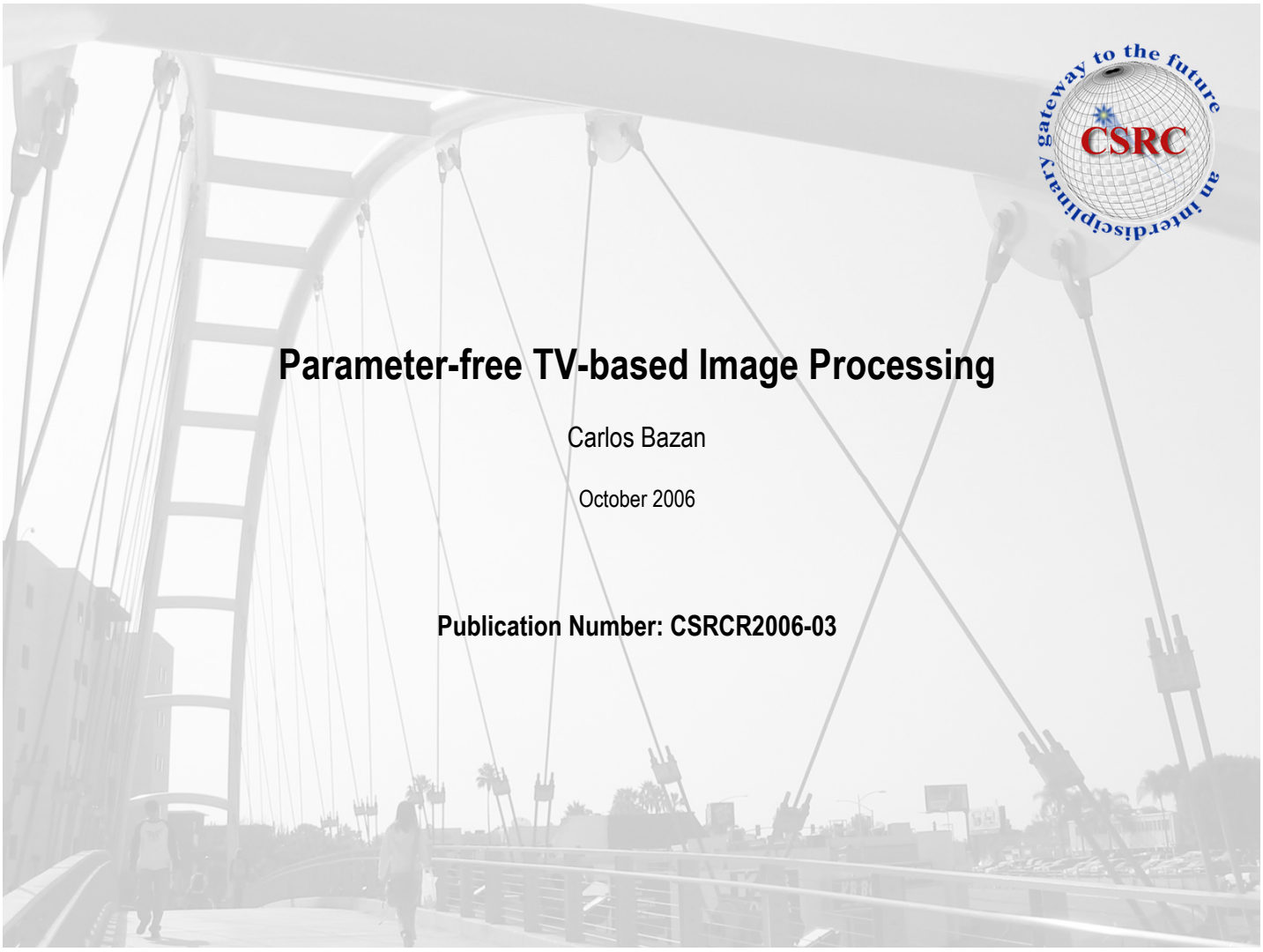


Parameter-free TV-based Image Processing

Carlos Bazan

October 2006

Publication Number: CSRCR2006-03

Computational Science & Engineering Faculty and Students Research Articles

Database Powered by the Computational Science Research Center Computing Group

COMPUTATIONAL SCIENCE & ENGINEERING

SAN DIEGO STATE UNIVERSITY

Computational Science Research Center
College of Sciences
5500 Campanile Drive
San Diego, CA 92182-1245
(619) 594-3430

Qualifying Exam Report

Carlos Bazan

Committee

Dr. Peter Blomgren, Chair

Dr. Robert Mellors

Dr. Ali Nadim

Dr. Alpan Raval

Dr. Peter Salamon

October 17, 2006

Index

0. Introduction

In this report we review some of the most influential studies on PDE-based approaches to image processing. It will illustrate the state of the art in the field and set the appropriate background for the study of dynamic and parameter-free total variation based image processing. Our work builds on the efforts of many scientists who contributed to the development of this fascinating field of image processing.

Mathematically any image can be represented as a function $u_0 : \Omega \to \mathbb{R}^N$, where Ω is a bounded subset of $\mathbb{R}^2$ (the domain). The integer N represents the number of color channels in the image [Xu06]. *I.e.* for $N = 1$ we have a grayscale image while for $N = 3$ we have a color image. In practice, images that we observe are seldom the true or clean images. They get perturbed by both noise and blur, thus the observed image u_0 can be decomposed as $u_0 = u + \eta$ or $u_0 = K * u + \eta$, where u denotes the true image, K represents a convolution kernel (blur), and η represents some additive noise (usually modeled as Gaussian white noise).

The main objective of this research is to devise efficient ways for decomposing the observed image as $u_0 = u + \eta$, whereby we can obtain a good approximation to the true image u. We propose an iterative method that will allow image reconstruction with no human intervention. This approach targets the incorporation of the method into imaging machines, *e.g.* ultrasound, MRI, electron microscopy, etc.

1. Physical Background

The diffusion process (mass transfer) is the movement of matter from high concentration to low concentration regions. The equilibrium property is expressed by Fick's first law [Co05]:

$$q = -D \cdot \nabla u \,. \tag{1.1}$$

The concentration gradient ∇u generates a flux q which attempts to compensate for the gradient. The relation between the gradient ∇u and the flux q is described by a positive definite symmetric matrix D, the diffusion tensor. For the isotropic case, one can replace the diffusion tensor with a scalar-valued − the diffusivity d − that describes the diffusion rate. For the most general anisotropic case, one will have to use the full diffusion tensor D.

The phenomenon described above represents the transport of mass without creating or destroying any mass. Therefore, one can state the following continuity (conservation) equation

$$\partial_t u = -\nabla \cdot q \,. \tag{1.2}$$

Substituting (1.1) in (1.2) yields the diffusion equation

$$\partial_t u = \nabla \cdot \left(D \cdot \nabla u \right). \tag{1.3}$$

In the context of (single channel) image processing the concentration represents the values of the amplitude of the image's gray-level intensities or tones of gray. The diffusion tensor D (or diffusivity d) is commonly a function of the concentration u and/or its derivatives, which leads to a nonlinear (isotropic or anisotropic) diffusion process [We97].

2. Linear Diffusion Process

Given a grayscale image u_0 represented by a real-valued mapping $u_0 \in L^1\left(\mathbb{R}^2\right)$, a classic way of smoothing the image u_0 is by applying a convolution of the form [GW02]

$$\nabla G_\sigma * u_0(x) = \int_{\mathbb{R}^2} \nabla G_\sigma(x-\xi) u_0(\xi)\, d\xi\,, \tag{2.1}$$

where $G_\sigma \in C^\infty\left(\mathbb{R}^2\right)$ is a smooth kernel, $\int_{\mathbb{R}^2} G_\sigma(x)\,dx = 1$, $\int_{\mathbb{R}^2} \nabla G_\sigma dx \le C_\sigma$, $G_\sigma(x) \to \delta_x$ for $\sigma \to 0$, and where δ_x is the Dirac delta at point x. Usually G_σ will be a two-dimensional Gaussian of width (standard deviation) $\sigma > 0$:

$$G_\sigma = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-|x|^2}{2\sigma^2}\right). \tag{2.2}$$

This approach is the simplest and most well known method for treating an image, and it has proven to have excellent smoothing characteristics [WS00]. We can observe in Figure 2.1 that the Gaussian smoothing process acts as a low-pass filter that suppresses the high frequencies of the image as σ increases.

(a) Grayscale true image and its contours.

(b) Convolved image with Gaussian of width $\sigma = 0.5$ and its contours.

(c) Convolved image with Gaussian of width $\sigma = 1.0$ and its contours

(d) Convolved image with Gaussian of width $\sigma = 1.5$ and its contours

Figure 2.1: Gaussian smoothing process. **(a)** Grayscale true image and its contours. **(b)** Convolved image with Gaussian of width $\sigma = 0.5$. **(c)** Convolved image with Gaussian of width $\sigma = 1.0$. **(d)** Convolved image with Gaussian of width $\sigma = 1.5$. *Original image courtesy of Dan Massey*

It is interesting to note that this convolution process is intimately related to solving a parabolic partial differential equation [Wi83] [Ko84] [Hu86]. It can be found in partial differential equations textbooks [BJS64] [He64] that the following linear heat equation

$$\partial_t u - \nabla^2 u = 0, \tag{2.3}$$

for an initial condition $u(0,x) = u_0(x)$ where $u_0 \in C\left(\mathbb{R}^2\right)$, has a unique solution

$$u(t,x) = \begin{cases} u_0(x) & t = 0 \\ \left(G_{\sqrt{2t}} * u_0\right)(x) & t > 0, \end{cases} \tag{2.4}$$

provided that the function satisfies

$$\left|u(t,x)\right| \le M \exp\left(a\left|x\right|^2\right) \qquad M > 0,\ a > 0, \tag{2.5}$$

that it depends continuously on the initial condition u_0 with respect to $\left\|\cdot\right\|_{L_\infty\left(\mathbb{R}^2\right)}$, and that it meets the maximum-minimum principle [We96a]

$$\inf_{\mathbb{R}^2} u_0 \le u(t,x) \le \sup_{\mathbb{R}^2} u_0 \qquad \text{on} \qquad \mathbb{R}^2 \times [0,\infty). \tag{2.6}$$

This similarity allows us to solve the diffusion equation (2.3) and obtain similar results to the Gaussian filtering, as long as we use a constant diffusion coefficient $d = 1$ and stop the diffusion process when we reach the scale state $t = 1/2\sigma^2$. In this context, the first attempts to use PDEs for image processing were conducted by Witkin [Wi83] and Koenderink [Ko84].

The simplicity and effectiveness of the Gaussian smoothing makes it an attractive tool for image noise removal. However, it also presents at least a couple of serious drawbacks: (i) the first main issue relates to the fact that Gaussian smoothing does not only smooth the noise but it also smoothes everything else along with it; and (ii) the second relates to the scale-space property, where Gaussian smoothing tends to dislocate edges when one moves from a finer to a coarser scale [Wi83] [Wi84]. Most of the shortcomings of linear diffusion processes can be avoided through inhomogeneous (or anisotropic) diffusion and nonlinear diffusion models. The

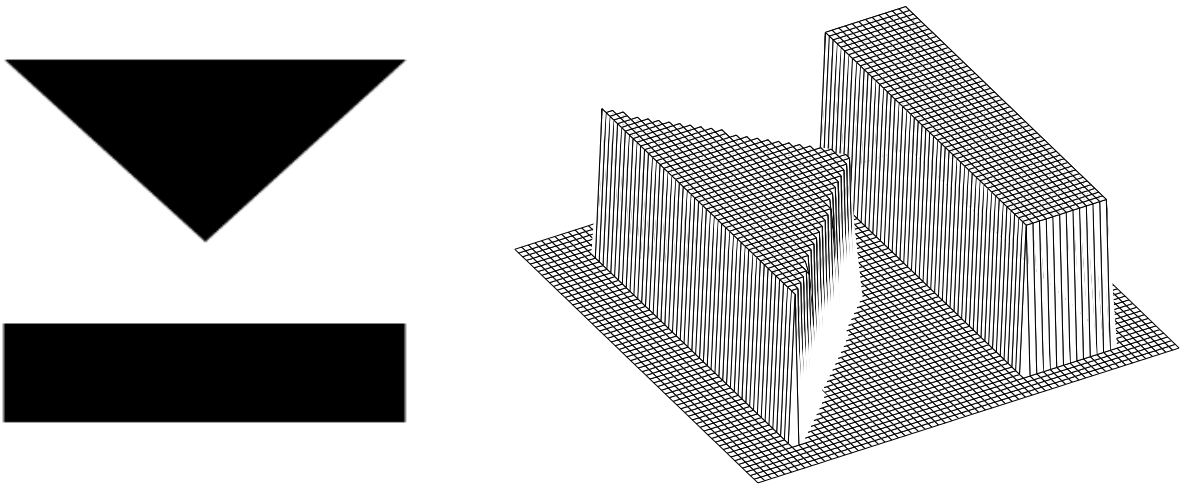
latter is the subject of the next section, while one of the earliest results of the former approach is due to Fritsch *et al* [FPC91] [Fr92]. They apply a simple way for including *a priori* knowledge about the image. It uses an edge detector based on the gradient of the observed image $\|\nabla u_0\|$, *i.e.* locations with large $\|\nabla u_0\|$ have a high likelihood to be an edge, and one can reduce the diffusivity for large values of $\|\nabla u_0\|$. Charbonnier *et al* [CBAB94] propose to use the following inhomogeneous (isotropic) diffusion term, where λ is a free parameter

$$g\left(\|\nabla u_0\|^2\right) = \frac{1}{\sqrt{1+\|\nabla u_0\|^2/\lambda^2}}, \qquad \lambda > 0\,. \tag{2.7}$$

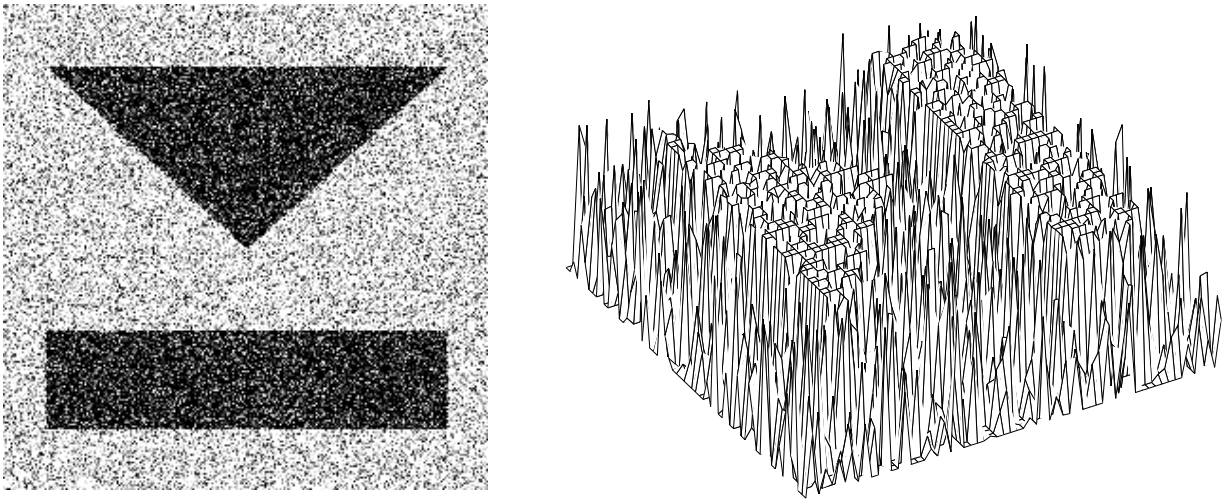
Note that the diffusion equation will remain linear despite the inhomogeneous (pixel-wise) diffusion process. Note also that the restriction $\lambda > 0$ is unnecessary provided that the parameter λ is sufficiently away from zero. In Figure 2.2 we compare Fritsch's model to the Gaussian smoothing process. For measuring the noise applied to the images we use the following definition of signal-to-noise (SNR) ratio

$$SNR(u_0) = 20\log_{10}\left(\frac{\|u-\bar{u}\|}{\|\eta-\bar{\eta}\|}\right),$$

