$$

it is easy to see that the Perona-Malik equation brings a biased competition between tangential and normal diffusion, which depends on the gradient of u and the mathematical expression of conductivity function g as shown in (2). Although this model expresses a relevant contribution in the field of image smoothing, it presents practical and theoretical difficulties (it is an ill-posed problem). For instance, when applied to images with high concentration of noise, the method (1) tends to produce high gradients which implies $g \approx 0$ and, consequently, the filtering does not eliminate the noise.

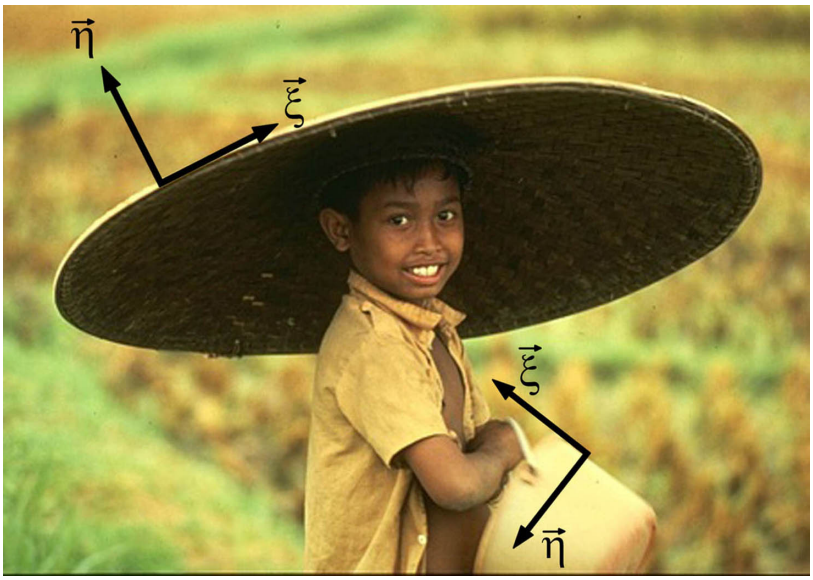


Figure 4: Representation of tangential and orthogonal vector field directions in an image.

In order to convert the ill-posed Perona-Malik problem into a more stable formulation, an interesting mechanism of regularization was introduced by (Catté *et al.*, 1992). They have modified the conductivity coefficient g by changing the classical edge detector based on ∇u to a preprocessed smooth term $\nabla(G_\sigma * u)$, where $*$ means the convolution operator and G_σ is the well-known gaussian kernel. Following these ideas, the authors proposed a successful regularized model as follows

$$\frac{\partial u}{\partial t} = div(g(|\nabla(G_\sigma * u)|)\nabla u), \tag{3}$$

with σ being the standard deviation of the noisy image. The Boundary conditions are the Neumann type.

The method solves the deficiency of Perona-Malik for denoising images with high presence of noise, because it employs a preliminary noise reduction which is performed by the gaussian filtering. Besides, the result of this regularization in g makes the PDE (3) a well-posed problem in the sense of existence, uniqueness and stability of solutions. On the other hand, the Gaussian regularization can blur edges and destroy the texture patterns contained in the image.

In the spirit of regularization strategies, various other regularized diffusion techniques have been proposed for the noise removal task and many of them rely on geometric points of view. An expressive example that defines a diffusive process based on a geometric perspective can be found in (Alvarez *et al*., 1992). The authors considered the first term of (2), which is responsible for accomplishing the orthogonal smoothing regarding the edges, combined by a multiplicative factor, given by the tuning function $g(|\nabla(G_\sigma * u)|)$ proposed in (3). Furthermore, regularized geometric-based techniques have a superior performance if compared to the techniques already described, but they still damage fine edges and sharp corners, in addition to the textures.

Another interesting way of improving the quality of denoising and simultaneously preserving the original structures of the image is to combine selective diffusion with reaction/forcing functions. The model proposed in (Nordstrom, 1991) represents one of the leading exponents of this kind of modeling. The main goal is to solve a minimization problem for a specific energy functional wherein an anisotropic PDE-based method is derived. This equation consists of Perona-Malik model (1) by adding the forcing term $(u-I)$ in its formulation; in other words:

$$\frac{\partial u}{\partial t} = div(g(k|\nabla u|)\nabla u) + (u-I). \tag{4}$$

Here, the Boundary conditions are the Neumann type, k is a constant and the diffusivity function $g = g(s) \in (0,1]$ is given by

$$g(s) = \frac{1}{1+s^2}. \tag{5}$$

The reaction term $(u-I)$ aims to hold the scaled image u close to the input image I since it holds and preserves the natural structures of the image. In addition, the PDE (4) has the property of possessing a nontrivial stationary state, in this way, eliminating the necessity of a stop time criterium. According to (Nordstrom, 1991), it was proved that the insertion of the forcing term $(u-I)$ does not change the theoretical results already achieved for (1).

With the intention to maintain edges and boundaries almost intact while denoising, (Barcelos *et al*., 2003) proposed the following parabolic anisotropic equation which relies on the ideas presented in (Nordstrom, 1991; Perona & Malik, 1990; Alvarez *et al*., 1992):

$$\frac{\partial u}{\partial t} = g|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right) + \alpha(1-g)(I-u), \tag{6}$$

wherein α is a tuning parameter and the Boundary conditions for the above problem are the Neumann type. $g = g(s)$ is given by (5) with $s = \sqrt{\kappa}|\nabla(G_\sigma * u)|$, where G_σ is the Gaussian kernel, σ denotes the standard deviation of the image I and κ is a constant. Similarly to the previous model, g is used to detect edges and locally control the degree of diffusion incidence, that is, if ∇u has a small average value near the neighborhood of a point x (average calculated by the convolution product $G_\sigma * u$), then x is considered an interior point and the applied diffusion by equation (6) will be intense. On the other hand, if the average

value in the neighborhood of x is large, then x is considered as being a point belonging to a boundary region and the diffusion will be small, smoothing homogeneous regions and preserving edges. Accordingly, the anisotropic filter (6) has the ability to smooth the homogeneous parts and preserve boundaries.

Indeed the model (6) does not produce satisfactory results from images with textures, because the diffusivity term g acts locally on the point x (based on the calculation of the "average value" among its neighboring points) and points contained in texture regions possess sudden value oscillations, even in small neighborhoods, which can possibly interfere in obtaining that average.

Unlike the previous model, which considers only the aspect of local edge-detection for eliminating noise, in (Mishra *et al*., 2010) a novel approach for denoising was successfully devised, where a general Bayesian least-squares estimation problem is solved. Its major advantage is the use of a non-local stochastic perspective for detecting and maintaining the fundamental image structures, since the process of noise removal is performed by generating a space scale which has the capability of holding those structures. A quasi-random density estimation model is defined in order to compute the posterior distribution for successive steps of the scale space. To control the speed of smoothing, the authors utilize a set of three objective functions: intensity, gradient and location, which minimize the loss of resolution at each level of the scale space decomposition.

Although the framework (Mishra *et al*., 2010) is not effectively a PDE, its behavior is similar to that of PDE, since it introduces the use of quasi-random scale space theory to represent different scale levels during the denoising process, i.e., a space scale is created such as occurs with PDE-based methods. Due to global aspect built by the stochastic modeling and the three penalizing functions previously mentioned, this approach works very well for images with moderate presence of texture, but in some cases it may produce some artifacts when the image has high-level of texture.

2.2 Texture sensitive denoising

In (Ma & Plonka, 2007), the authors devised a PDE-based technique directed to the problem of textured images. The proposed model was

$$\frac{\partial u}{\partial t} = div(g \nabla (Pu)), \tag{7}$$

with Neumann Boundary conditions, $g = g(|\nabla (Pu)|)$ given by the diffusivity term presented in (5) and P representing the projection operator which is defined by a specific curvelet shrinkage (Cand□s & Donoho, 2004; Cand□s *et al*., 2006). The equation (7) may be seen as the TV diffusion (Rudin *et al*., 1992) managed by the regularization term Pu, which identifies the texture in u. Here, the edge detection provided by g is more sensitive than that presented in the previous model, since it relies on a diffusivity term that best captures texture, due to curvelet transforms. The idea of the authors was to use the diffusive process to minimize the pseudo-Gibbs artifacts produced by the curvelets.