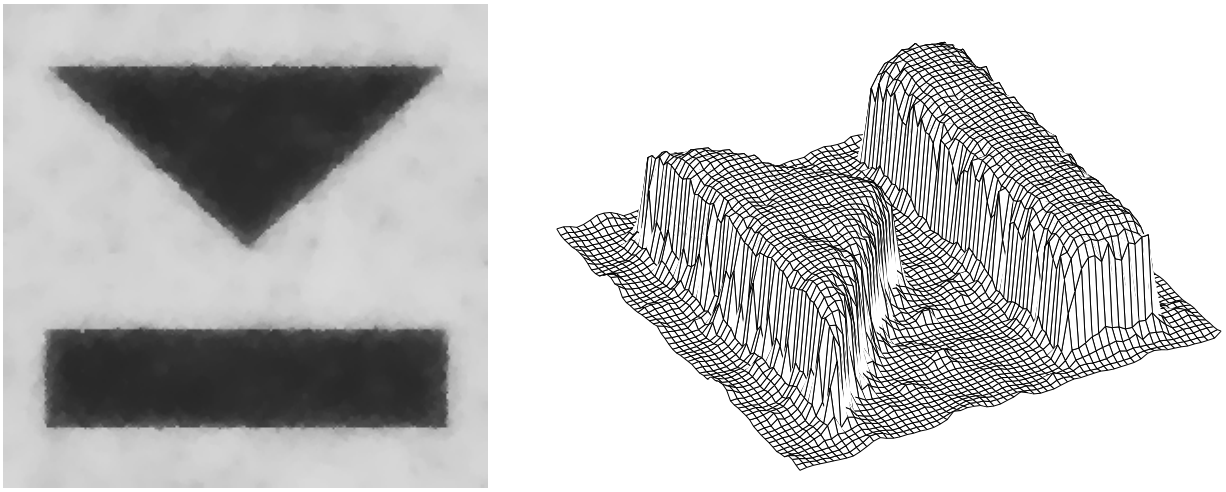
where $\bar{u}$ and $\bar{\eta}$ are the averages of the grayscale intensities of the true image u and the noise η respectively.



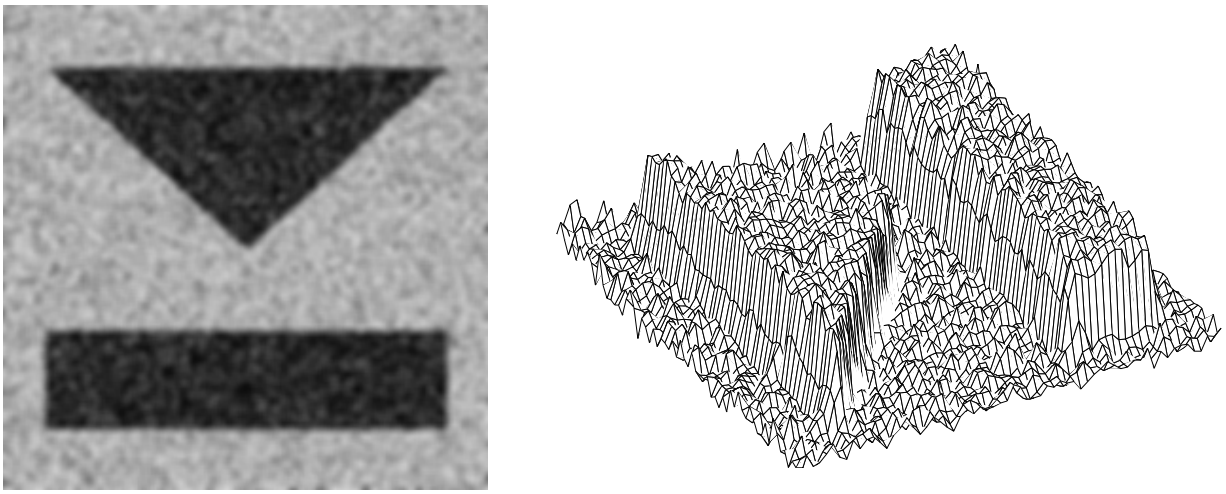
(a) Binary synthetic true image and its elevation.



(b) Noisy image with $SNR = 15.0$ dB and its elevation.



(c) Treated image with Fritsch/Charbonnier model and parameter $\lambda = 1.0$ and its elevation.



(d) Convolved image with Gaussian of width $\sigma = 2.0$ and its elevation.

Figure 2.2: Comparison of Gaussian smoothing and Fritsch's inhomogeneous diffusion. **(a)** Binary synthetic true image and it elevation. **(b)** Noisy image with $SNR = 15.0$ dB and its elevation. **(c)** Treated image with Fritsch/Charbonnier model and parameter $\lambda = 1.0$ and its elevation. **(d)** Convolved image with Gaussian of width $\sigma = 2.0$ and its elevation.

3. Nonlinear Diffusion Process

The first attempt to derive a model that incorporates current local information from an image within a PDE framework was conducted by Perona and Malik [PM90]. They propose a nonlinear diffusion model (which they call 'anisotropic') in order to avoid the blurring of edges and other localization problems presented by linear diffusion models. The model accomplishes this by applying a process that reduces the diffusivity in places having higher likelihood of being edges. This likelihood is measured by a function of the current local gradient $\|\nabla u\|^2$. The model can be written as

$$\partial_t u - \nabla \cdot \left(g\left(\|\nabla u\|^2 \right) \nabla u \right) = 0 . \tag{3.1}$$

In this model the diffusivity has to be such that $g\left(\|\nabla u\|^2 \right) \to 0$ when $\|\nabla u\|^2 \to \infty$ and $g\left(\|\nabla u\|^2 \right) \to 1$ when $\|\nabla u\|^2 \to 0$. One of the diffusivities Perona and Malik propose is

$$g\left(\|\nabla u\|^2 \right) = \frac{1}{1 + \|\nabla u\|^2 / \lambda^2} \qquad \lambda > 0 . \tag{3.2}$$

Needless to say, despite the name 'anisotropic', the model uses a scalar-valued diffusivity and not a diffusion tensor. The behavior of the above diffusivity (3.2) is depicted in Figure 3.1 for various values of the parameter λ. We can see that $0 \le g\left(\|\nabla u\|^2 \right) \le 1$ and that it goes to zero very rapidly for small values of the parameter λ. In practice, the higher the gradient (*i.e.* near edges), the lower the diffusion for all values of λ. Now, the higher the parameter λ, the higher the diffusivity for the same gradient. Notice that, since the parameter λ is squared in the diffusivity term, it can take on negative values, as long as its absolute value is reasonably greater than zero.

It has been argued in [We96a] that the parameter λ plays the role of a contrast parameter separating forward $|u_x| \le \lambda$ (low contrast) from backward $|u_x| > \lambda$ (high contrast) diffusion areas. *I.e.* locations with $|u_x| > \lambda$ are regarded as edges where the smoothing process is inhibited, while locations with $|u_x| \le \lambda$ are assumed to belong to the interior of a segment [WS00]. And that this fact is the main contributor to the good restoration properties of this model.

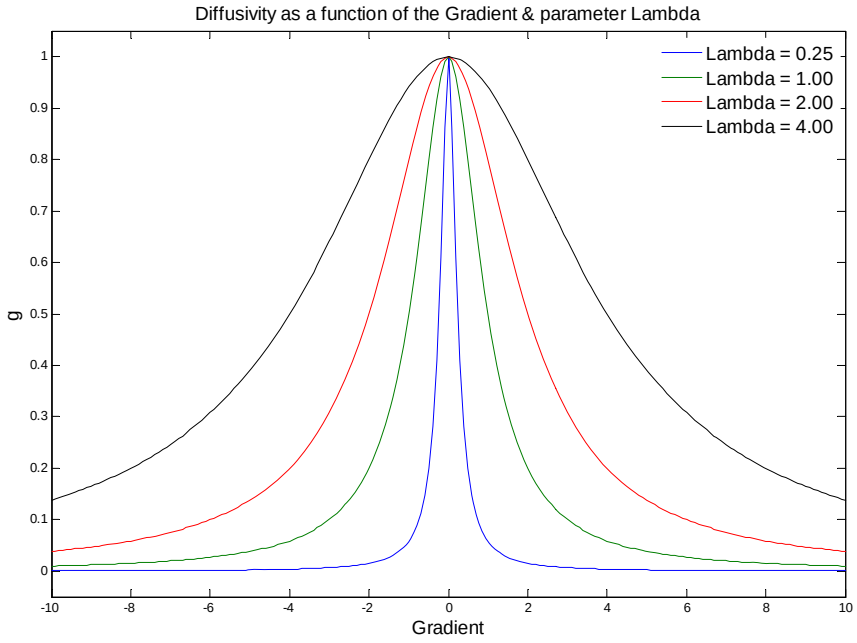
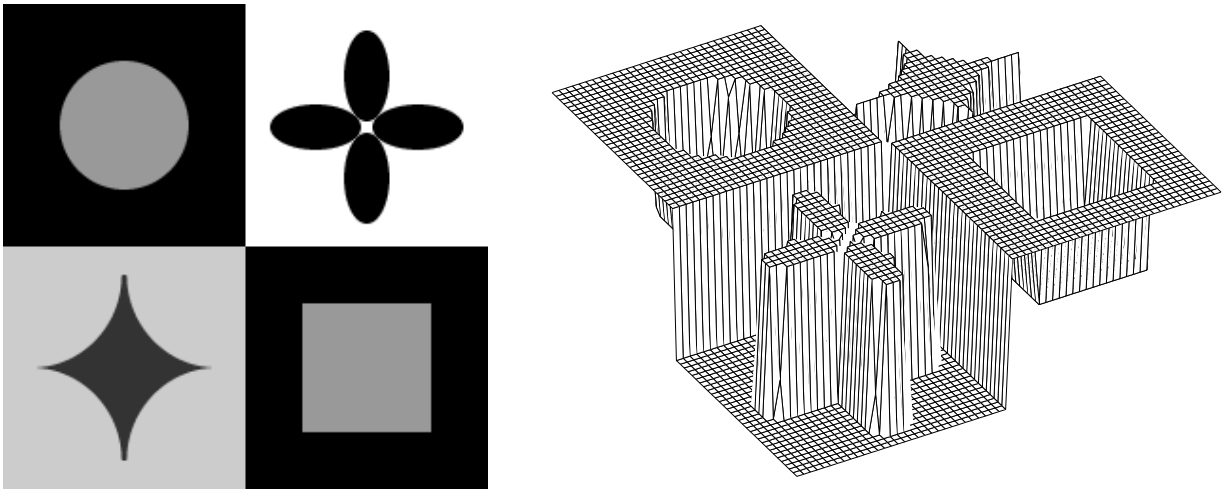


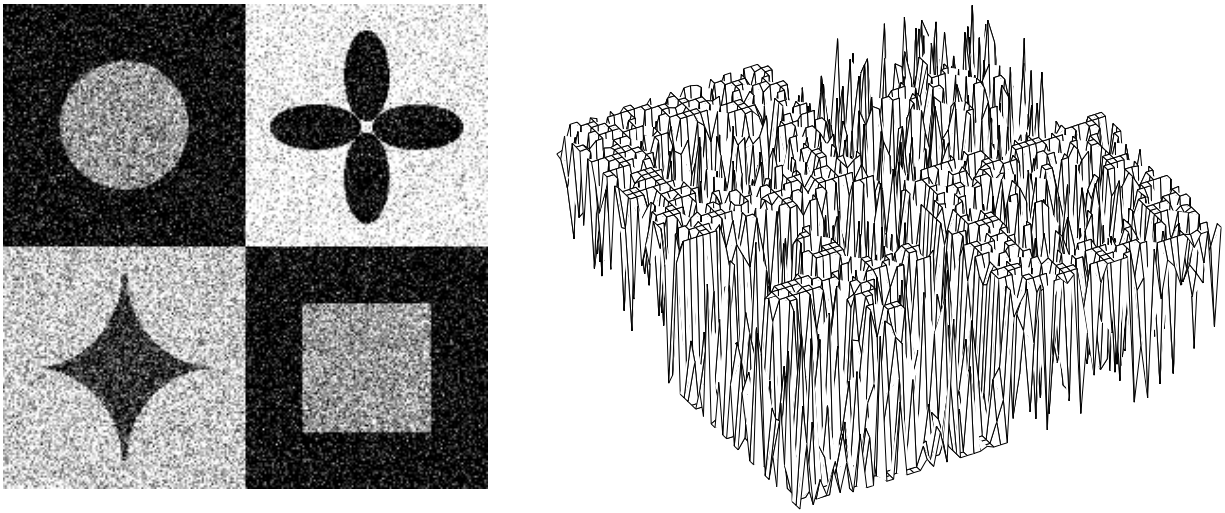
Figure 3.1: Diffusivity (3.2) for various values of the parameter λ.

The results obtain by Perona and Malik are visually very impressive, and it has awaken the interest of the image processing community towards the incorporation of current local information in the image processing models. The model accomplishes the long sought effect of blurring small fluctuations (possible noise) while sharpening edges. Results using the Perona-Malik model for various values of the parameter λ are shown in Figure 3.2. Despite the practical success of the Perona-Malik model, it presents some serious theoretical problems: (i) None of the classical well-posedness frameworks is applicable to the Perona-Malik model, *i.e.*

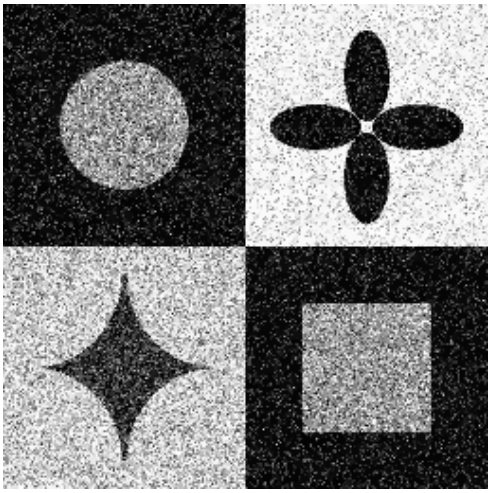
we can not ensure well-posedness results [WS00] [NS92]; (ii) Uniqueness and stability with respect to the initial image should not be expected, *i.e.* solvability is a difficult problem, in general [Ki97] [HN83] [Ho83] [PSM94] [CLMC92]; (iii) The regularizing effect of the discretization plays too much of an important role in the solution [FW94] [Be94]. The latter is perhaps the key element in the success or failure of the model. Most practical applications work very well provided that the numerical schemes stabilize the process through some implicit regularization.

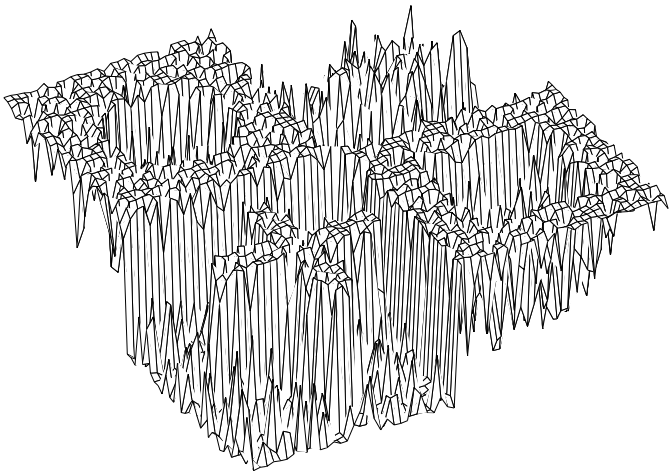


(a) Synthetic grayscale true image and its elevation.