A year later, a variant of equation (7) was proposed by the same authors in (Plonka & Ma, 2008), which modified the diffusion filter (7) by adding the reaction term $(SI - u)$, intending to maintain and improve data labeled as texture. In this case, S is a texture-preserving operator that can be defined from a wavelet, curvelet or wave atom shrinkage. The conception of anisotropic diffusion guided by harmonic analysis techniques in order to preserve and improve the characteristics of an image may be also found in (Liu *et al*., 2008) and (Liu, 2011).

Another important technique in the literature is the model based on the behavior of local adaptability devised by (Gilboa *et al*., 2006), which is defined in relation to the classical approach of Rudin-Osher-Fatemi

(total variation) proposed in (Rudin *et al*., 1992):

$$\inf_{(u,I)\in BV(\Omega)\times L^2(\Omega)} \left\{ \int_\Omega |\nabla u| dx + \frac{\lambda}{2} \int_\Omega |I-u|^2 dx \right\}, \tag{8}$$

where $BV(\Omega)$ denotes the space of functions of bounded variation, Ω the image domain and λ is a parameter to be chosen. The first term in (8) brings the total variation of u while the second is a data-fidelity term. The authors of (Gilboa *et al*., 2006) investigated the model (8) and proposed a new version having the capacity to preserve textures during the diffusion process. Instead of taking λ constant as it was used in (8), they defined $\lambda = \lambda_I(x)$ variable in all Ω, attributing to λ the capacity to represent local image properties such as contrast and textures. The PDE obtained in (Gilboa *et al*., 2006) relies on a modified version of (8) and it is given by the following regularized equation:

$$\frac{\partial u}{\partial t} = div\left(\Phi'(\nabla u)\frac{\nabla u}{|\nabla u|}\right) + \lambda_I(x)(I-u+C), \tag{9}$$

where C is a constant and Φ is a regularization function of ∇u. The factor $div\left(\Phi'(\cdot)\frac{\nabla u}{|\nabla u|}\right)$ represents the diffusion adaptive process and $\lambda_I(x)(I-u+C)$ is the reaction term locally weighted by λ. Therefore, the coefficient λ performs the function of contributing to the maintenance of the texture in u. Other recent models of texture-preserving denoising derived from the ideas presented in (Gilboa *et al*., 2006) can be found in (Gilles, 2007) and (Li *et al*., 2009).

3 Regularized nonlinear anisotropic equation

In this section we discuss the regularized anisotropic equation recently presented in (Casaca & Boaventura, 2009). Some of its fundamental mathematical properties have been investigated here such as regularization, reaction-diffusion strategy, texture and edge detecting, existence and uniqueness of solutions, among others.

3.1 Mathematical background and regularized PDE-based filtering

Let $u = u(\mathbf{x},t)$ be a grayscale image which is time-varying and $I = I(\mathbf{x})$ the observed image (contaminated by noise), represented respectively by the real functions

$$u(\cdot,t) : \Omega \times [0,+\infty] \to \mathbb{R} \quad \text{and} \quad I : \Omega \to \mathbb{R}.$$

We assume that the domain Ω is a rectangular region of $\mathbb{R}^2$, while the degradation due to noise in I is additive (see Figure 1), wherein $\mathfrak{n}$ represents the noise function modeled by a Gaussian statistical distribution with mean 0 and variance σ^2,

$$\int_\Omega \mathfrak{n}(\mathbf{x}) d\mathbf{x} = 0 \quad \text{and} \quad \int_\Omega \mathfrak{n}^2(\mathbf{x}) d\mathbf{x} = \sigma^2 . \tag{10}$$

We suppose also that the time-varying image u contains oscillatory patterns such as texture, fine details and high-frequency information. The goal here is to reduce the contamination by noise from the target image I; in other words: the impact of the noise $\mathfrak{n}(\mathbf{x})$ must be minimum in the output image, making it visually closer to the associated noise-free image without destroying intrinsic textures and other significant image features.

In the spirit of (Barcelos *et al*., 2003; Gilboa *et al*., 2006; Plonka & Ma, 2008), recently a new regularized anisotropic diffusion filtering has been proposed successfully in (Casaca & Boaventura, 2009), which addresses the problem of denoising for textured images. This filtering is performed taking into account the mixing among oriented diffusion, regularization, weighing-based strategy and reaction-based scheme into a unified Partial Differential Equation:

$$\frac{\partial u}{\partial t} = h|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right) + \alpha(1-h)(I-u) + \beta(1-h)(SI-u), \ \mathbf{x} \in \Omega, \ t > 0, \tag{11}$$

$$u(\mathbf{x},0) = I(\mathbf{x}), \ \mathbf{x} \in \Omega, \quad \frac{\partial u}{\partial \vec{n}} = 0, \ x \in \partial\Omega, \ t > 0,$$

where h is the diffusivity term, S is an operator based on texture descriptor and α and β are tuning parameters. As observed in (11), the Boundary conditions are the Neumann type and the initial condition to PDE (at $t = 0$) is set by using the input noisy image I. Figure 5 shows an illustrative example of applying the equation (11).

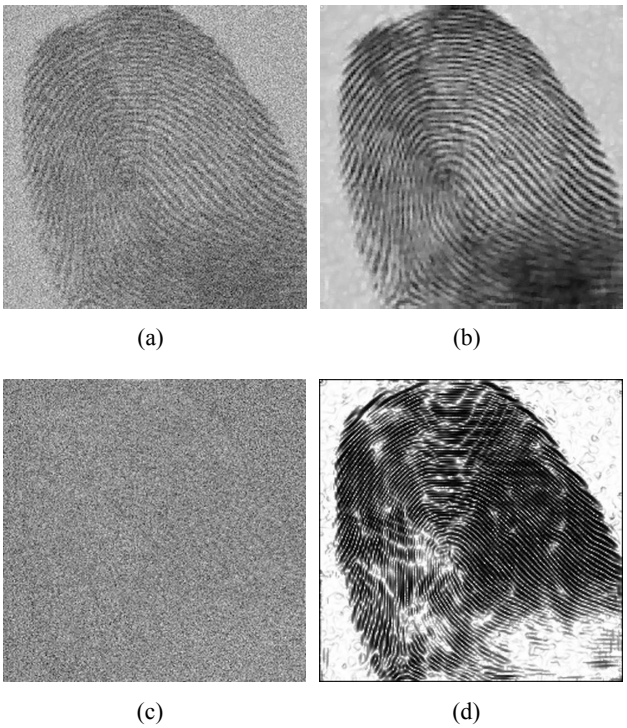


Figure 5: Example of denoising performed by our approach. (a) Noisy fingerprint image, (b) reconstructed image, (c) residual image (difference between (a) and (b)) and (d) representation of the mapping term h.

The PDE (11) is regularized by employing a penalization term h (see Figure 5(d)) that ensures a good trade off between noise removal and texture preserving. Besides, it also achieves an improvement on information patterns, which are naturally observed in an image like oriented texture and other specific features.

3.2 Evaluating the contribution of diffusion

Let us consider the level curves regarding to u in the scale t, $S_\varepsilon(t) := \{\mathbf{x} \in \Omega : u(\mathbf{x},t) = \varepsilon\}$. The first non-regularized component on equation (11),

$$|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right) \tag{12}$$

employs a diffusive process in direction tangential to level sets $S_\varepsilon(t)$. This statement can be shown by taking the level set directions $\overrightarrow{\xi} = \frac{\nabla^{\perp} u}{|\nabla u|}$ (because the gradient vector ∇u is ever normal to level curves) and computing the second-order derivative of u relative to related direction:

$$\frac{\partial^2 u}{\partial \overrightarrow{\xi}^2} = \overrightarrow{\xi}\,\mathcal{H} u\,\overrightarrow{\xi}^T = \frac{\frac{\partial^2 u}{\partial x^2}\left(\frac{\partial u}{\partial y}\right)^2 - 2\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial xy} + \frac{\partial^2 u}{\partial y^2}\left(\frac{\partial u}{\partial x}\right)^2}{|\nabla u|^2} = |\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right), \tag{13}$$

where the operator $\mathcal{H}$ is the Hessian matrix. Since the factor

$$\rho(\mathbf{x},t) = div\left(\frac{\nabla u(\mathbf{x},t)}{|\nabla u(\mathbf{x},t)|}\right)$$

represents the curvature for the level curve $S_\varepsilon(t)$ at point x in the time t, the degenerate divergent (12) performs a curvature-driven smoothing on the image. The expression (12) is usually mentioned in the literature as *Mean Curvature Flow* (MCF) and it is responsible for the diffusion step associated to PDE (11). According to the geometric interpretation provided by (13), the MCF diffuses u into the orthogonal direction of its gradient vector ∇u and does not diffuse it in any other direction.

3.3 Understanding the regularization term *h*

The strength of the diffusion is handled by applying the regularized factor $h = h_\mu(\sqrt{\delta}|\nabla Su|)$, which is defined as:

$$h = \begin{cases} g(\sqrt{\delta}|\nabla Su|), & \text{if } g(\sqrt{\delta}|\nabla Su|) \le \mu \\ g(\sqrt{\kappa}|\nabla u|), & \text{if } g(\sqrt{\delta}|\nabla Su|) > \mu \end{cases}, \tag{14}$$

where g relies on the diffusivity expression (5), δ and κ are weighting parameters, and μ is a threshold value that controls the capturing of gray shades in relation to split function $g(\sqrt{\delta}|\nabla Su|)$.

We may also rewrite the equation (14) in an explicit way at the point $\mathbf{x} \in \Omega$ as follows:

$$h_\mu(\mathbf{x}) = g(O_1(\mathbf{x}))\chi_{D_\mu} + g(O_2(\mathbf{x}))\chi_{D_\mu^c}, \tag{15}$$

wherein $O_1(\mathbf{x}) = \sqrt{\delta}|\nabla Su|$ and $O_2(\mathbf{x}) = \sqrt{\kappa}|\nabla u|$ are operators, $D_\mu = \{\mathbf{x} \in \Omega : g(O_1(\mathbf{x})) \le \mu\}$ determines the threshold region, D_μ^c expresses the complementary of D_μ and μ is a suitable fixed threshold value. Obviously, we have $\Omega = D_\mu \cup D_\mu^c$.

Using a hybrid diffusivity factor h instead of taking only $g(O_1(\mathbf{x}))$ or $g(O_2(\mathbf{x}))$ can be justified by the following statements:

1. While $g(O_1(\mathbf{x}))$ adequately captures texture pattern points into the image (that is, $g(O_1(\mathbf{x})) \sim 0 \Rightarrow (1 - g(O_1(\mathbf{x}))) \sim 1$ at those points), for the homogeneous regions, where the effects of diffusion must be more intense, $g(O_1(\mathbf{x}))$ also sets a value close to zero (see Figure 6(a) for an illustration).

2. In an opposite way, the term h sets $g(O_2(\mathbf{x})) \sim 1$ to homogeneous parts, but in the regions characterized by texture and other high-level details, $g(O_2(\mathbf{x}))$ does not efficiently detect these information (see the example of Figure 6(b)).

3. We would like $g(O_1(\mathbf{x}))$ as well as $g(O_2(\mathbf{x}))$ to get near both extremities (black and white), but they should not reach such values as this would annul completely the contribution from the diffusion or two reaction terms considered in (11), creating abrupt image discontinuities.

Figure 6 clearly illustrates the summarized issues above.

(a) (b)

(c)

Figure 6: The proposed diffusivity term $h \in (0, 1]$ applied in image 3(b), where the color mapping can be understood by 0 ~ black and 1 ~ white. (a) Component $g(O_1(\mathbf{x}))$, (b) $g(O_2(\mathbf{x}))$ and (c) h, the balancing between both results.

3.4 Balancing the equation through reaction terms

Equation (11) can be understood as a balance between smoothing and "keeping close" to I and SI components. This balancing is governed by the adaptive term h, which may be interpreted as an edge-and-texture detector, in addition to control the velocity of the diffusion.

As described previously, it may be observed that the value of h is quite high for the nontextured regions. Then, $(1-h) \sim 0$ and, hence, the forcing terms

$$(I-u) \quad \text{and} \quad (SI-u)$$

act in a practically insignificant way in the composition of (11). Consequently, the MCF expression (12) presented in the first part of (11) will considerably smooth the image. On the other hand, in regions wherein h is small, we have $(1-h) \sim 1$. Thus, the difference $(I-u)$ holds edges and boundaries from the noisy image I in synergy with the reaction term $(SI-u)$, which aims to preserve and reconstruct oscillatory characteristics of the image including texture and other details.

Therefore, adding both reaction components in the equation (11), we can ensure simultaneously the task of reducing the degenerative effects of the diffusion process and recovering important image features.

3.5 Capturing patterns and textures

In the formulation of base equation (11), we employ the operator S to compute one of the reaction terms and to identify regions defined by texture patterns. In our work, S relies on recent and useful harmonic analysis tool called *wave atoms* (Demanet & Ying, 2007a; Demanet & Ying, 2007b).

Wave atom is a variant of *2D wavelet packets* obeying the important relationship of the wavelength $wavelenght \sim (diameter)^2$ parabolic scale, which improves the sparse representation of certain oscillatory patterns. This means that distorted oscillatory functions (e.g., oriented textures) have a significant sparser expansion in wave atom modeling than in other representations such as *pure wavelets*, *Gabor atoms* or *curvelets*, for instance. Wave atoms composition elements have high directional sensitivity and are anisotropic, which makes them ideal for applications where the goal is to identify regions characterized by oscillatory patterns such as textures and oscillating details.

Although wave atoms can be used successfully to accomplish a great number of image processing tasks, for the denoising problem, it is recommendable to use wave atom shrinkage. In most cases, it is calculated by the following formulation:

$$u_c = \sum_{\mu} \Theta(c_{\mu}^{(1)}(u))\varphi_{\mu}^{(1)} + \Theta(c_{\mu}^{(2)}(u))\varphi_{\mu}^{(2)}, \tag{16}$$

where $\varphi_{\mu}^{(i)}, i = 1,2$ are $2D$ versions of a frame of the wave atom derived from wave packets (see (Demanet & Ying, 2007b) or (Casaca & Boaventura, 2010) for a detailed mathematical background) and $\Theta = \Theta_{\theta}$ denotes a thresholding function. For example, we may take

$$\Theta_{\theta}(x) = \begin{cases} x, & |x| \geq \theta \\ 0, & |x| < \theta, \end{cases} \tag{17}$$

with θ being the threshold value.