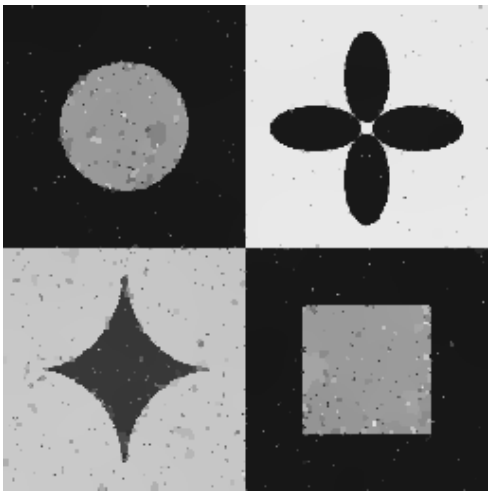


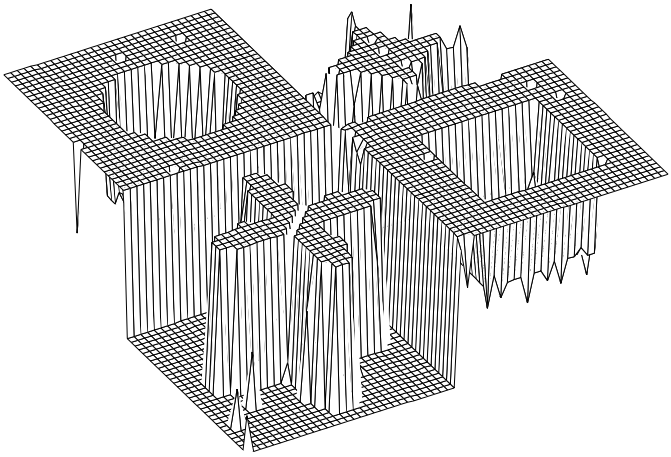
(b) Noisy image with $SNR = 18.0$ dB and its elevation.



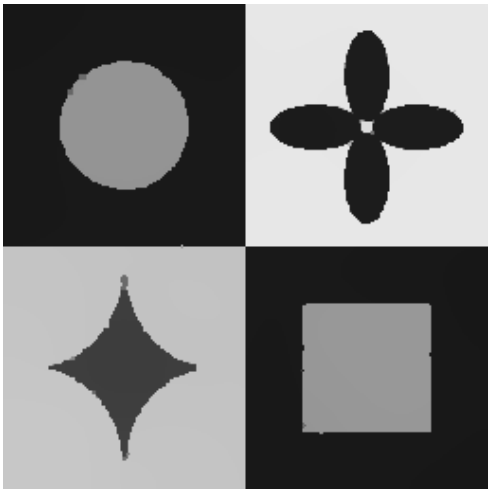


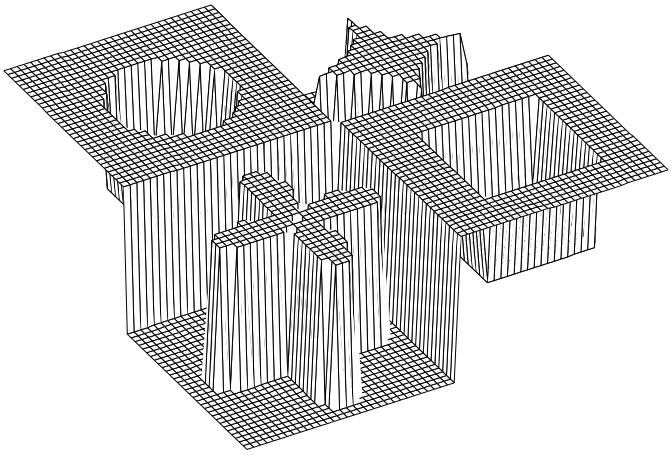
(c) Treated image with parameter $\lambda = 0.25$ and its elevation.



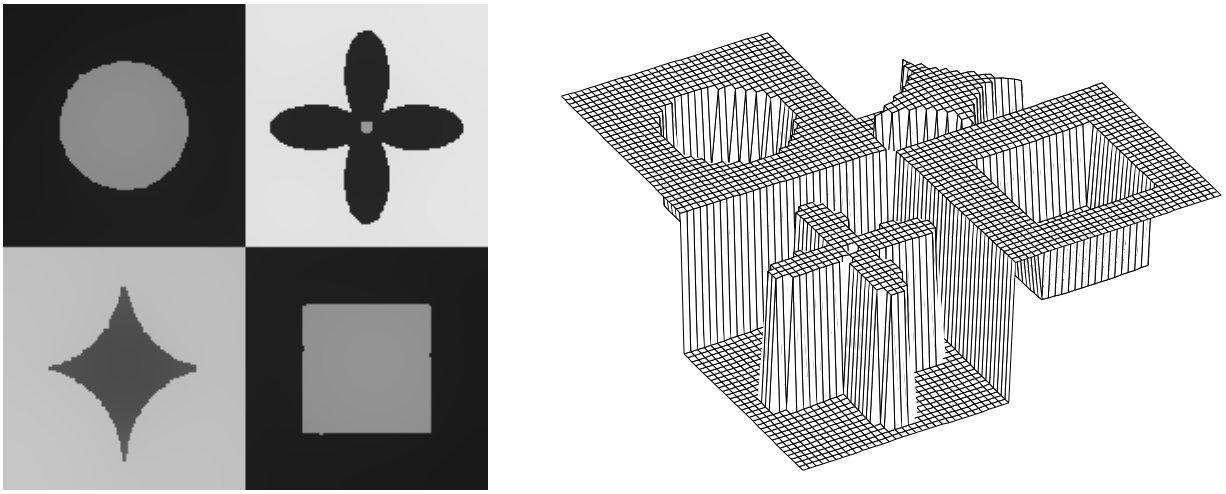


(d) Treated image with parameter $\lambda = 1.0$ and its elevation.





(e) Treated image with parameter $\lambda = 2.0$ and its elevation.



(f) Treated image with parameter $\lambda = 4.0$ and its elevation.

Figure 3.2: Perona-Malik model for various values of the parameter λ. **(a)** Synthetic grayscale true image and its elevation. **(b)** Noisy image with $SNR = 18.0$ dB and its elevation. **(c)** Treated image with parameter $\lambda = 0.25$ and its elevation. **(d)** Treated image with parameter $\lambda = 1.0$ and its elevation. **(e)** Treated image with parameter $\lambda = 2.0$ and its elevation. **(f)** Treated image with parameter $\lambda = 4.0$ and its elevation.

This observation motivated much research towards the introduction of the regularization directly into the PDE to avoid the dependence on the numerical schemes [CLMC92] [NS92]. A variety of spatial, spatio-temporal, and temporal regularization procedures have been proposed over the years [BBD93] [CLMC92] [We01a] [We96b] [WP93] [LC94]. The one that has attracted much attention is the mathematically sound formulation due to Catté, Lions, Morel and Coll [CLMC92]. They propose to replace the diffusivity $g\left(\|\nabla u\|^2\right)$ of the Perona-Malik model by a slight variation $g\left(\|\nabla u_\sigma\|^2\right)$ with $u_\sigma = G_\sigma * u$, where G_σ is the smooth kernel described in Section 2. The proposed model is therefore

$$\partial_t u - \nabla \cdot \left(g\left(\|\nabla u_\sigma\|^2\right) \nabla u \right) = 0 . \tag{3.3}$$

We should note that this spatial regularization model belongs to a class of well-posed problems (existence and uniqueness were proven in [CLMC92]), and that its successful implementation is contingent to the choosing of an appropriate value for the additional regularization parameter σ.

Whitaker and Pizer [WP93] and Li and Chen [LC94] suggest making the parameters σ and λ time-dependent, and Benhamouda [Be94] performs a systematic study of the influence of these parameters for the one-dimensional case.

As for the diffusivity $g\left(\left\|\nabla u_\sigma\right\|^2\right)$, it has to be chosen as a (rapidly) decreasing function of the edge detector $\left\|\nabla u_\sigma\right\|^2$. Weickert [We01a] proposes to use

$$g\left(\left\|\nabla u_\sigma\right\|^2\right) = \begin{cases} 1 & \left\|\nabla u_\sigma\right\|^2 = 0 \\ 1 - \exp\left(\dfrac{-c}{\left(\left\|\nabla u_\sigma\right\| / \lambda\right)^8}\right) & \left\|\nabla u_\sigma\right\|^2 > 0, \end{cases} \tag{3.4}$$

for which he argues that good results are obtained by choosing $c \approx 3.315$. After some time this filter creates segmentation-like results which are piecewise almost constant [We01b]. For $t \to \infty$, however, the image becomes completely flat [We98]. Well-posedness results for this filter can be found in [CLMC92] [We98] and a scale-space interpretation in terms of an extremum principle as well as decreasing variance, decreasing energy, and increasing entropy is given in [We98].

A different way of introducing regularization to the Perona-Malik model is through anisotropic diffusion. We should recall that despite its name, the Perona-Malik model is actually (inhomogeneous) isotropic, since it utilizes a scalar-valued diffusivity g which is adapted to the underlying image structure. The main advantage of anisotropic diffusion models over their isotropic counterparts is that they not only account for the modulus of the edge detector, but also its direction. Isotropic diffusion will inhibit diffusion near edges, making it hard to eliminate noise near them. Anisotropic diffusion, in the other hand, will allow diffusion along edges while avoiding diffusing perpendicular to them. The basic idea behind anisotropic models is to

construct the orthogonal system of eigenvectors v_1, v_2 of the diffusion tensor D in such way that they will reveal the presence of edges, *i.e.* $v_1 \perp \nabla u_\sigma$ (perpendicular) and $v_2 \| \nabla u_\sigma$ (parallel). And choose appropriate corresponding eigenvalues that will allow smoothing along the edges and avoid doing so across them. *I.e.* the diffusion tensor D steers the diffusion process in such a way that the eigenvectors prescribe the diffusion directions and the corresponding eigenvalues determine the amount of diffusion along these directions. In this context, Weickert [We96b] [We94] proposes to choose the corresponding eigenvalues λ_1 and λ_2 as

$$\lambda_1 = g\left(\left\|\nabla u_\sigma\right\|^2\right), \qquad \lambda_2 = 1. \tag{3.5}$$

While Cottet and Germain [CG93] propose constructing a model with corresponding eigenvalues

$$\lambda_1 = 0, \qquad \lambda_2 = \frac{\gamma\left\|\nabla u_\sigma\right\|^2}{1+\left(\left\|\nabla u_\sigma\right\|/\sigma\right)^2} \qquad \eta > 0, \tag{3.6}$$

which is a model that only diffuses along edges, making it very suitable for processing one-dimensional features. One should note that this model involves the choosing of an additional free parameter γ.

A very interesting variation of the Perona-Malik model has been proposed by Nordström [No90], who considers diffusion-reaction methods for the reconstruction of degraded images. Such an approach leads to Euler-Lagrange equations of the form

$$\partial_t u - \nabla\cdot\left(g\left(\left\|\nabla u_\sigma\right\|^2\right)\nabla u\right) = \beta\left(u_0 - u\right), \tag{3.7}$$

which is identical to the model we have been considering with an additional bias term $\beta\left(u_0 - u\right)$. In principle, the bias term, should spare the user from choosing an stopping time [1] (which is a

[1] In practice, the bias term forces the steady-state solution to stay close to the observed image.

function of σ), by the choosing of an additional free parameter β. This type of diffusion-reaction models have been study further and improved upon in [GY91] and [Sc94].

4. Total Variation Based Methods

Rudin, Osher and Fatemi [ROF92] propose image noise removal by minimizing the total variation (TV) norm of the estimated solution. They derive a constrained minimization algorithm as a time-dependent nonlinear PDE, where the constraints are determined by the noise statistics. They state that the space of bounded total variation is the proper class for many basic image processing tasks. Thus, the restored image is the solution of

$$\min_u \int_\Omega \|\nabla u\| \, dx, \tag{4.1}$$

subject to the following constraint involving the observed image u_0

$$\frac{1}{2}\int_\Omega \left(u - u_0\right)^2 dx = \sigma^2 . \tag{4.2}$$

This constraint uses *a priori* information that the standard deviation of the noise is σ (it is also assumed that the noise is normally distributed with mean zero, *i.e.* $\int u \, dx = \int u_0 \, dx$). In most practical cases this parameter will not be known and the success of the method will require a good estimate of its value.

To solve this minimization problem, one would usually solve its Euler-Lagrange equation, namely

$$-\nabla \cdot \left(\frac{\nabla u}{\|\nabla u\|} \right) + \lambda \left(u - u_0\right) = 0, \qquad \text{in } \Omega \tag{4.3}$$

$$\partial_n u = 0, \qquad \text{on } \partial\Omega . \tag{4.4}$$

The solution procedure proposed in [ROF92] uses a parabolic equation with time as an evolution (scale) parameter, or equivalently, the gradient descent method. This is

$$\partial_t u - \nabla \cdot \left(\frac{\nabla u}{\|\nabla u\|} \right) + \lambda \left(u - u_0\right) = 0, \qquad \text{in } \Omega, \text{ for } t > 0 \tag{4.5}$$

$$\partial_n u = 0, \qquad \text{on } \partial\Omega, \tag{4.6}$$

and $u(0,x) = u_0$ is the observed image used as initial condition. For the parameter λ they suggest a dynamic value $\lambda(t)$ estimated by Rosen's gradient-projection method, which as $t \to \infty$ converges to (with a little abuse in notation)

$$\lambda = -\frac{1}{2\sigma^2}\int_\Omega \left[\sqrt{u_x^2 + u_y^2} - \left(\frac{(u_0)_x u_x}{\sqrt{u_x^2 + u_y^2}} + \frac{(u_0)_y u_y}{\sqrt{u_x^2 + u_y^2}}\right)\right] dx\,. \tag{4.7}$$

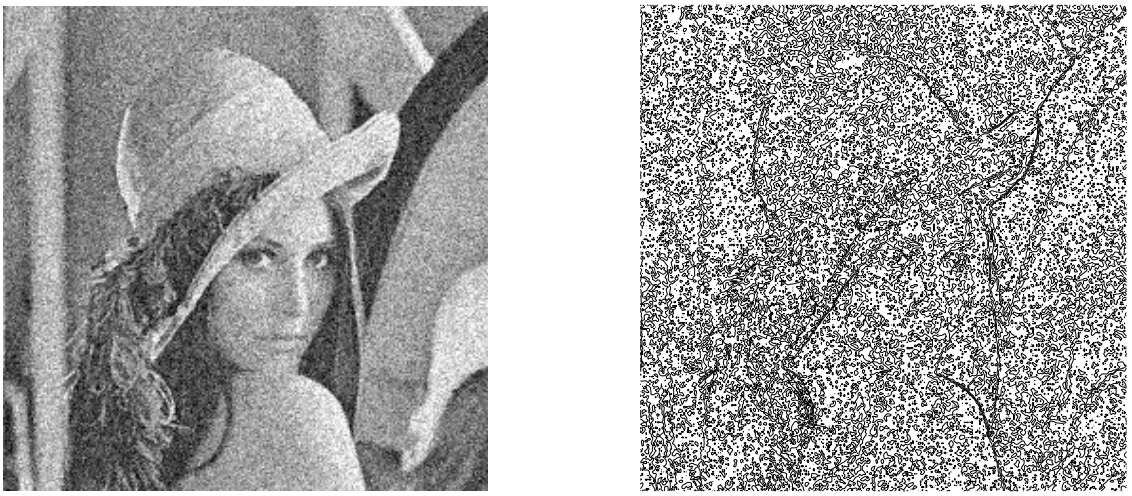
Existence and uniqueness results for this constrained nonlinear PDE have been obtained by Lions, Osher and Rudin [LOR93].