3.6 Mathematical validity of the main equation

The steps for proving the theoretical validity of PDE (11) are: firstly, we have to examine the parabolicity of equation (11) for a predefined function h respecting some restrictions and after, to prove the existence and uniqueness of solutions. To ascertain the first assumption, the goal is to write the current equation through the general second-order PDE formula:

$$\frac{\partial u}{\partial t} = \sum_{i,j=1}^{2} a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{j=1}^{2} b_j \frac{\partial u}{\partial x_j} + cu + f, \tag{18}$$

with a_{ij}, b_j and c representing coefficient functions (which depends on the variables embedded on the target PDE) and $f = f(\cdot, t)$. After, we may examine the parabolicity of equation (11) by the following definition (Evans, 2010):

Definition 1. *The generic PDE* (18) *is called parabolic for given coefficients* a_{ij}, b_j, c, $(i, j = 1, 2)$, *if and only if the coefficient matrix* $A = (a_{ij})$ *is positive definite for the problem domain, in short words:*

$$\sum_{i,j=1}^{2} a_{ij} \rho_i \rho_j > 0 \quad \forall\ (\mathbf{x}, t) \in \Omega \times (0, +\infty),\ \nabla u \neq (0,0) \ \textit{and}\ \rho = (\rho_1, \rho_2) \in \mathbf{R}^2.$$

Manipulating (11) in a convenient way and using the relationship (13) we have:

$$\frac{\partial u}{\partial t} = h\left(\frac{\partial u}{\partial y}\right)^2 \frac{\partial^2 u}{\partial x^2} - 2h \frac{\partial u}{\partial x} \frac{\partial u}{\partial y} \frac{\partial^2 u}{\partial x \partial y} + h\left(\frac{\partial u}{\partial x}\right)^2 \frac{\partial^2 u}{\partial y^2} + (h-1)(\alpha+\beta)u + \alpha(1-h)I + \beta(1-h)SI =$$

$$a_{11} \frac{\partial^2 u}{\partial x^2} + 2a_{12} \frac{\partial^2 u}{\partial x \partial y} + a_{22} \frac{\partial^2 u}{\partial y^2} + cu + f,$$

where $a_{11} = h\left(\frac{\partial u}{\partial y}\right)^2$, $a_{12} = -h \frac{\partial u}{\partial x} \frac{\partial u}{\partial y}$, $a_{22} = h\left(\frac{\partial u}{\partial x}\right)^2$, $c = (1-h)(\alpha+\beta)$ and $f = \alpha(1-h)I + \beta(1-h)SI$.

According to parabolicity *Definition 1*, we may observe that for all $\rho \in \mathbb{R}^2$:

$$\sum_{i,j=1}^{2} a_{ij} \rho_i \rho_j = h \langle \nabla^{\perp} u, \rho \rangle^2 \geq 0, \quad \text{if } h > 0.$$

Hence, we can conclude that the PDE (11) is parabolic with *degeneration possibilities*, since the function g defined in h is strictly positive for all $(\mathbf{x}, t) \in \Omega \times (0, +\infty)$.

About the existence and uniqueness of solutions to parabolic nonlinear equation (11), it may be verified in the sense of *viscosity solutions* (Crandall *et al*., 1992) following the ideas of (Alvarez *et al*., 1992; Barcelos *et al*., 2003).

The theory of viscosity solutions is essentially based on the study of *Maximum Principle* and it can be applied successfully to solve both linear and nonlinear PDE of any order. To investigate the solution of the current equation with respect to viscosity theory, first of all, we suppose the Boundary conditions as being Neumann type and then, from the original formulation (11), an equivalent expression is derived. Without loss of generality we may take $\mu = \delta = \kappa = 1$ in h and $\alpha = \beta = \frac{1}{2}$ in (11). Thus,

$$\frac{\partial u}{\partial t} = h|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right) + \frac{1}{2}(1-h)(I-u) + \frac{1}{2}(1-h)(SI-u) = h|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right) + \frac{1}{2}(1-h)(I-2u+SI)$$

$$= h|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right) + (1-h)\left(\frac{1}{2}(I+SI) - u\right) = h|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right) + (1-h)(\widetilde{I}-u), \tag{19}$$

where $\widetilde{I} = \frac{1}{2}(I+SI)$ is the average between the noisy image I and the transformed image SI: it is a fixed image and does not vary when t changes.

Equation (19) can be seen as a variant of parabolic PDE (6). Since the wave atom shrinkage embedded in h works like a smoothing kernel, which appears in the diffusivity coefficient (6), both existence and uniqueness may be rigorously checked in the sense of viscosity solutions by following the same ideas developed in (Barcelos *et al*., 2003): to build a Cauchy problem and to exploit the properties of viscosity sub-solution and super-solution considering the verification of Maximum Principle.

4 Numerical approximation

We assume that all the involved digital images are numerically represented by $2D$ matrices, where each element u_{ij} denotes the intensity value with respect to continuous image $u(x,y)$ at the point

$$(x,y) = (x_i, y_j), \text{ with } x_i = i\Delta x \text{ and } y_j = j\Delta y.$$

In order to simplify the notation, we set $\Delta x = \Delta y = 1$. We denote $u(x_i, y_j, t_n)$ by u_{ij}^n, where $t_n = n\Delta t$. Here, t_n represents the scalar "temporal" parameter, Δt is its step and $n = 1, 2, \ldots, N$, with N being the number of iterations associated to the discrete process of obtaining numerical solutions.

4.1 Computing the approximation to PDE

The derivatives in (11) are numerically approximated by using finite difference techniques, which are quite precise when implementing computational algorithms.

The temporal derivative $\frac{\partial u}{\partial t}$ is calculated at the point (x_i, y_j, t_n) by the advanced difference

$$\frac{\partial u}{\partial t} \sim \frac{u_{ij}^{n+1} - u_{ij}^n}{\Delta t}, \quad \text{which implies that} \quad u_{ij}^{n+1} \approx u_{ij}^n + \Delta t \frac{\partial u}{\partial t}, \tag{20}$$

while the MCF diffusion expression

$$|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right) = \frac{\frac{\partial^2 u}{\partial x^2}\left(\frac{\partial u}{\partial y}\right)^2 - 2\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial xy} + \frac{\partial^2 u}{\partial y^2}\left(\frac{\partial u}{\partial x}\right)^2}{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2}$$

is numerically approximated by the employment of central and second-order difference schemes as follows:

$$\frac{\partial u}{\partial x} \sim \frac{u_{i+1j}^n - u_{i-1j}^n}{2\Delta x}, \quad \frac{\partial u}{\partial y} \sim \frac{u_{ij+1}^n - u_{ij-1}^n}{2\Delta y}, \quad \frac{\partial^2 u}{\partial x^2} \sim \frac{u_{i+1j}^n - 2u_{ij}^n + u_{i-1j}^n}{\Delta x}, \quad \frac{\partial^2 u}{\partial y^2} \sim \frac{u_{ij+1}^n - 2u_{ij}^n + u_{ij+1}^n}{\Delta y};$$

$\frac{\partial^2 u}{\partial xy} \sim \frac{u_{i+1j+1}^n - u_{i-1j+1}^n - u_{i+1j-1}^n + u_{i-1j-1}^n}{4\Delta x \Delta y}$. In some cases, the accuracy of the approximations may be improved by setting the values already calculated in step $n+1$ rather than n.

In (14), both $|\nabla u|$ and $|\nabla Su|$ are computed through the calculation of the Euclidian norm with quantization of derivatives, similar to finite difference equations used to approximate the MCF expression. To compute the signals derived from S we propose a wave atom shrinkage as follows:

$$Sf = (WA)^{-1}\Theta_\theta(WA)(f), \tag{21}$$

where f can be taken as u or I, (WA) denotes the wave atom transform originated from (16) (Demanet & Ying, 2007b), $(WA)^{-1}$ is its the inverse transform and Θ_θ is the hard threshold function (17). For the numerical implementation of wave atoms transforms, we have employed the discretization suggested in (Demanet & Ying, 2007b), where the authors proposed a model based on the fast Fourier transform and a wrapping trick. A computational prototype (written in C and MATLAB) called *WaveatomLab packet* of discrete wave atoms can be found in the homepage *www.waveatom.org* for non-commercial evaluation purposes.