This evolution scheme is not trivial to solve since it is highly nonlinear and not well-posed in strong sense [So03]. It also can run into trouble when $\nabla u \to 0$ beyond machine accuracy. When the scheme converges it does it at a linear rate. Further, direct application of classical schemes, *e.g.* affine invariant form of the damped Newton method as described in Deuflhard [De74] generally run into convergence problems due to the ill-conditioning of the problem introduced by the non-linearity. In practice it is common to use a slightly modified version of the TV norm [AV94]

$$\int_\Omega \sqrt{\|\nabla u\|^2 + \varepsilon}\, dx\,, \tag{4.8}$$

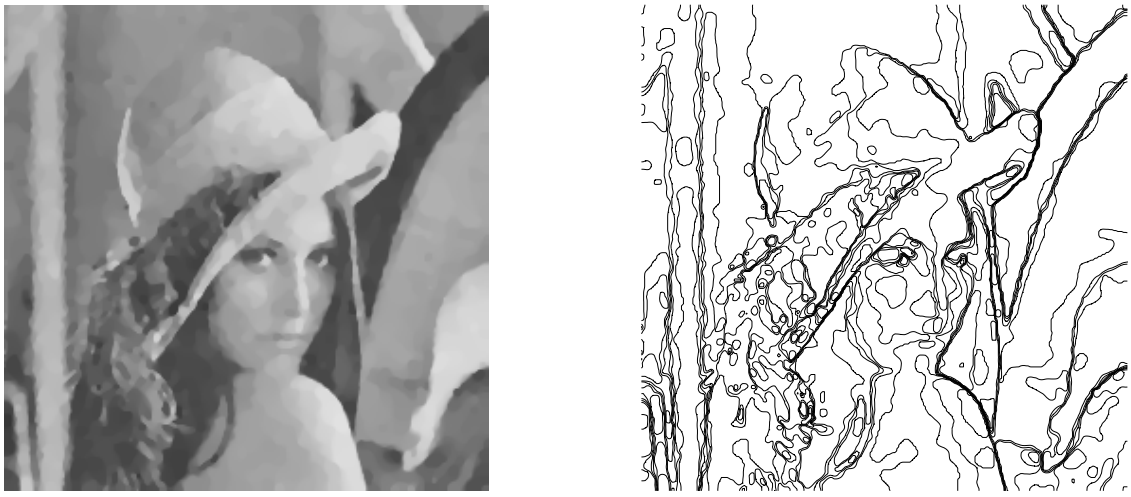
where ε is a small positive number which smoothes out the "corner" at $\|\nabla u\| = 0$. Also, when ε is very small, the Newton method does not work satisfactorily. Figure 4.1 depicts the implementation of Rudin *et al*'s model with the modification proposed in [AV94] for different values of the parameter σ.

(a) Grayscale true image and its contours.

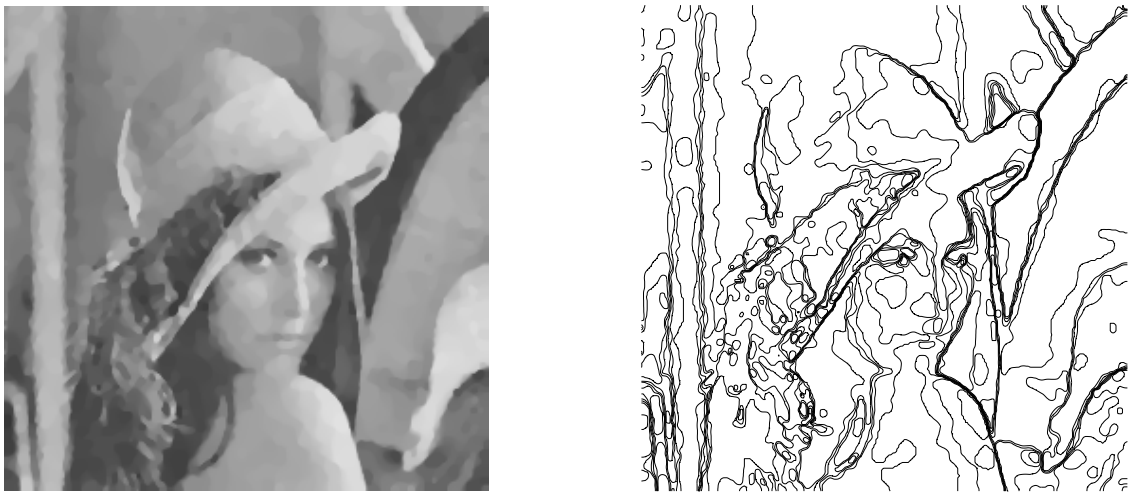


(b) Noisy image with $SNR = 22.1$ dB and its contours.

(c) Treated image with parameter $\sigma = 20$ and its contours.



(d) Treated image with parameter $\sigma = 40$ and its contours.



(e) Treated image with parameter $\sigma = 60$ and its contours.

Figure 4.1: Rudin-Osher-Fatemi model for various values of the parameter σ. **(a)** Grayscale true image and its contours. **(b)** Noisy image with $SNR = 22.1$ dB and its contours. **(c)** Treated image with parameter $\sigma = 20$ and its contours. **(d)** Treated image with parameter $\sigma = 40$ and its contours. **(e)** Treated image with parameter $\sigma = 60$ and its contorurs.

We observe that the images experience some loss of contrasts, which is one of the important limitations of the standard Rudin-Osher-Fatemi model [CE04]. This property was extensively studied by Strong and Chan [SC03].

To overcome the problems presented by the highly nonlinearity of the model Vogel and Oman [VO96] propose a fixed point lagged diffusivity iteration scheme,

$$-\nabla\cdot\left(\frac{\nabla u^{k+1}}{\left\|\nabla u^{k}\right\|}\right)+\lambda\left(u^{k+1}-u_0\right)=0\,. \tag{4.9}$$

This is a robust scheme but it is only linearly convergent. Golub, Chan and Mulet [GCM99] use interior-point primal-dual implicit method to solve the Euler-Lagrange equation by introducing a new dual variable

$$w=\frac{\nabla u}{\sqrt{\left\|\nabla u\right\|^2+\varepsilon}}, \tag{4.10}$$

and writing the problem as a system of nonlinear partial differential equations as follows

$$\begin{cases} -\nabla\cdot w+\lambda\left(u-u_0\right)=0 \\ w\sqrt{\left\|\nabla u\right\|^2+\varepsilon}-\nabla u=0 \end{cases}. \tag{4.11}$$

All these approaches, which attempt to solve the original TV-minimization problem, lead to solutions which exhibit the "staircase effect", *i.e.* a strong preference for piecewise constant patches.

Marquina and Osher [MO00] propose a different version of the transient parabolic equation that helps speed up the convergence of the time-marching scheme. The new evolution equation is

$$\partial_t u-\left\|\nabla u\right\|\nabla\cdot\left(\frac{\nabla u}{\left\|\nabla u\right\|}\right)+\left\|\nabla u\right\|\lambda K*\left(K*u-u_0\right)=0,\qquad \text{in } \Omega,\ \text{for } t>0\,, \tag{4.12}$$

where $K\left(x\right)$ is a blurring operator (heat kernel). The well-posedness of this equation in the sense that there is a maximum principle that determines the solution is shown in reference [OS88]. This approach fixes the staircase problem of the original scheme and is used for removal of both blur and noise. Strong and Chan [SC96] introduce the weighted total variation functional for spatially adaptive image restoration

$$TV_\alpha = \int_\Omega \alpha(x) \|\nabla u\| dx . \tag{4.13}$$

The function $\alpha(x)$ is such that it is large (more diffusion) away from possible edges and smaller (less diffusion) near likely edges. Blomgren, Chan and Mulet [BCM97] propose a new approach considering regularizing functionals of the type

$$R(u) = \int_\Omega \Phi(\|\nabla u\|) dx , \tag{4.14}$$

for suitable real functions Φ. They consider the functional in (4.14) for $\Phi(z) = z^p$, and for $p \in [1,2]$

$$R(u) = \int_\Omega \|\nabla u\|^p dx . \tag{4.15}$$

For an exponent $p = 1$, one has the TV-norm and when $p = 2$, one would be using the L_2-norm. Song [So03], in his dissertation, pursues this approach further and renames it "Adaptive TV Model". The model considers

$$\min_u \frac{1}{p} \int_\Omega \|\nabla u\|^p dx, \qquad 1 < p < 2 , \tag{4.16}$$

subject to $\frac{1}{2}\int (u - u_0)^2 dx = \sigma^2$. The Euler-Lagrange equation for this model is

$$\partial_t u - \nabla \cdot \left(\|\nabla u\|^{p-2} \nabla u \right) + \lambda (u - u_0) = 0 . \tag{4.17}$$

The proof of the uniqueness of the solution is given in [Ka92].

Levine, Chen and Stanich [LCS04] and Chen, Levine and Rao [CLR05] propose a variant to that of Blomgren *et al*, where they define the exponent p based on the observed data u_0, their model is:

$$J(u) = \int_\Omega \phi(x, \nabla u) dx + \frac{\lambda}{2} \int_\Omega |u_0 - u|^2 dx \tag{4.18}$$

where

$$\phi(x,r)=\begin{cases} \frac{1}{p(x)}|r|^{p(x)} & \text{if } |r|<\mu \\ |r|-\frac{\mu p(x)-\mu^{p(x)}}{p(x)} & \text{if } |r|\geq \mu \end{cases} \tag{4.19}$$

Here, $\mu>0$ is fixed, and $p(x)$ is based on a smoothed version of the observed image u_0,

$$p(x)=\frac{1}{1+\kappa\left|\nabla G_\sigma * u_0(x)\right|^2} \tag{4.20}$$

where κ and σ are adjustable parameters, and $G_\sigma(x)$ is a Gaussian smoothing kernel similar to (2.2). The authors show existence and uniqueness of minimizers for this functional, and develop a numerical method for computing them based on gradient descent.

Chambolle [Ch04] also touches upon this subject where he combines two functionals $\int|\nabla u|$ and $\int|\nabla u|^2$ as

$$F(u)=\frac{1}{2\omega}\int_{|\nabla u|<\omega}|\nabla u|^2\,dx+\int_{|\nabla u|\geq\omega}\left(|\nabla u|-\frac{\omega}{2}\right)dx+\int_\Omega|u-u_0|^2\,dx \tag{4.21}$$

where ω is an adjustable parameter to be chosen. The Euler-Lagrange equation for this functional resembles that of the models discussed in this section. Schults, Bollt, Chartrand, Esedoglu and Vixie [SBCEV05] has recently revisited the subject and they suggest to minimize the following functional,

$$\min_u J(u)=\int_\Omega|\nabla u|^{p(|\nabla u|)}\,dx+\frac{\lambda}{2}\int_\Omega|u_0-u|^q\,dx, \qquad q=1 \text{ or } 2 \tag{4.22}$$

For two cases: case 1, $p(x)=P\left(\left|\nabla\left(G_\sigma * u_0\right)(x)\right|\right)$; and case 2, $p(x)=P\left(\left|\nabla\left(G_\sigma * u\right)(x)\right|\right)$.

They prove existence in both cases, and uniqueness in the case of $q=2$.

More recently, Osher, Burger, Goldfarb, Xu and Yin [OBGXY05], introduced an innovative iterative variant to the classic Rudin-Osher-Fatemi model aiming at improving its restoration capabilities and generalizations. The new model, instead of stopping after recovering

the image u, it is used to compute new iterates u_k using the Bregman distance [Br67]. They define the Bregman distance between u and v as $D(u,v)$ associated with the functional J, and design the following algorithm

$$u_k = \arg\min_{u \in BV(\Omega)} \left\{ D(u, u_{k-1}) + \lambda \|u_0 - u\|_2^2 \right\}. \tag{4.23}$$

They obtain a sequence $\{u_k\}$ which they show converges monotonically to u_0, the noisy image. However, as k increases, for λ sufficiently small, the values u_k also converge monotonically to $\tilde{u}$, the true noise-free image, until

$$\|u_{\tilde{k}} - u_0\|_2 < \tau \|\tilde{u} - u_0\|_2, \tag{4.24}$$

for any $\tau > 1$. According to the authors, the ideal situation is to take λ small and $\tilde{k}$ large so that $\lambda\tilde{k}$ converges to a critical time $\tilde{t}$ at which the estimate (4.24) is satisfied. It is also shown in [OBGXY05] that the new model can be used for restoring blurry and noisy images very efficiently.

Years ago, Chambolle and Lions [CL97] established that the optimization problem (4.1) constraint by (4.2), is equivalent to the following unconstrained optimization

$$\inf_{u \in L^2(\Omega)} \int_\Omega \|\nabla u\| dx + \lambda \int_\Omega \|u - u_0\|^2 dx, \tag{4.25}$$

where $\lambda \geq 0$ is a Lagrange multiplier. The first integral in the functional is the regularization term which disfavors oscillations and is responsible for the smoothing of the noise. The second integral in the functional is the fidelity term which encourages the solution u to stay close to the observed image u_0. The work of Meyer [Me02] motivated much interest into understanding the role of the fidelity term. Vese and Osher [VO03] and Osher, Sole and Vese [OSV03] propose to replace the original fidelity form by weaker forms. These variants allow better separation of

texture and outline (cartoon) making it easier to treat both parts separately. More recently, Chan and Esedoglu [CE04] study a simple variant to the original Rudin-Osher-Fatemi to accomplish contrast invariance. They replace the squared L_2 norm in the fidelity term (4.25) by the L_2 norm instead. The energy is then

$$\int_\Omega \|\nabla u\| dx + \lambda \int_\Omega \|u - u_0\| dx .$$

The authors argue that this model is more effective than the original Rudin-Osher-Fatemi model in the presence of certain types of noise, such as salt and pepper.

5. Finite Element Implementation

Digital images are given on discrete (regular) grids. This lends itself for discretizing the PDEs to obtain numerical schemes that can be solved on a computer. Because of their favorable stability and efficiency properties, semi-implicit schemes have been the methods of choice for the scale discretization [CLMC92] [KM95] [BM97] [BM01] [WRV98] [We99] [MR01] [KM02] [KM00] [PR99] [MSL98] [HMS99] [HMS03] [HMS02] [DPRS01]. As for the space discretization, the most popular choices are finite difference [CLMC92] [WRV98] [We99] and finite element methods [KM95] [BM01] [BM97] [DPRS01] [PR99] (in that order of preference). We choose the finite element method for our discretization in space since it will give us more flexibility for our adaptive formulations.

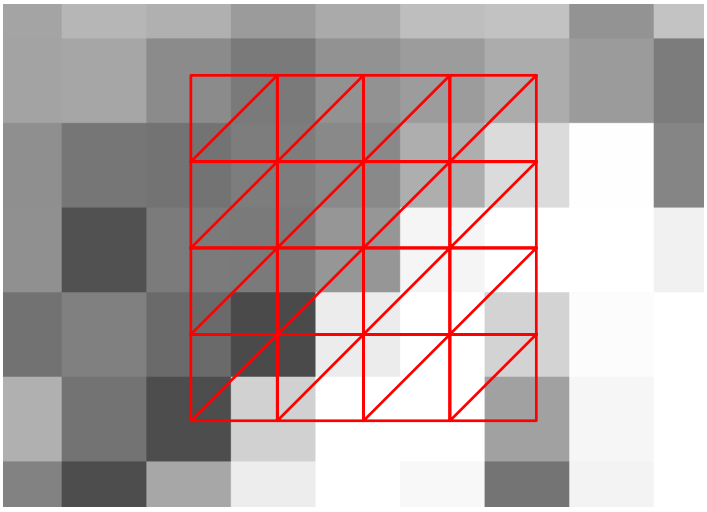


Figure 5.1: Triangulation for the finite element method. The nodes are located at the centers of the pixels.

Consider again the model due to Perona and Malik [PM90],

$$\partial_t u - \nabla \cdot \left(g\left(\|\nabla u\|^2 \right) \nabla u \right) = 0 . \tag{5.1}$$

The starting point for the finite element method is to partition the geometry (domain) into small units (elements or cells) of simple shape joined together at the vertices (nodes). This will

constitute our finite element space (mesh or grid). Once we have our mesh (see Figure 5.1), the idea is to approximate the dependent variables with functions that we can describe with a finite number of parameters (degrees of freedom DOF). Inserting this approximation into the weak form of equation (5.1) generates a system of equations for the degrees of freedom [Co05].

In our case, we need to perform discretizations in scale and space. We perform the semi-discretization in scale by letting $N \in \mathbb{N}$, and $k = T/N$ be fixed numbers [2], and letting $u(0,x) = u_0(x)$ in Ω. Then, we can look for a function u_n for every $n = 1,\ldots,N$, such that it is a solution to the equation

$$\frac{u_n - u_{n-1}}{k} - \nabla \cdot \left(g\left(\left\| \nabla u_{n-1} \right\|^2 \right) \nabla u_n \right) = 0 . \tag{5.2}$$

It is shown in [KM95] [BM01] that there exist unique variational solutions u_n of (5.2) at every discrete scale step, for which the following stability estimates hold:

$$\begin{aligned} &\left\| u_n \right\|_2 \le \left\| u_0 \right\|_2 , \qquad \left\| u_n \right\|_\infty \le \left\| u_0 \right\|_\infty , \qquad \text{for } n = 1,\ldots,N \quad \text{on } \Omega \\ &\sum_{n=1}^{N} \left\| \nabla u_n \right\|_2^2 h \le C, \qquad \sum_{n=1}^{N} \left\| u_n - u_{n-1} \right\|_2^2 \le C, \qquad \text{on } \Omega , \end{aligned} \tag{5.3}$$

where C is a general (large) constant [3]. To discretize the problem in space we can take advantage of the pixel structure of the image. For our case, the finite element method assumes that the approximation of the solution of the partial differential equations are continuous piecewise linear. *I.e.* the discrete intensity values are regarded as approximations of the continuous intensity function in the center of the pixels (see Figure 5.1). We can multiply equation (5.2) by an arbitrary test function $v \in V$, where V is the Sobolev space $W^{1,2}(\Omega)$ of

[2] Here, T represents the last scale state we want to reach.
[3] Here, h represents a typical element size.

$L_2(\Omega)$–functions with doubly integrable weak derivatives, and integrate (using Green's theorem and homogeneous Neumann boundary conditions) to obtain the weak form [Mi02]

$$\int_\Omega u_n v\,dx + k\int_\Omega g\left(\left\|\nabla u_{n-1}\right\|^2\right)\nabla u_n \nabla v\,dx = \int_\Omega u_{n-1} v\,dx\,. \tag{5.4}$$

Then, for each scale step n, we look for a continuous piecewise linear function $u_{n,h} \in V_h$ that satisfies

$$\int_\Omega u_{n,h} v_h\,dx + k\int_\Omega g\left(\left\|\nabla u_{n-1,h}\right\|^2\right)\nabla u_{n,h} \nabla v_h\,dx = \int_\Omega u_{n-1,h} v_h\,dx \tag{5.5}$$

for all $v_h \in V_h$. Considering the standard Lagrangian base functions $\phi_q \in V_h$, $q = 1,\ldots,M$, given by $\phi_q(x_p) = \delta_{qp}$ (Kronecker delta) for all nodes, the function $u_{n,h}$ is given by

$$u_{n,h} = \sum_{p=1}^{M} u_{n,p}\phi_p\,. \tag{5.6}$$

Substituting (5.6) in (5.5) and considering as test functions $v_h = \phi_q$ for $q = 1,\ldots,M$, we get the Ritz-Galerkin equation for the nodal values $u_{n,p}$, of the piecewise linear function $u_{n,h}$:

$$\sum_{p=1}^{M}\left(\int_\Omega \phi_p\phi_q\,dx + k\int_\Omega g\left(\left\|\nabla u_{n-1,h}\right\|^2\right)\nabla\phi_p\nabla\phi_q\,dx\right)u_{n,p} = \int_\Omega u_{n-1,h}\phi_q\,dx, \qquad q = 1,\ldots,M\,. \tag{5.7}$$

Then, in each scale step we need to assemble and solve a linear system of the form

$$\left[\mathbf{M} + k\mathbf{A}\left(g\left(\left\|\nabla u_{n-1}\right\|^2\right)\right)\right]\mathbf{u}_n = \mathbf{f}_{n-1}, \tag{5.8}$$

for the vector of unknowns (DOF) $\mathbf{u}_n$.

6. Dynamic/Adaptive Implementation

We implement a variation of Blomgren *et al*'s [BCM97] version of the fully nonlinear Euler-Lagrange equation (4.12),

$$\partial_t u - \|\nabla u\| \nabla \cdot \left(L\left(\|\nabla u\|^{p-2} \right) \nabla u \right) + \Lambda \left(u - u_0 \right) = 0 , \tag{6.1}$$

defined in the domain Ω with boundary conditions $\partial_n u = 0$ on $\partial\Omega$ (where n is the unit normal vector to the boundary of the domain Ω). The Neumann boundary conditions should guarantee that the filtering does not significantly affect the average gray value of the image. The initial condition is the observed image $u(0,x) = u_0(x)$ in Ω.

The model (6.1) can be regarded as "adaptive TV model with morphological convection and anisotropic diffusion." Unlike the approach in [ROF92], we implement a user-independent choice of all the parameters in the model. We start by estimating the unknown parameter σ, namely the standard deviation of the noise. Since we consider that the image has been perturbed by additive Gaussian noise, $u_0 = u + \eta$, then the variance of the noisy image has to be equal to the sum of the variance of the true image and the variance of the noise, $\sigma_{u_0}^2 = \sigma_{G_\sigma * u_0}^2 + \sigma_\eta^2$. Here, the variance of the (unknown) true image is approximated by the variance of the convolved noisy image. This parameter will be updated iteratively as we will see below.

For the parameter λ, we implement a variation of the method suggested in [ROF92]. Instead of integrating (or summing) over the domain Ω, we implement a node-wise $\Lambda \equiv \|\nabla u\| \lambda$ as (again with a little abuse of notation)

$$\Lambda = -\frac{\sqrt{u_x^2 + u_y^2}}{2\sigma^2} \left[\sqrt{u_x^2 + u_y^2} - \left(\frac{(u_0)_x u_x}{\sqrt{u_x^2 + u_y^2}} + \frac{(u_0)_y u_y}{\sqrt{u_x^2 + u_y^2}} \right) \right],$$

$$\Lambda = -\frac{1}{2\sigma^2}\left[\left(\sqrt{u_x^2+u_y^2}\right)^2 - \left(\frac{(u_0)_x\, u_x\sqrt{u_x^2+u_y^2}}{\sqrt{u_x^2+u_y^2}} + \frac{(u_0)_y\, u_y\sqrt{u_x^2+u_y^2}}{\sqrt{u_x^2+u_y^2}}\right)\right],$$

$$\Lambda = -\frac{1}{2\sigma^2}\left[u_x^2+u_y^2-(u_0)_x\, u_x-(u_0)_y\, u_y\right],$$

$$\Lambda = -\frac{1}{2\sigma^2}\left[u_x\left(u_x-(u_0)_x\right)+u_y\left(u_y-(u_0)_y\right)\right]. \tag{6.2}$$

The dynamic parameter Λ has the following attributes:

1. The smaller the value of Λ, the more the diffusion contributed by the forcing term. Analogously, the larger the value of Λ, the lesser the diffusion contributed by the forcing term.
2. At the beginning of the scale-marching iterations the gradients $u_x \approx (u_0)_x$ and $u_y \approx (u_0)_y$, therefore the terms $u_x-(u_0)_x$ and $u_y-(u_0)_y$ are very small and the forcing term tends to contribute more to the diffusion process. In areas where u_x and u_y are large (*i.e.* near edges), these values compensate for the small terms $u_x-(u_0)_x$ and $u_y-(u_0)_y$.
3. As iterations evolve the terms $u_x-(u_0)_x$ and $u_y-(u_0)_y$ get larger. Near edges, the forcing term prevents diffusion and helps reach convergence.

We can also get an *a posteriori* estimate to the variance of the noise σ^2 by integrating (or summing) over the domain after convergence,

$$\sigma^2 = -\frac{1}{2}\int_\Omega \frac{1}{\Lambda}\left[u_x\left(u_x-(u_0)_x\right)+u_y\left(u_y-(u_0)_y\right)\right]dx. \tag{6.3}$$

This will be an improved value that can be used to run the model with a better estimate of the unknown parameter σ.

The diffusion tensor $L\left(\left\|\nabla u\right\|^{p-2}\right)$ incorporates the parameter $1 \le p \le 2$, as suggested in [BCM97]. The diffusion tensor becomes

$$L\left(\left\|\nabla u\right\|^{p-2}\right) = \begin{bmatrix} \left\|\nabla u\right\|^{p_x-2} & -\left\|\nabla u\right\|^{p_{xy}-2} \\ -\left\|\nabla u\right\|^{p_{xy}-2} & \left\|\nabla u\right\|^{p_y-2} \end{bmatrix} \tag{6.4}$$

where p_x, p_y, p_{xy}, are the following unnormalized Gaussians:

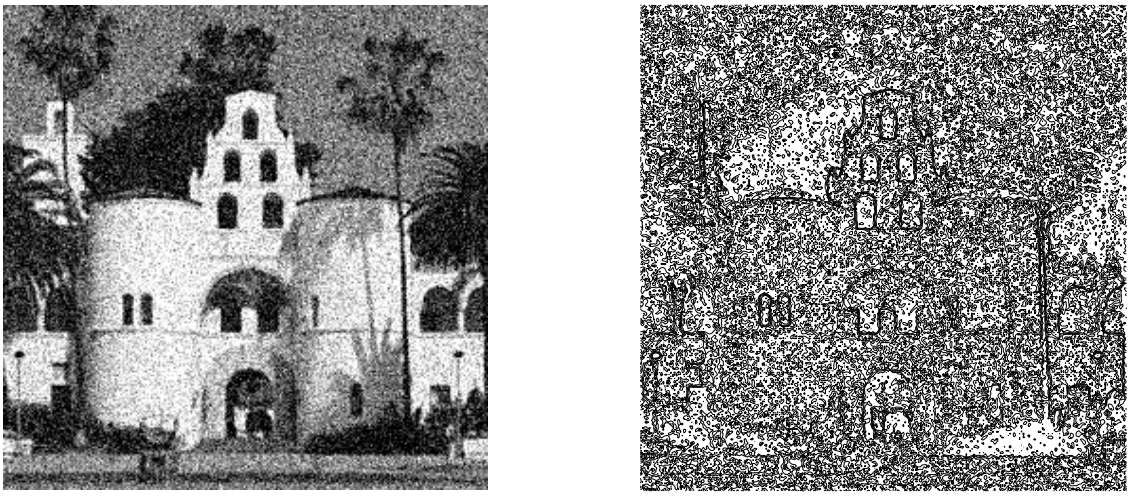
$$\begin{aligned} p_x &= 1 + e^{-\hat{u}_x^2/4\sigma^2} \\ p_y &= 1 + e^{-\hat{u}_y^2/4\sigma^2} \\ p_{xy} &= 1 + e^{-\left(\hat{u}_x^2+\hat{u}_y^2\right)/4\sigma^2}. \end{aligned} \tag{6.5}$$

In equation (6.5) above, $\hat{u}_x$ and $\hat{u}_y$ are the gradients of the convolved noisy image $G_\sigma * u_0$ used to estimate the unknown parameter σ. The dynamic parameters p_x, p_y, p_{xy}, have the following attributes:

1. For every pixel in the image, the parameters take values $1 \le p_x \le 2$, $1 \le p_y \le 2$ and $1 \le p_{xy} \le 2$.
2. When $p_x = 1$, $p_y = 1$ or $p_{xy} = 1$ the model uses the TV-norm in the corresponding direction, and when $p_x = 2$, $p_y = 2$ or $p_{xy} = 2$, the model uses the L_2-norm in the corresponding direction.
3. When the parameters $1 \le p_x \le 2$, $1 \le p_y \le 2$ and $1 \le p_{xy} \le 2$, the model interpolates between both norms.

Figures 6.1, 6.2 and 6.3 show examples of the implementation of the proposed model. We can also observe the dynamic parameters Λ, p_x, p_y and p_{xy}.

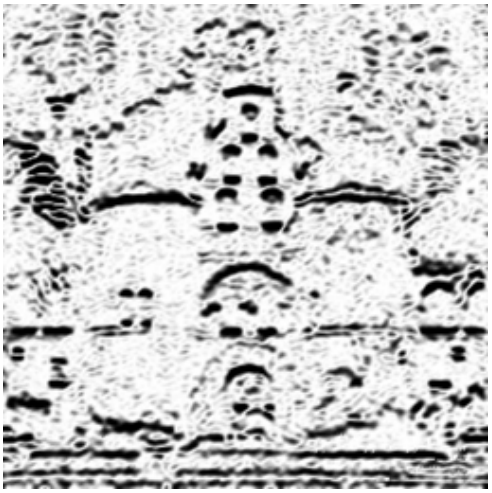
(a) Grayscale true image and its contours.



(b) Noisy image with $SNR = 23.4$ dB and its contours.

(c) Treated image and its contours.

(d) Parameter p_x.

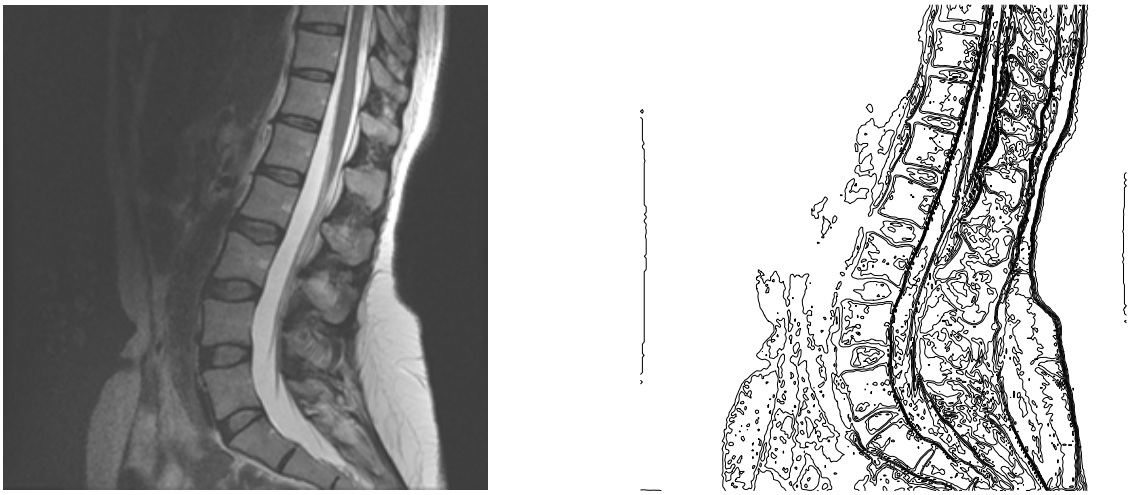


(e) Parameter p_y.