Using Neumann boundary conditions mixed with the expression (20), we may calculate u_{ij}^{n+1}, $n = 0, 1, 2, ..., N-1$, by the associated discrete formula of (11):

$$u_{ij}^{n+1} = u_{ij}^{n} + \Delta t\, \Gamma(u_{ij}^{n}), \;\; \forall\, (i,j) \in \Omega, \tag{22}$$

wherein $u_{ij}^{0} = I(x_i, y_j)$, N denoting the number of iterations and $\Gamma(u)$ representing the right side of (11), except for calculation of h, which is computed previously for each iteration of (22). The complete algorithm of our technique is given as follows:

```
1  Input: I (noisy image), Δt (temporal step), N (number of iterations), α
       ↪  and β (PDE weights), δ and κ (weights of the mapping term h), μ
       ↪ (threshold of h) and θ (wave atom parameter).
2  Output: ũ (denoised image).
3  compute SI by using Equation (21), set v^0 = SI;
4  for n = 0 : N − 1 do\\
5     update boundary pixels of u by applying Neumann boundary conditions;
6     compute O_1(i,j) = √δ |∇v_ij^n|, O_2(i,j) = √κ |∇u_ij^n| and h = h_ij according to
          ↪ expression (14), ∀ (i,j) ∈ Ω;
7     update u_ij^{n+1} by Equation (22), taking into account the computation of
          ↪ h in the previous step;
8     update v^{n+1} = Sv^n by Equation (21);
9  endFor
```

Listing 1: Algorithm for the Numerical computation of Equation (11)

4.2 Using genetic algorithm-based strategy to improve the accuracy

Genetic Algorithms (Holland, 1975) are search and non-deterministic optimization techniques inspired by the biological principles of evolution and natural selection of species. These methods seek to find optimal solutions to a given problem through a probabilistic search in a predefined search space. Thus, they have been largely exploited in optimization problems. The possibility of finding optimal or appropriate solutions

for complex problems without using specific information about the problem is one of its main features (Goldberg, 1989).

Each solution of a problem is represented by an individual in a population (a set of possible solutions) and the feature of an individual are typified by a chromosome. New solutions are generated within the search space by modifying the current individuals through selection, recombination and mutation operators, that simulate the reproductive process. In addition, a *fitness function* is used to simulate the natural selection stages and to assign a quality measure for each individual. Therefore, individuals with best fitness values have more chance of being selected for reproduction and also to remain in the next generation. The flowchart in Figure 7 illustrates the main steps of the genetic algorithm.

In our problem, a real-coded genetic algorithm was used in order to obtain appropriate values for the set of parameters to recursive formulation (22). Individuals correspond to a given set of parameters and its chromosome is encoded by the real vector

$$v = (N, \alpha, \delta, \kappa, \mu, \theta), \tag{23}$$

where N, α, δ are embedded in the central operator Γ, κ, μ are defined in (14) and θ is used in (21). The fitness function here used was the classical denoising metric *Peak Sinal Noise-to-Ratio* (PSNR), however, it is possible to use a mixing among some other fitness measure functions as suggested in recent literature of Genetic Algorithms applied to image processing. Since a more sophisticated modeling with Genetic Algorithms is beyond the scope of this work we did not employ multiple fitness-based strategies in our estimations.

Still on the details of using the GA in our numerical approximations, we have used the *Feasible Population* function to create a random initial population that satisfies bounds constraints. Raw fitness values that are returned by the fitness function are in a range that is appropriate for the selection function. Thus, a *Rank* fitness scaling function can be employed to scale the raw scores. This one is based on the rank of each individual instead of its fitness values. The rank of an individual is its position in the sorted fitness values and an individual with rank r has scaled a score proportional to $1/\sqrt{r}$.

The selection operator defines how the genetic algorithm chooses individuals who will be used to form the new population while the *Tournament* function chooses each individual taking the best element of the set composed of four players (the default tournament size) who are chosen randomly and, hence, was utilized as the selection function.

In our experiments, we have applied the genetic algorithm taking into account two elite children, mutation and crossover operators to create new individuals. The mutation operator mutates the individual's value with small random changes to set mutation children. The *Adaptive Feasible* generates random directions that are adaptive with respect to the last generation and may be used as mutation function. The *Crossover operator* specifies how the genetic algorithm combines two individuals, or parents, to make a crossover child for the next generation. Finally, the *Heuristic* returns a new individual that lies on the line containing the two parents, with a small distance away from the parent with better fitness value toward the other parent, which was here utilized as crossover function.

The algorithm employs default stopping criteria, except for the values of maximum number of iterations and the function tolerance which were defined as 20 and 0.1 respectively. Implementation of the Genetic Algorithm was performed by using the *Global Optimization Toolbox of Matlab*® R2011b (The MathWorks, 2011) with default values, except to small modifications regarding implementation details, explained above.

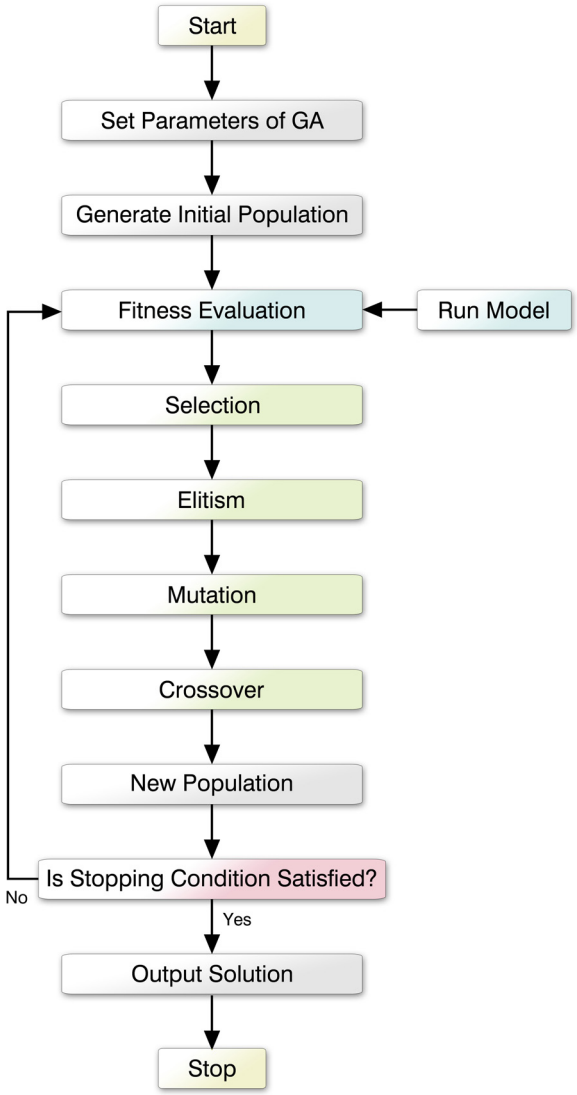


Figure 7: Genetic Algorithm flowchart shows the main steps of the GA taking into account the interaction on fitness evaluation.

5 Experimental results

In this section we present the numerical results provided by our model in addition to the comparative study to some recent texture-sensitive denoising techniques described in the literature. Both gray-scale and color RGB images were used in our experiments and, for this last category, the algorithm was performed for each channel of the image. Here, the evaluated images are represented by matrices of dimensions 128×128 and 256×256, (see the Table 1). In our anisotropic filter, the fixed parameters $\beta = 1$ and $\Delta t = 0.1$ were set in all experiments.