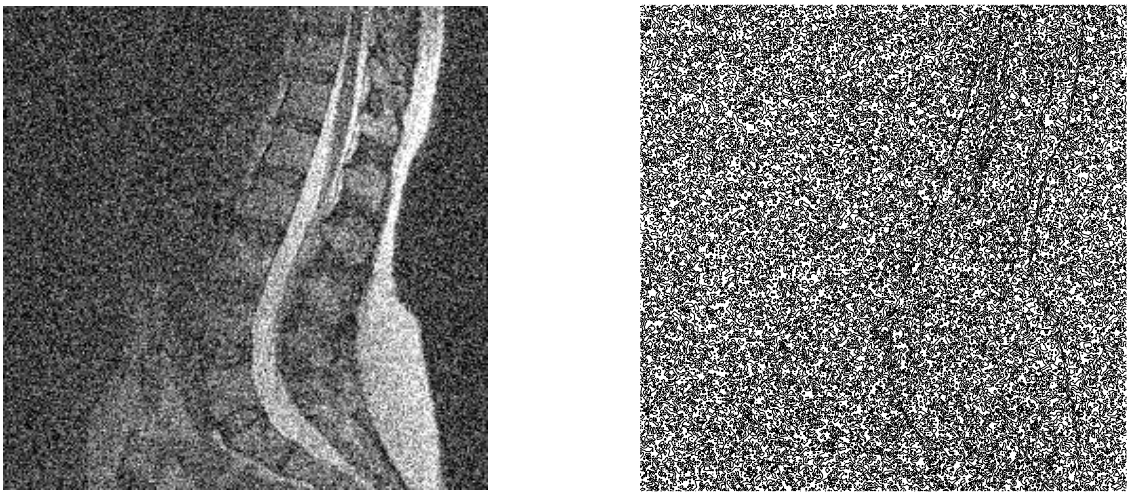
(f) Parameter p_{xy}.

(g) Parameter Λ.

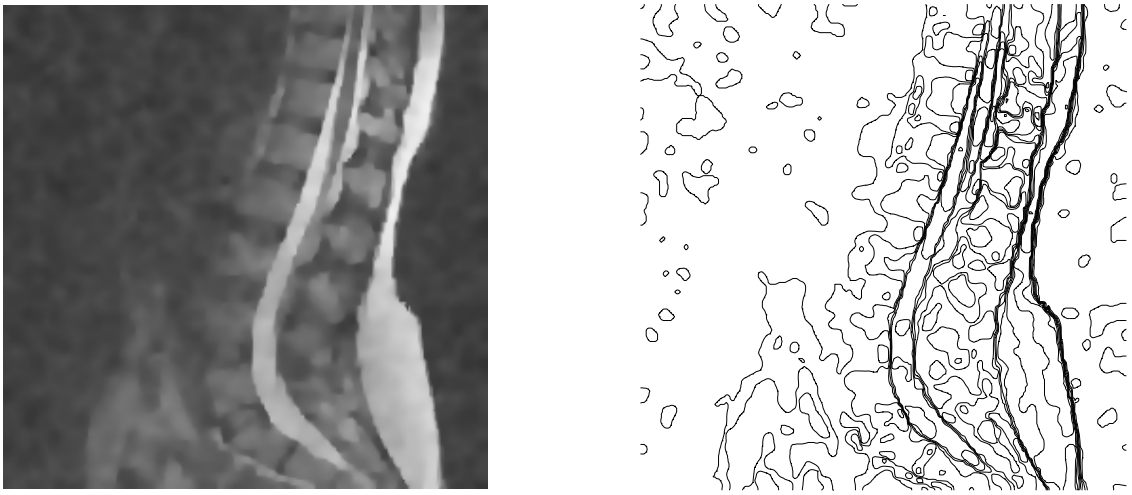
Figure 6.1: Dynamic restoration of a grayscale image. **(a)** Grayscale true image and its contours. **(b)** Noisy image with SNR = 23.4 dB and its contours. **(c)** Treated image and its contours. **(d)** Parameter p_x. **(e)** Parameter p_y. **(f)** Parameter p_{xy}. **(g)** Parameter Λ.



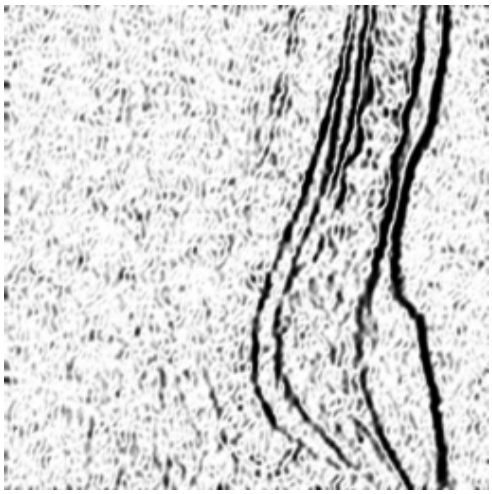
(a) Grayscale true medical image and its contours.



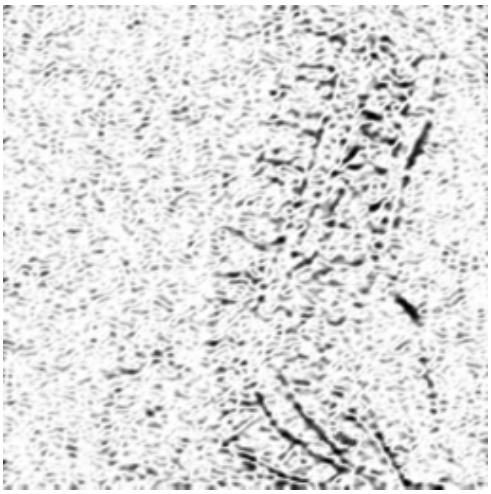
(b) Noisy image with $SNR = 20.1$ dB and its contours.



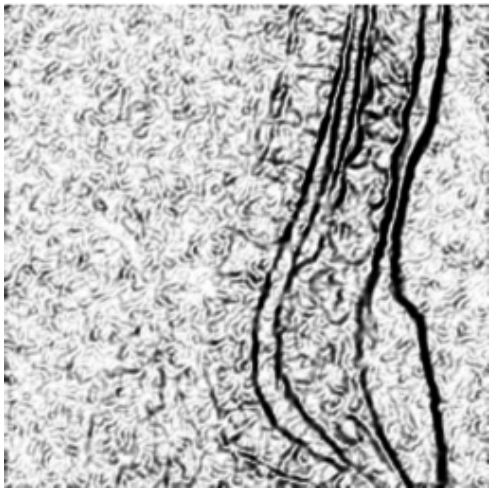
(c) Treated image and its contours.



(d) Parameter p_x .



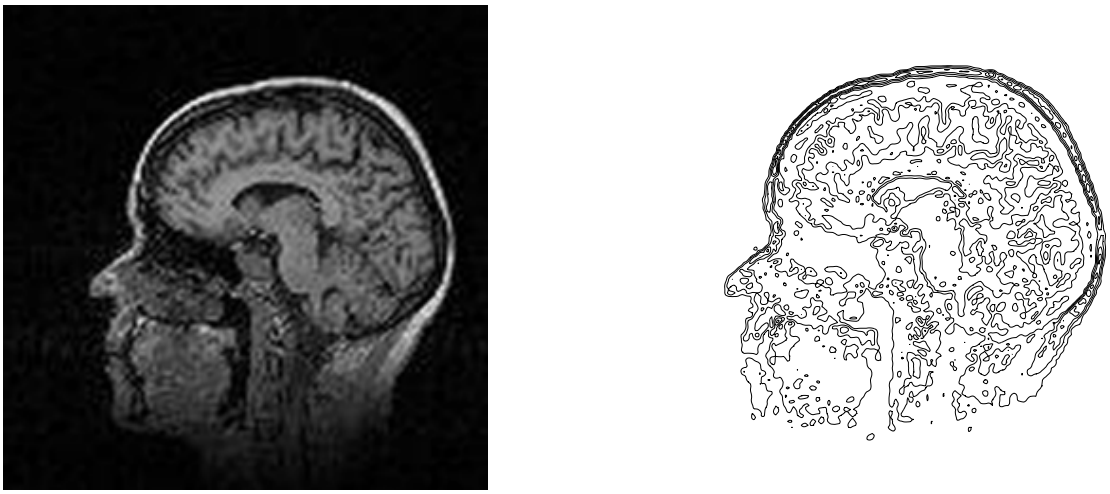
(e) Parameter p_y .



(f) Parameter p_{xy} .

(g) Parameter Λ

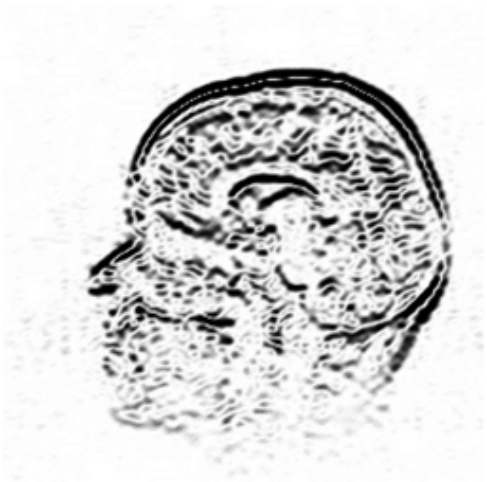
Figure 6.2: Dynamic restoration of a medical image. **(a)** Grayscale true medical image and its contours. **(b)** Noisy image with $SNR = 20.1$ dB and its contours. **(c)** Treated image and its contours. **(d)** Parameter p_x . **(e)** Parameter p_y . **(f)** Parameter p_{xy} . **(g)** Parameter Λ . *Original image courtesy of Robert E. Edelman*



(a) Noisy medical image with unknown SNR and its contours.

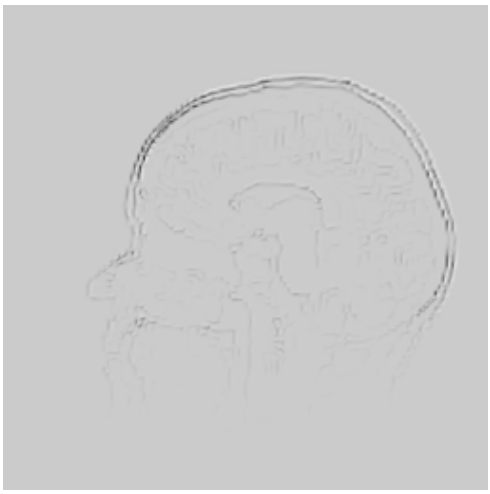
(c) Treated image and its contours.

(d) Parameter p_x.



(e) Parameter p_y.

(f) Parameter p_{xy}.



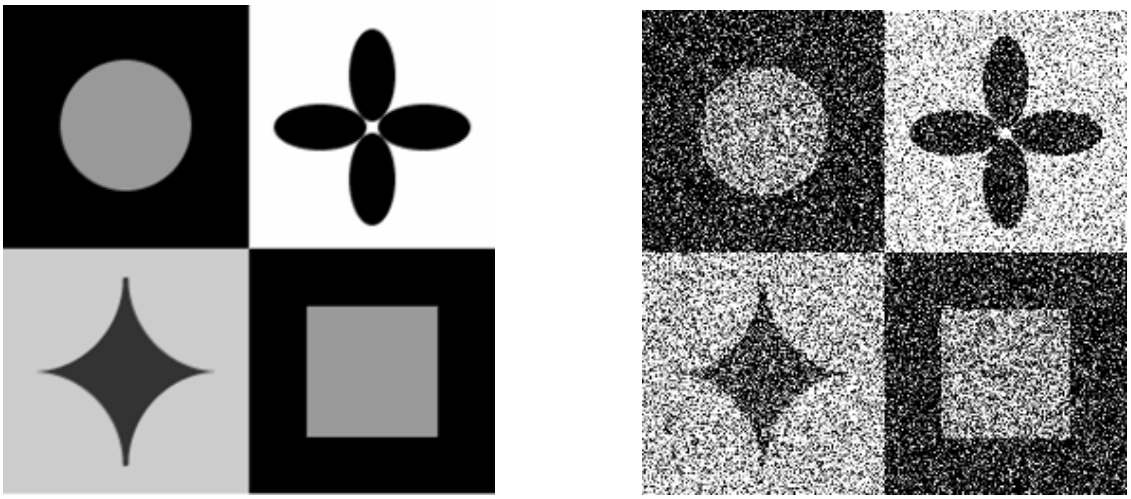
(g) Parameter Λ.

Figure 6.3: Dynamic restoration of a medical image. **(a)** Noisy medical image with unknown SNR and its contours. **(b)** Treated image and its contours. **(c)** Parameter p_x. **(d)** Parameter p_y. **(e)** Parameter p_{xy}. **(f)** Parameter Λ.

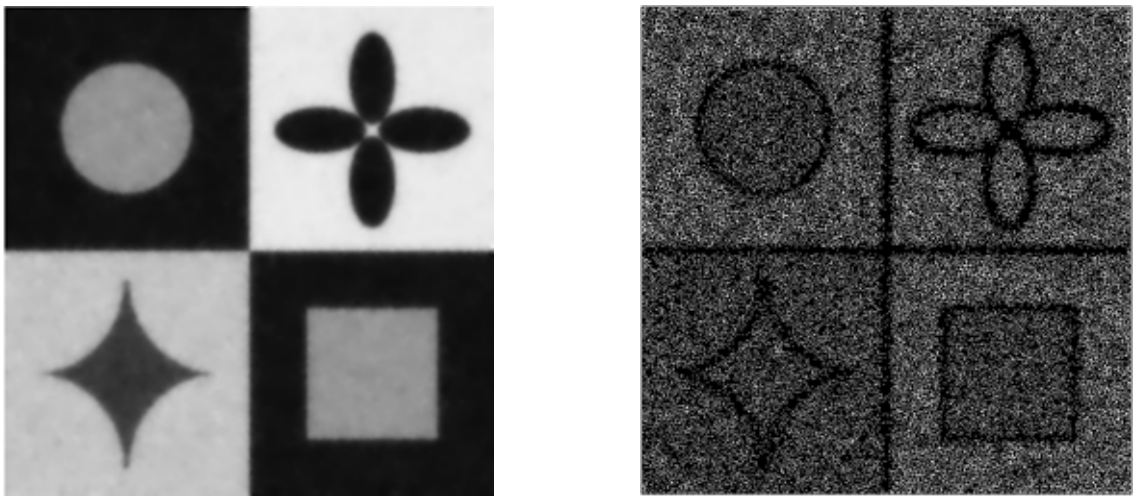
In an earlier stage of our research we implemented a scheme that aims at using information from the images to adjust the numerical solution via adaptive grids [BB05]. The employment of adaptive grid proved to be a very efficient approach where considerably fewer DOF are necessary to produce similar results to the regular grid case. By using the remeshing approach based on the L_2-norm, nodes are placed following the edges of the images which allows very good edge preservation. Some examples of this proof of concept using the Catté-Lions-Morel-Coll model are depicted in Figures 6.4 and 6.5. The computational effort necessary to run the experiments shown in Figure 6.4 on a computer equipped with a 1.50GHz Intel® Pentium® processor and 1.25 GB of RAM is shown in Table 6.1.