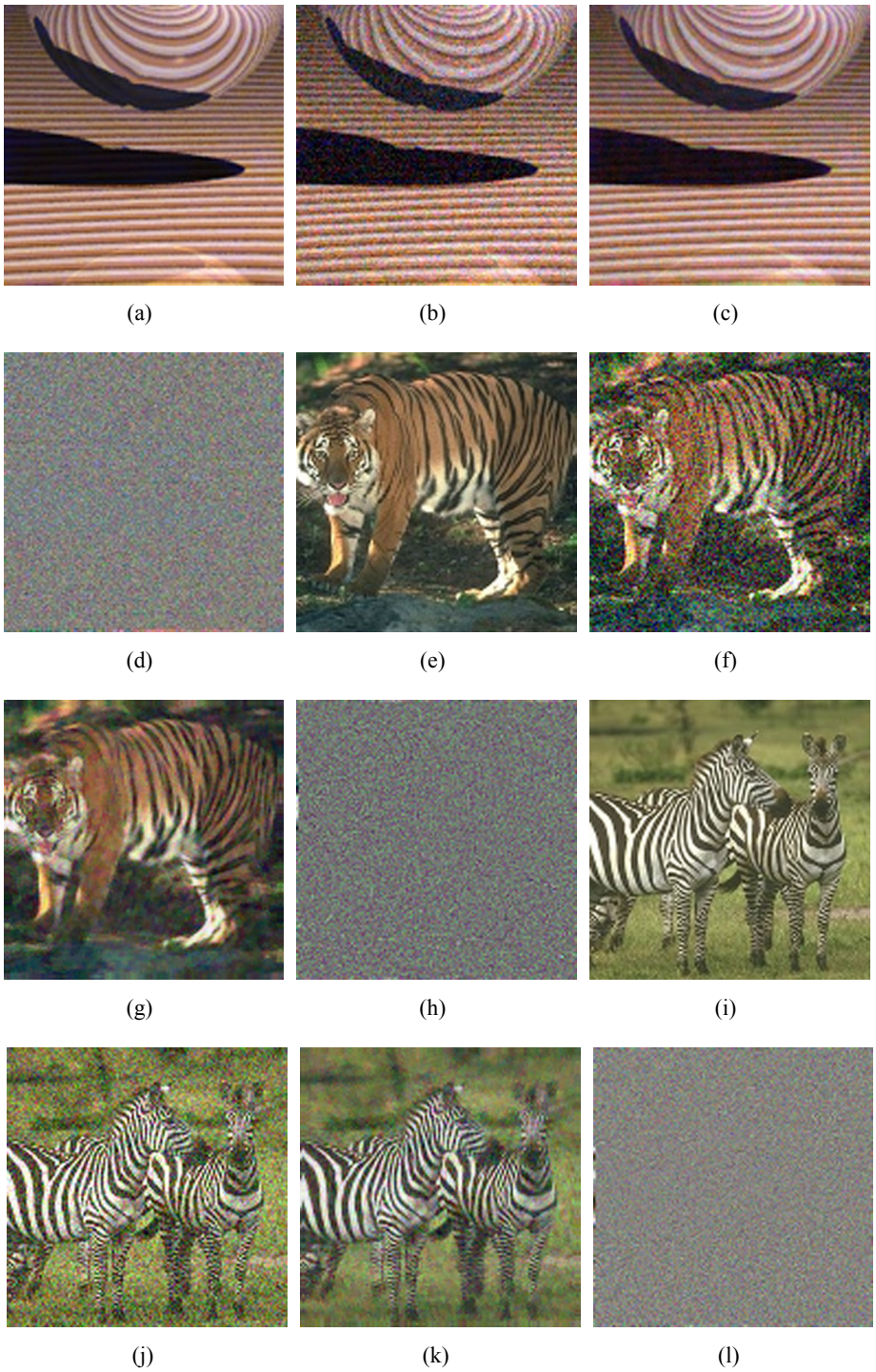


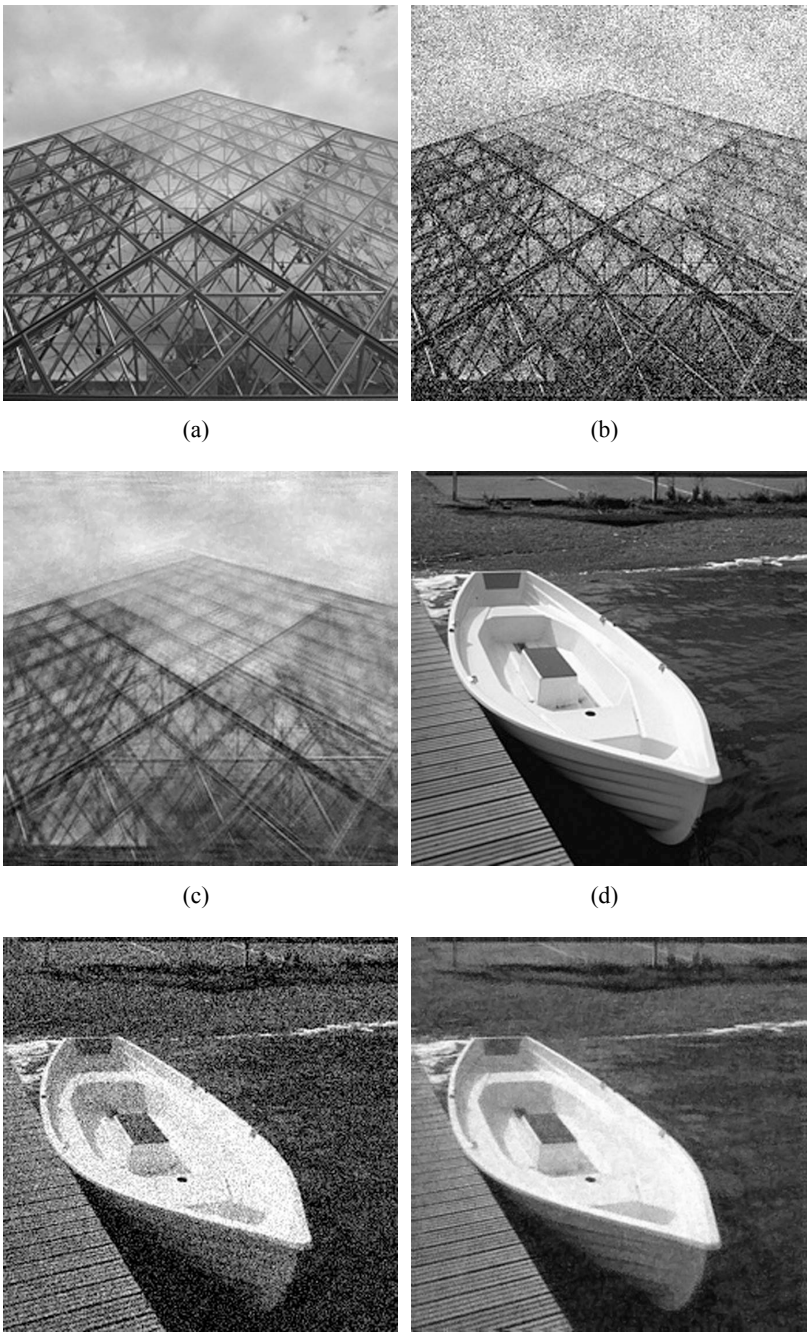
Figure 8: Denoising of color images. (a)-(e)-(i) Original images, (b)-(f)-(j) its noise version, (c)-(g)-(k) denoised images and (d)-(h)-(l) the residual components.

Image	Size	RGB/Gray	Param vec. v
Ball	128×128	RGB	$(20, 0.20, 0.04, 0.001, 0.80, 0.05)$
Tiger	128×128	RGB	$(20, 0.22, 0.07, 0.020, 0.83, 0.08)$
Zebra	128×128	RGB	$(20, 0.15, 0.17, 0.016, 0.67, 0.07)$
Louvre	256×256	Gray	$(27, 0.16, 0.47, 0.002, 0.34, 0.15)$
Boat	256×256	Gray	$(30, 0.44, 0.23, 0.002, 0.77, 0.17)$
Nanotechnology	256×256	Gray	$(65, 0.20, 0.04, 0.001, 0.70, 0.15)$
Barbara face	256×256	Gray	$(20, 0.80, 0.09, 0.001, 0.25, 0.80)$

Table 1: Informative table containing the properties of the evaluated images. Here, all damaged images are artificially generated from their noiseless original versions, except the *Nanotechnology image*. The vector v represents the parameters that we used in our model (see the expression (23)).

Figure 8 displays the visual results by applying the current technique for color images. In both cases, the method produced good results since the residual matrices (Fig.8(d) and Fig.8(h)) did not hold the essential image features, only the noise was retained. Figure 8(i) shows another photography which the task of reconstructing was accomplished successfully. In the experiments of Figure 9 we show textured photographs where the filter (11) has been applied. It is notorious the quality of restoring with respect to cognition and texture-preserving aspects. For example, in Fig.9(a) (*Louvre museum*), the oriented patterns (defined by the textured lines) were maintained almost intact on the restored image Fig.9(c). With relation to Fig.9(d), we can observe that both patterns consisting of ”wood” and the ”boat” were denoised in a satisfactory way (see Fig.9(f)). The result of Figure 10 addresses a practical application where we have not the desired image to determine a visual approximation, more specifically: this example brings a heavy noisy image originated through a nanotechnology sample. According to Figure 10(b), we can see that our filter achieved a good denoising result maintaining, in this way, all the relevant data of the image.