Table 6.1

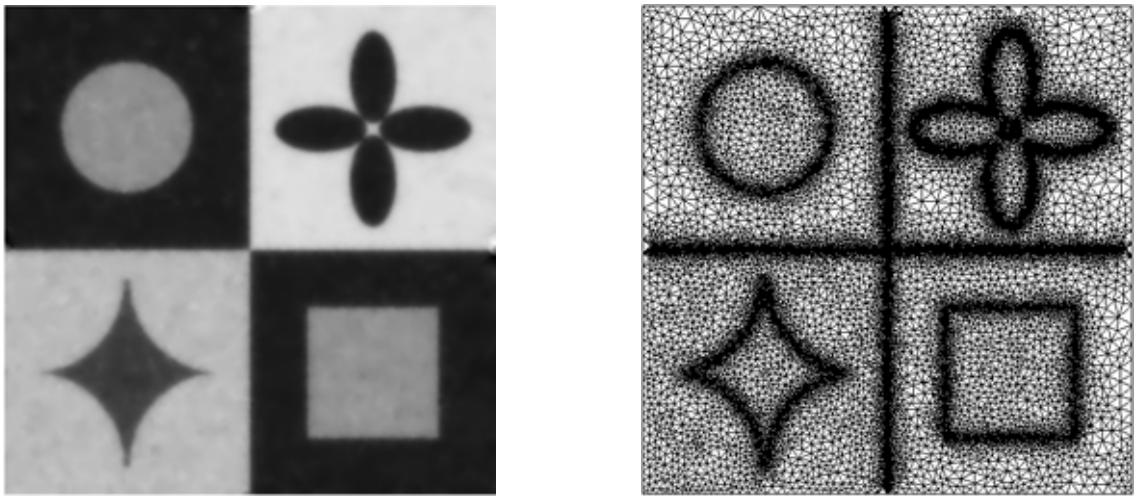
Experiment Case	CPU time
Regular grid, 47 iteration	285.230 s
Adaptive grid, 1 iteration	116.668 s
Adaptive grid, 2 iterations	153.230 s
Adaptive grid, 5 iterations	274.124 s



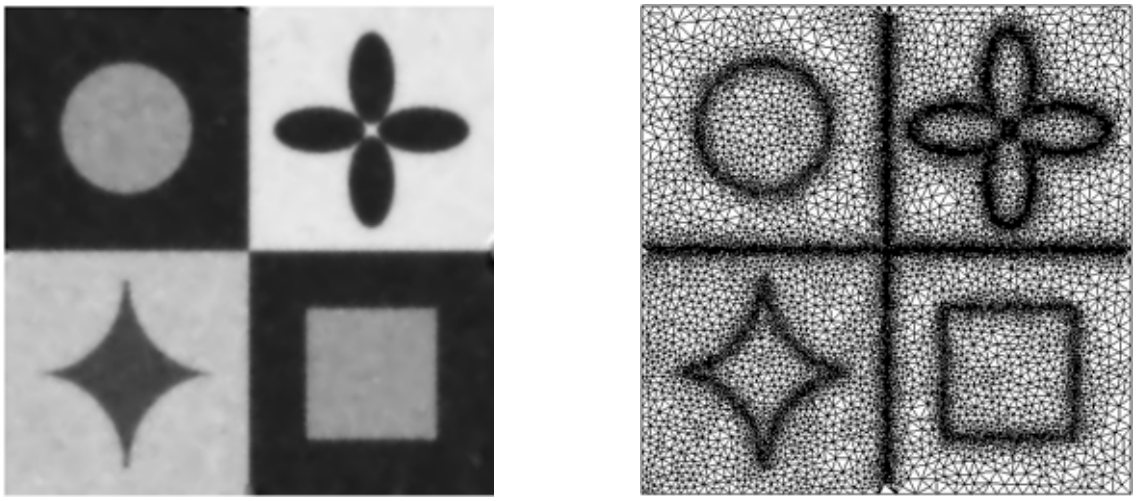
(a) Synthetic grayscale true image and noisy image with $SNR = 14.8$ dB (65,536 DOF).



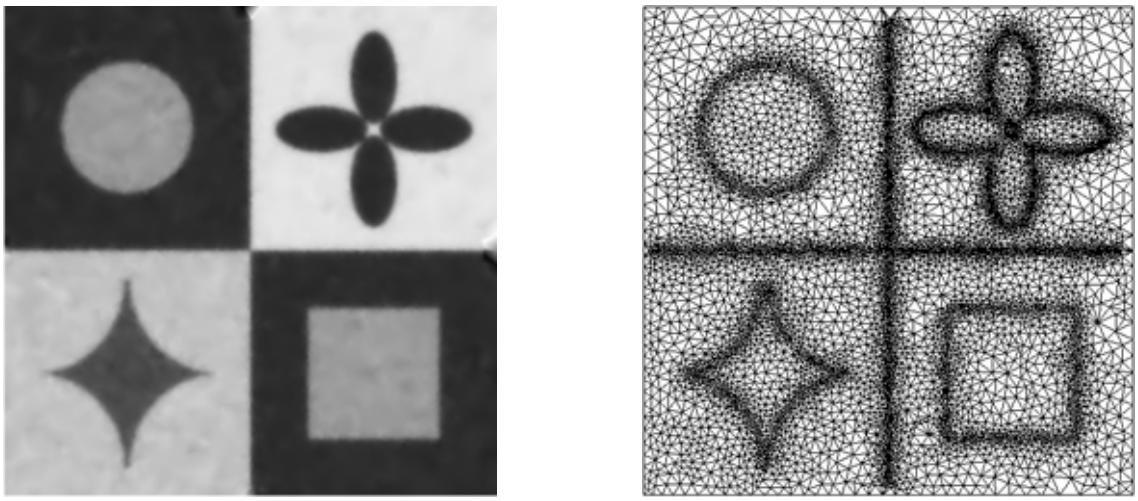
(b) Treated image and mesh after iteration 1 (45,696 DOF).



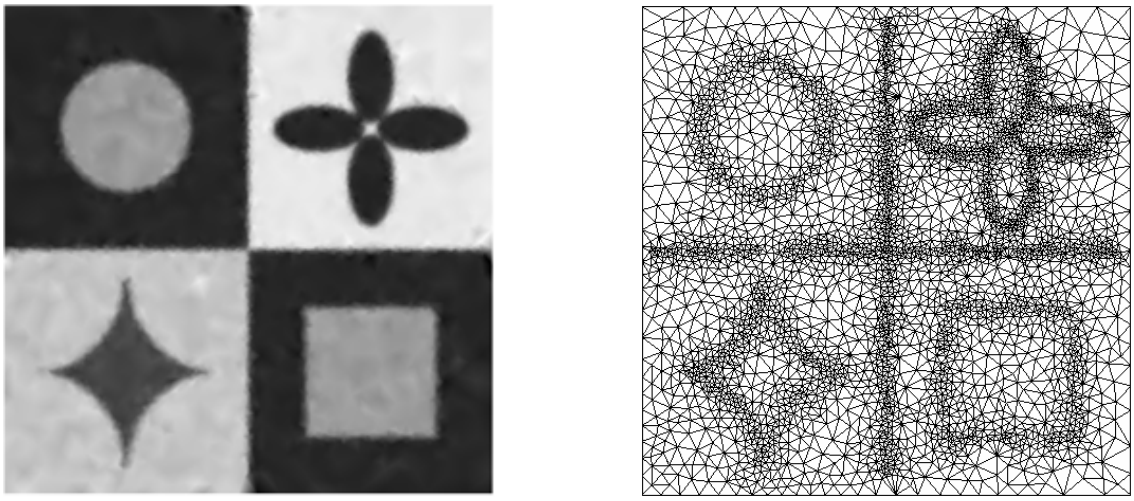
(c) Treated image and mesh after iteration 2 (25,401 DOF).



(d) Treated image and mesh after iteration 3 (20,159 DOF).

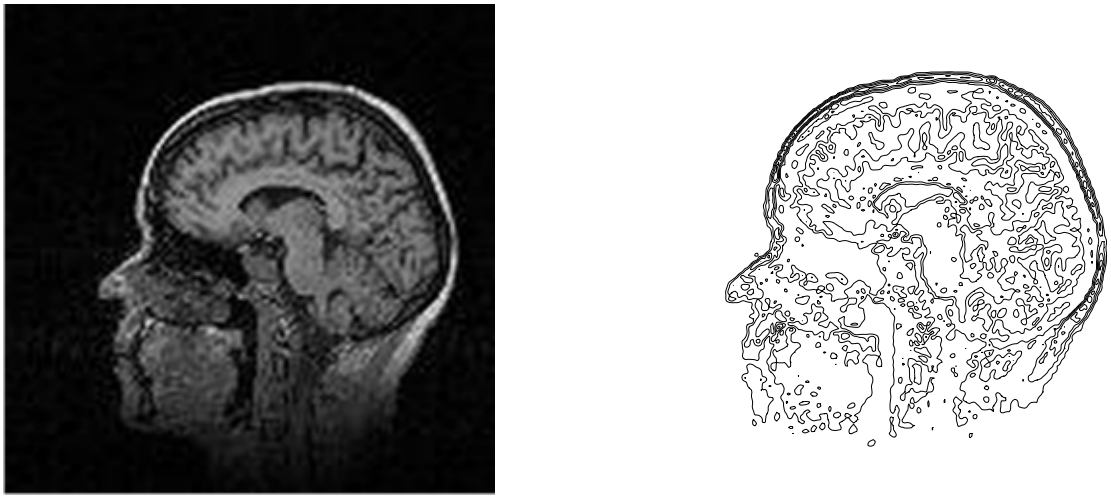


(e) Treated image and mesh after iteration 4 (12,780 DOF).

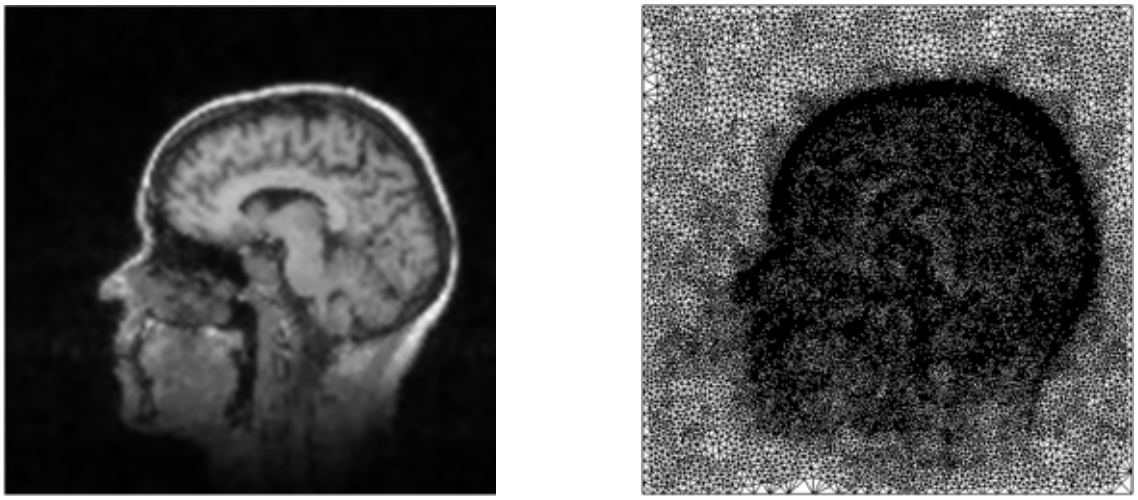


(f) Treated image and mesh after iteration 5 (6,417 DOF).

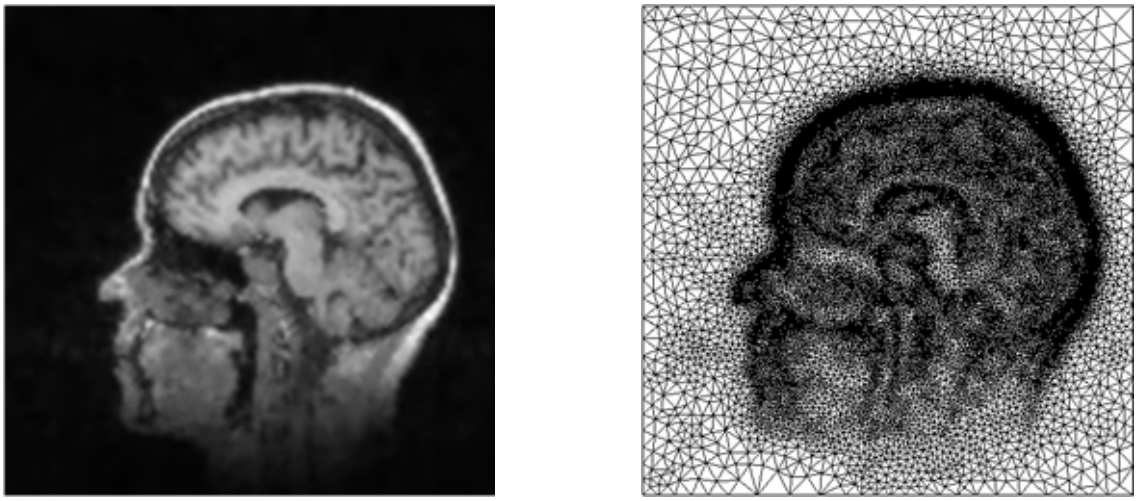
Figure 6.4: Adaptive mesh implementation of the Catté- Lions-Morel-Coll model. **(a)** Synthetic grayscale true image and noisy image with $SNR = 14.8$ dB (65,536 DOF). **(b)** Treated image and mesh after iteration 1 (45,696 DOF). **(c)** Treated image and mesh after iteration 2 (25,401 DOF). **(d)** Treated image and mesh after iteration 3 (20,159 DOF). **(e)** Treated image and mesh after iteration 4 (12,780 DOF). **(f)** Treated image and mesh after iteration 5 (6,417 DOF).



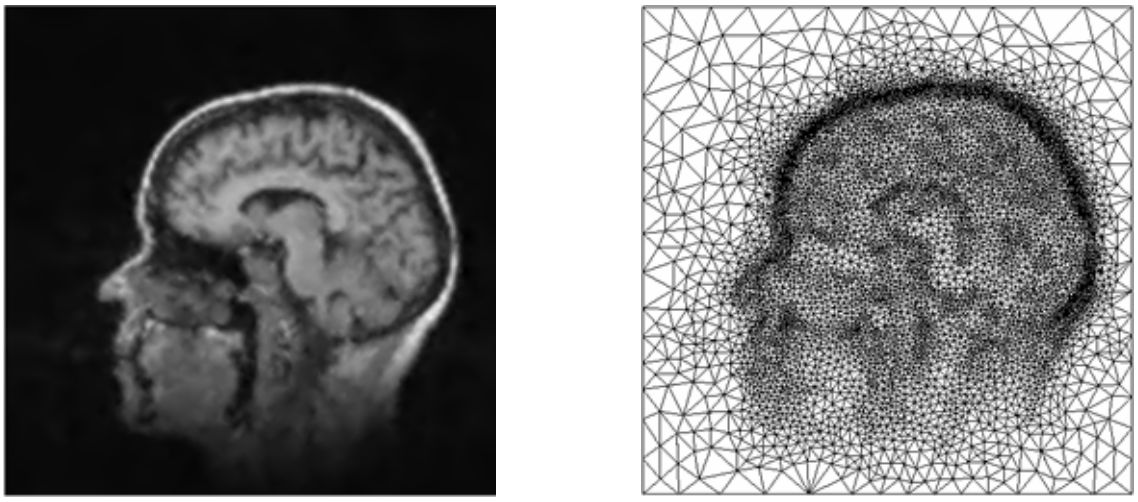
(a) Noisy grayscale medical image and it contours.