We finish our numerical experiments evaluating the current filter with other ones in the literature. To accomplish this task, we took a piece of Barbara's well-known photography (Figure 3). Fig.11(a) contains important features to be preserved and recovered: the texture above the neck, the background and some intrinsic details like eyes, nose and the hand. Fig.11(b) and Fig.11(c) were obtained using fast discrete curvelet transform (Cand□s & Donoho, 2004; Cand□s *et al*., 2006) and $2D$ discrete wave atom-based technique (Demanet & Ying, 2007b), respectively. The texture was not appropriately recovered by using the curvelet, in addition to the emergence of some artifacts. On the other hand, the wave atom transform restored all the texture but produced a blurred image. Fig.11 (d) shows the denoised image with the model based on adaptive fidelity term (9) using $C = 0$, while Fig.11(e) displays the reconstruction generated by diffusion combined with curvelets (7), with step 0.0005 and 11 iterations. Although the first method had retrieved parts of the texture, important details of the image were excessively smoothed (e.g., face and hand). In contrast, the second method did not produce any excessive smoothing, but still kept a lot of noise in the processed image. Fig.11(f) is the output image using our technique. Both texture and image details (intrinsic data) are recovered, and the noise level was minimized without causing excessive smoothing. The parameters used in all the related techniques were chosen according to authors' works or through the best visual quality obtained by applying each of these models in an extensive way (taking into account the PSNR metric (Table 2)).

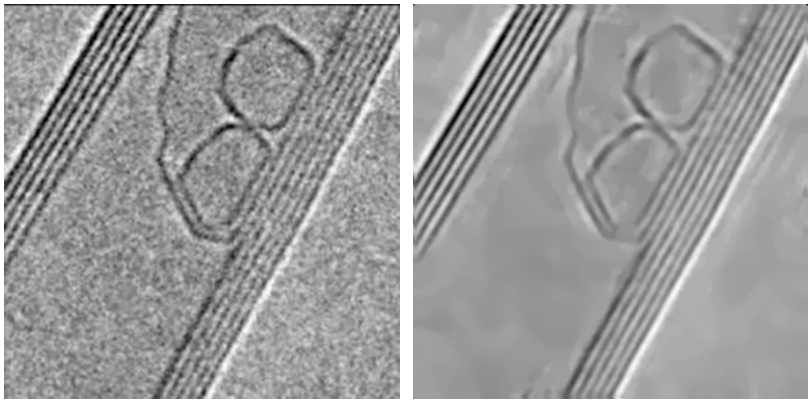


(a) (b)

(c) (d)

(e) (f)

Figure 9: Experimental results performed by our technique. (a)-(d) Original images, (b)-(e) contaminated version by heavy degradation of noise and (c)-(f) recovered images.



(a) (b)

(c)

Figure 10: Practical application (nanotechnology image). (a) Image to be reconstructed, (b) its recovered version and (c) the texture representation through diffusivity term in the final iteration of the algorithm.

Figure	**Technique**	**PSNR (dB)**
Fig. 11(b)	Curvelet	22.80
Fig. 11(c)	Wave atom	23.54
Fig. 11(d)	Gilboa et al.	24.25
Fig. 11(e)	Ma et al.	23.43
Fig. 11(f)	Our	25.51

Table 2: Quantitative measure calculated for all methods with respect to Barbara's original image and the restored results. PSNR between the original image and noisy image is 22.69 dB.

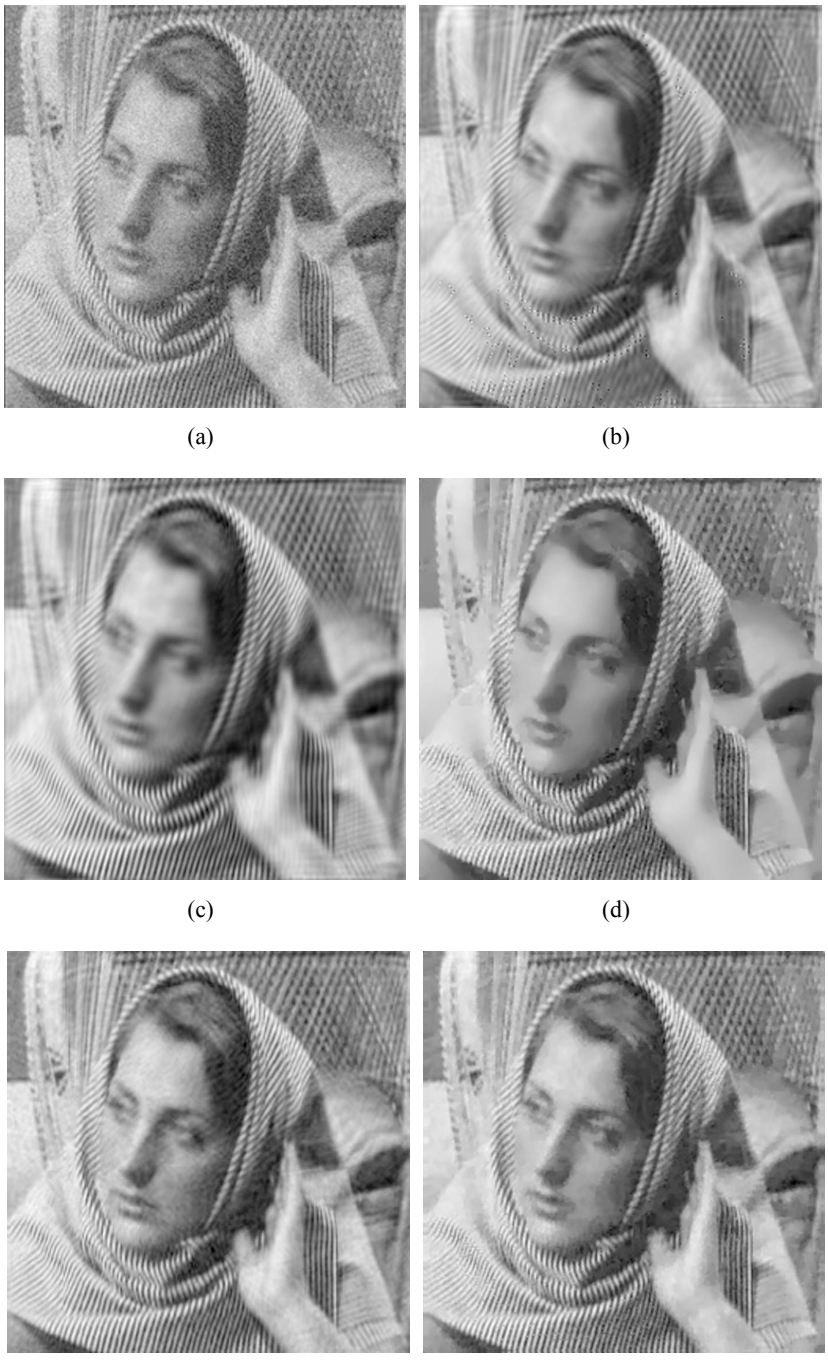


Figure 11: Comparative results of Barbara's photography. (a) Target image, (b) generated by curvelets, (c) by wave atoms, (d) by diffusion with adaptive-fidelity term, (e) by TV diffusion combined with curvelet and (f) by proposed equation.

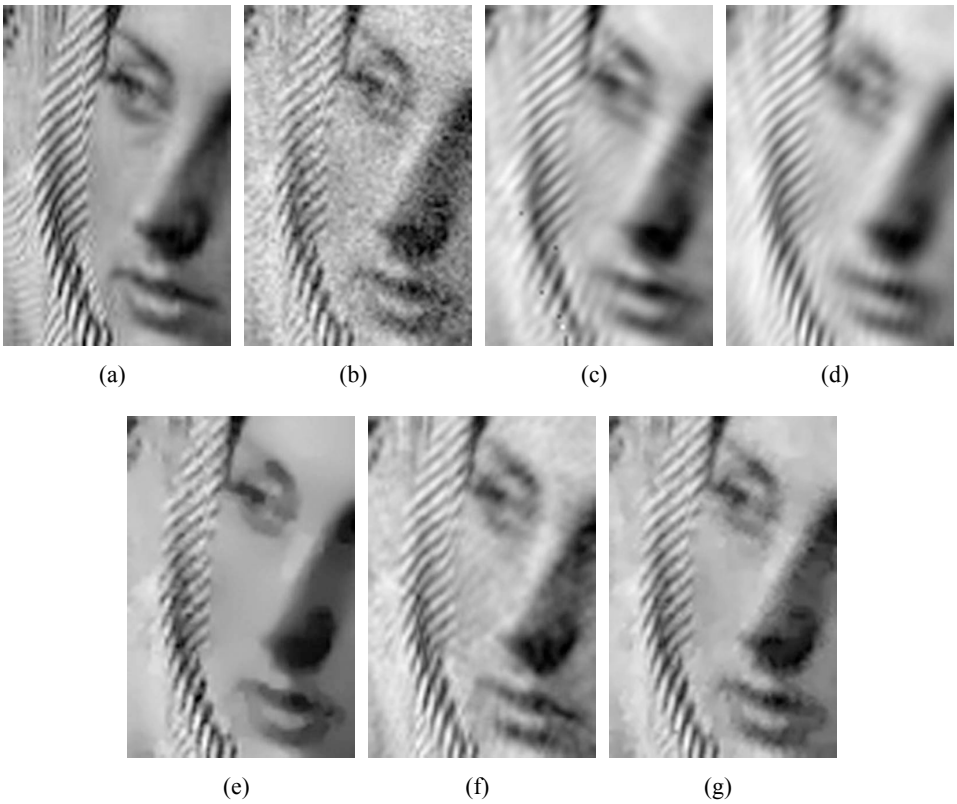


Figure 12: Close up on the Barbara face. (a) Original image, (b) version with noise, (c) generated by curvelets, (d) by wave atoms, (e) by diffusion with adaptive-fidelity term, (f) by TV diffusion combined with curvelet, (g) by proposed model.

In order to improve the cognitive analysis with regard to results of Figure 11, Fig.12 shows a close up of Barbara's right-eye, which is extracted from Fig.11(b)-(f). We can clearly see that in the images 12(c), 12(d) and 12(f), the texture is erroneously extended to the face. Furthermore, Fig.12(e) is quite smoothed, in addition to the loss of pieces of the oriented texture. Finally, in Fig.12(g) both texture and intrinsic details contained on the face are recovered.

6 Conclusions

In this work we propose a new Partial Differential Equation directed for denoising high-detail images contaminated by Gaussian noise. The current method combines anisotropic diffusion filtering with harmonic analysis methodologies, specifically by means of wave atom systems. Moreover, we propose a new regularization mechanism which can detect complex oscillatory patterns on the image. The proposed approach works in a selective way; that is: it applies diffusion while recovering edges and textured parts in accordance with a mapping factor (the regularization term) and by the action of forcing terms. Regularized PDE for treating textured images may be used on a whole range of practical applications; for instance: *finger-*

print processing, *engineering surfaces*, *geophysical data*, *inpainting*, etc. Experimental results show the efficiency of our method when compared to recent techniques in the literature.

References

Alvarez, L., Lions, P., & Morel, J. (1992). Image selective smoothing and edge detection by nonlinear diffusion. SIAM J. Numer. Anal., 29, 845–866.

Barcelos, C. A. Z., Boaventura, M., & Jr, E. S. (2003). A well-balanced flow equation for noise removal and edge detection. IEEE Transactions on Image Processing, 12, 751–763.

Burger, M., Gilboa, G., Osher, S., & Xu, J. (2006). Nonlinear inverse scale space methods. Communications in Mathematical Sciences, 4, 179–212.

Cand□s, E. J. & Donoho, D. L. (2004). New tight frames of curvelets and optimal representations of objects with piecewise c^2 singularities. Communications on Pure and Applied Mathematics, 57, 219–266.

Cand□s, E., Demanet, L., Donoho, D., & Ying, L. (2006). Fast discrete curvelet transforms. SIAM J. on Multiscale Model. Simul., 5, 861–899.

Casaca, W. C. O. & Boaventura, M. (2009). A regularized nonlinear diffusion approach for texture image denoising. In XXII Brazilian Symposium on Computer Graphics and Image Processing (Sibgrapi) (pp. 164–171).: IEEE Computer Society.

Casaca, W. C. O. & Boaventura, M. (2010). A decomposition and noise removal method combining diffusion equation and wave atoms for textured images. Mathematical Problems in Engineering, 2010, 1–21.

Catté, F., Lions, P., Morel, J., & Coll, T. (1992). Image selective smoothing and edge detection by nonlinear diffusion. SIAM J. Numer. Anal., 29, 182–193.

Chan, T. F., Osher, S., & Shen, J. (2001). The digital tv filter and nonlinear denoising. IEEE Transactions on Image Processing, 10, 231–241.

Crandall, M. G., Ishii, H., & Lions, P. L. (1992). User's guide to viscosity solutions of second order partial differential equations. Bulletin of the American Mathematical Society, 27, 1–67.

Demanet, L. & Ying, L. (2007a). Curvelets and wave atoms for mirror-extended images. In Proc. SPIE Wavelets XII conf, volume 6701 (pp. 67010J).

Demanet, L. & Ying, L. (2007b). Wave atoms and sparsity of oscillatory patterns. Appl. Comput. Harmon. Anal., 23, 368–387.

Evans, L. D. (2010). Partial Differential Equations. AMS.

Gilboa, G., Sochen, N., & Zeevi, Y. Z. (2006). Variational denoising of partly textured images by spatially varying constraints. IEEE Transactions on Image Processing, 15, 2281–2289.

Gilles, J. (2007). Noisy image decomposition: a new structure, texture and noise model based on local adaptivity. Journal of Mathematical Imaging and Vision, 22, 285–295.

Goldberg, D. E. (1989). Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Longman Publishing Co., Inc.

Holland, J. H. (1975). Adaptation in Natural and Artificial Systems. University of Michigan Press.

Li, F., Shen, C., Shen, C., & Zhang, G. (2009). Variational denoising of partly textured images. Journal of Visual Communication and Image Representation, 20, 293–200.

Liu, G. (2011). Variational image decomposition using wave atoms. International Conference on Optoelectronics and Image Processing (ICOIP'10), 1, 588–591.

Liu, G., Feng, X., & Bai, J. (2008). Variational image decomposition using wave atoms. Current Development in Theory and Applications of Wavelets, 2, 277–291.

Ma, J. & Plonka, G. (2007). Combined curvelet shrinkage and nonlinear anisotropic diffusion. IEEE Transactions on Image Processing, 16, 2198–2206.

Mishra, A., Wong, A., Clausi, D. A., & Fieguth, P. W. (2010). Quasi-random nonlinear scale space. Pattern Recogn. Lett., 31, 1850–1859.

Nordstrom, K. (1991). Biased anisotropic diffusion: a unified regularization and diffusion approach to edge detection. Image Vision Comput., 8, 318–327.

Osher, S., Burger, M., Goldfarb, D., Xu, J., & Yin, W. (2005). An iterative regularization method for total variation-based image restoration. SIAM Multiscale Model. and Simu., 4, 460–489.

Perona, P. & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell., 12, 629–639.

Plonka, G. & Ma, J. (2008). Nonlinear regularized reaction-diffusion filters for denoising of images with textures. IEEE Transactions on Image Processing, 17, 1283–1294.

Ross, S. M. (2009). Introduction to Probability Models. Academic Press; 10 edition.

Rudin, L., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. Physica D, 60, 259–268.

Starck, J. L., Candès, E. J., & Donoho, D. L. (2002). The curvelet transform for image denoising. IEE Transactions on Image Processing, 17, 670–684.

The MathWorks, I. (2011). Matlab optimization toolbox user's guide (r2011b).