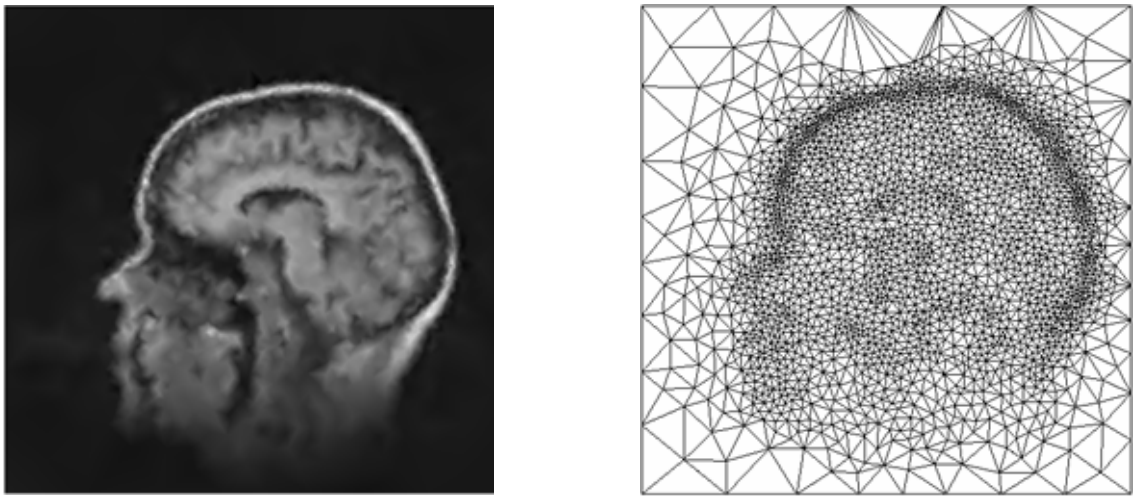
(b) Treated image and mesh after iteration 1.



(c) Treated image and mesh after iteration 2.



(d) Treated image and mesh after iteration 3.



(e) Treated image and mesh after iteration 4.

Figure 6.5: Adaptive mesh implementation of the Catté- Lions-Morel-Coll model. **(a)** Noisy grayscale medical image and its contours. **(b)** Treated image and mesh after iteration 1. **(c)** Treated image and mesh after iteration 2. **(d)** Treated image and mesh after iteration 3. **(e)** Treated image and mesh after iteration 4.

7. Summary

In this report we presented some of the most widely used PDE based image processing models. We briefly exposed their advantages and drawbacks and tried to describe their methodological (and chronological) evolutions. The models studied in this report are very well established in the image processing community and will serve as robust infrastructure whereupon we can built our own research with confident.

As a proof of concept, we started to unveil the direction of our future research, namely, dynamic and parameter-free total variation based image processing. It is our intent to further study efficient mechanisms for implementing total variation based image processing that will spare the user from having to adjust parameters for every occasion. Our dynamic/adaptive schemes aim at incorporating more information from the images themselves in the solution process and limit the use of *ad hoc* techniques.

References

[AV94] R. Acar, C.R. Vogel. Analysis of Total variation Penalty Methods for Ill-Posed Problems, *Inverse Problems*, **10**:1217–1229, 1994.

[BB05] C. Bazan, P. Blomgren. Adaptive Finite Element Method for Image Processing, *Proceedings COMSOL Multiphysics Conference*, 377–381, 2005.

[BBD93] G.I. Barenblatt, M, Bertsch, R. Dal Passo, M. Ughi. A Degenerate Pseudo-Parabolic Regularization of a Nonlinear Forward-Backward Heat Equation Arising in the Theory of Heat and Mass Exchange in Stably Stratified Turbulent Shear Flow, *SIAM Journal on Mathematical Analysis*, **24**(6):1414−1439, 1993.

[BCM97] P. Blomgren, T. Chan, P. Mulet. Extensions to Total Variation Denoising, *Proceedings of Society of Photo-Optical Instrumentation Engineers*, **3162**, 1997.

[Be94] B. Benhamouda. Parameter Adaptation for Nonlinear Diffusion in Image Processing, *Master Thesis, Department of Mathematics*, University of Kaiserslautern, Kaiserslautern, 1994.

[BJS64] L. Bers, F. John, M. Schechter. Partial Differential Equations, *Interscience*, New York, 1964.

[BM01] E. Bänsch, K. Mikula. Adaptivity in 3D Image Processing, *Computing and Visualization in Science*, **4**(1):21–30, 2001.

[BM97] E. Bänsch, K. Mikula. A Coarsening Finite Element Strategy in Image Selective Smoothing, *Computing and Visualization in Science*, **1**(1):53–61, 1997.

[Br67] L. Bregman. The Relaxation Method of Finding the Common Points of Convex Sets and Its Application to the Solution of Problems in Convex Programming,

Russian Academy of Science Journal of Computational Mathematics and Mathematical Physics, **7**:200-217, 1967.

[CBAB94] P. Charbonnier, L. Blanc-Féraud, G. Aubert, M. Barlaud. Two Deterministic Half-Quadratic Regularization Algorithms for Computed Imaging, *Proceedings of the IEEE International Conference on Image Processing*, **2**:168−172, 1994.

[CE04] T. Chan, S. Esedoglu. Aspects of Total Variation Regularized L^1 Function Approximation, *Techinal Report CAM 05-01*, *University of California Los Angeles*, Los Angeles, 2004.

[CG93] G.-H. Cottet, L. Germain. Image Processing Through Reaction Combined with Nonlinear Diffusion, *Mathematics of Computation*, **61**(204):659−673, 1993.

[Ch04] A. Chambolle. An Algorithm for Total Variation Minimization and Applications, *Journal of Mathematical Imaging and Vision*, **20**(1-2):89−97, 2004.

[CL97] A. Chambolle, P. Lions. Image Recovery via Total Variation Minimization and Related Problems, Numerical Mathematics, **76**:167−188, 1997.

[CLMC92] F. Catté P.-L. Lions, J.-M. Morel, T. Coll. Image Selective Smoothing and Edge Detection by Nonlinear Diffusion, *SIAM Journal of Numerical Analysis*, **29**(1):182−193, 1992.

[CLR05] Y. Chen, S. Levine, M. Rao. Variable Exponent, Linear Growth Functions in Image Restoration, submitted to *SIAM Journal on Applied Mathematics*, 2005.

[Co05] COMSOL AB. COMSOL Multiphysics Modeling Guide v3.2, *COMSOL AB*, Burlington, 2005.

[De74] P. Deuflhard. A Modified Newton Method for the Solution of Ill-Conditioned Systems of Nonlinear Equations with Application to Multiple Shooting,

Numerische Mathematik, **22**:289−315, 1974.

[DPRS01] U. Diewald, T. Preusser, M. Rumpf, R. Strzodka. Diffusion Models and Their Accelerated Solution in Image and Surface Processing, *Acta Mathematica Universitatis Comenianae*, **70**(1):15−34, 2001.

[FPC91] D.S. Fritsch, S.M. Pizer, J.M. Coggins. A Multiscale Medial Description of Greyscale Image Structure, *Advances in Intelligent Robotic Systems, Society of Photo-Optical Instrumentation Engineers*, 1991.

[Fr92] D.S. Fritsch. A Medical Description of Grayscale Image Structure by Gradient-Limited Diffusion, R.A. Robb (Ed.), *Visualization in Biomedical Computing*, *Society of Photo-Optical Instrumentation Engineers*, **1808**:105−117, 1992.

[FW94] J. Fröhlich, J. Weickert. Image Processing Using a Wavelet Algorithm for Nonlinear Diffusion, *Report No. 104, Laboratory of Technomathematics*, University of Kaiserslautern, Kaiserslautern, 1994.

[GCM99] G. Golub, T. Chan, P. Mulet. A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration, *SIAM Journal on Scientific Computing*, **20**(6):1964−1977, 1999.

[GY91] D. Geiger, A. Yuille. A Common Framework for Image Segmentation, *International Journal of Computer Vision*, **6**(3):227−243, 1991.

[GW02] R.C. Gonzalez, R.E. Woods. Digital Image Processing, Second Edition, *Pearson Education*, Singapore, 2002.

[He64] G. Hellwig, Partial Differential Equations, *Blaisdell*, New York, 1964.

[HMS03] A. Handlovičová, K. Mikula, F. Sgallari. Semi-Implicit Complementary Volume Scheme for Solving Level set Like Equations in Image Processing and Curve

Evolution, *Numerische Mathematik*, **93**(4):675−69, 2003.

[HMS02] A. Handlovičová, K. Mikula, F. Sgallari. Variational Numerical Methods for Solving Nonlinear Diffusion Equations Arising in Image Processing, *Journal for Visual Communication and Image Representation*, **13**(1-2):217−237, 2002.

[HMS99] A. Handlovičová, K. Mikula, A. Sarti. Numerical Solution of Parabolic Equations Related to Level Set Formulation of Mean Curvature Flow, *Computing and visualization in Science*, **1**(2):179−182, 1999.

[HN83] K. Höllig, J.A. Nohel. A Diffusion Equation with a Non-Monotone Constitutive Function, J.M. Ball (Ed.), *Proceedings of NATO/London Mathematical Society Conference on Systems of Partial Differential Equations*, 409−422, 1983.

[Ho83] K. Höllig. Existence of Infinitely Many Solutions for a Forward-Backward Heat Equation, *Transactions of the American Mathematical Society*, **278**:299−319, 1983.

[Hu86] R. Hummel. Representations Based on Zero-Crossings in Scale-Space, *IEEE Computer Vision and Pattern Recognition*, 204−209, 1986.

[Ka92] B. Kawohl. Remarks on the operator $\operatorname{div}\left(\nabla u/|\nabla u|\right)$. *Contemporary Mathematics*, **127**:69–83, 1992.

[Ki97] S. Kichenassamy. The Perona-Malik Paradox, *SIAM Journal of Applied Mathematics*, **57**(5):1328−1342, 1997.

[KM02] Z. Krivá, K. Mikula. An Adaptive Finite Volume Scheme for Solving Nonlinear Diffusion in Image Processing, *Journal for Visual Communication and Image Representation*, **13**(1/2) (2002):22−35, 2002.

[KM00] Z. Krivá, K. Mikula. An Adaptive Finite Volume Scheme in Processing of Color

Images, *Proceedings of ALGORITMY 2000, Conference on Scientific Computing*, Podbanské, Slovakia, 174−188, 2000.

[KM95] J. Kačur, K. Mikula. Solution of Nonlinear Diffusion Appearing in Image Smoothing and Edge Detection, *Applied Numerical Mathematics*, **17**(1):47–59, 1995.

[Ko84] J.J. Koenderink. The Structure of Images, *Biological Cybernetics*, **50**(5):363−370, 1984.

[LC94] X. Li, T. Chen. Nonlinear Diffusion with Multiple Edginess Thresholds, *Pattern Recognition*, **27**(8):1029−1037, 1994.

[LCS04] S. Levine, Y. Chen, J. Stanich. Image Restoration via Nonstandard Diffusion, *Technical Report #04-01*, *Department of Mathematics and Computer Science, Duquesne University*, Pittsburgh, 2004.

[LOR93] P.-L. Lions, S. Osher, L.I. Rudin. Denoising and Deblurring Images Using Constrained Nonlinear Partial Differential Equations, *Technical Report, Cognitech, Inc*., 1993.

[Me02] Y. Meyer. Oscillating Patterns in Image Processing and Nonlinear Evolution Equations, *AMS University Lecture Series*, **22**, 2002.

[Mi02] K. Mikula. Image Processing with Partial Differential Equations, *Modern Methods in Scientific Computing and Applications* (A.Bourlioux, M.Gander, Eds.), *NATO Science Series II*, **75**:283−322, 2002.

[MO00] A. Marquina, S. Osher. Explicit Algorithms for a New Time Dependent Model Based on Level Set Motion for a Nonlinear Debluring and Noise Removal, *SIAM Journal on Scientific Computing*, **22**:387−405, 2000.

[MR01] K. Mikula, N. Ramarosy. Semi-Implicit Finite Volume Schemes for Solving Nonlinear Diffusion Equations in Image Processing, *Numerische Mathematik*, **89**(3):561−590, 2001.

[MSL98] K. Mikula, A. Sarti, C. Lamberti. Geometrical Diffusion in 3-D-Echocardiography, *Proceedings of ALGORITMY'97, Conference on Scientific Computing*, (J. Kačur, K. Mikula, eds.), *Acta Mathematica Universitatis Comenianae*, **67**:167−181, 1998.

[No90] N. Nordström. Biased Anisotropic Diffusion: a Unified Regularization and Diffusion Approach to Edge Detection, *Image and Vision Computing*, **8**(11):318−327, 1990.

[NS92] M. Nitzberg, T. Shiota. Nonlinear Image Filtering with Edge and Corner Enhancement, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **14**:826−833, 1992.

[OBGXY05] S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Yin. An Iterative regularization Method for Total Variation-Based Image Restoration, *SIAM Multiscale Modeling and Simulation*, **4**(2):460−489, 2005.

[OS88] S. Osher and J.A. Sethian. Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulation, *Journal of Computational Physics,* **79**:12−49, 1988.

[OSV03] S. Osher, A. Sole, L. Vese. Image Decomposition and Restoration Using Total Variation Minimization and the H^{-1} Norm, *SIAM Journal on Multiscale Modeling and Simulation*, **1**(3):349−370, 2003.

[PM90] P. Perona, J. Malik. Scale Space and Edge Detection Using Anisotropic

Diffusion, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **12**:629−639, 1990.

[PSM94] P. Perona, T. Shiota, J. Malik. Anisotropic Diffusion, B.M. ter Haar Romeny (Ed.), *Geometry-Driven Diffusion in Computer Vision*, Kluwer, 72−92, 1994.

[PR99] T. Preusser, M. Rumpf. An Adaptive Finite Element Method for Large Scale Image Processing, *Proceedings of Scale-Space'99*, 223−234, 1999.

[ROF92] L. Rudin, S. Osher, E. Fatemi. Nonlinear Total Variation based Noise Removal Algorithm, *Physica D*, **60**:259−268, 1992.

[SBCEV05] P. Schults, E. Bollt, R. Chartrand, S. Esedoglu, K. Vixie. Graduated Adaptive Image Denoising: Local Compromise Between Total Variation and Isotropic Diffusion, *Preprint*, 2005.

[Sc94] C. Schnörr. Unique Reconstruction of Piecewise Smooth Images by Minimizing Strictly Convex Non-Quadratic Functionals, *Journal of Mathematical Imaging and Vision*, **4**(2):189−198, 1994.

[SC03] D. Strong, T. Chan. Edge-Preserving and Scale-Dependent Properties of Total Variation Regularization, *Inverse Problems*, **19**:165−187, 2003.

[SC96] D. Strong, T. Chan. Spatially and Scale Adaptive Total Variation Based Regularization and Anisotropic Diffusion in Image Processing*, Technical Report 46*, *University of California Los Angeles*, Los Angeles, 1996.

[So03] B. Song. Topics in Variational PDE Image Segmentation, Inpainting and Denoising. *PhD thesis*, University of California Los Angeles, 2003.

[VO03] L. Vese, S. Osher. Modeling Textures with Total Variation Minimization and Oscillating Patterns in Image Processing, *Journal of Scientific Computing*, **19**(1-

3):553−572, 2003.

[VO96] C.R. Vogel, M.E. Oman. Iterative Methods for Total Variation Denoising, *SIAM Journal on Scientific Computing*, **17**(1):227−238, 1996.

[We01b] J. Weickert. Applications of Nonlinear Diffusion in Image Processing and Computer Vision, *Acta Mathematica Universitatis Comenianae*, **LXX**(1):33−50, 2001.

[We01a] J. Weickert. Efficient Image Segmentation Using Partial Differential Equations and Morphology, *Pattern Recognition*, 34(9):1813−1824, 2001.

[We99] J. Weickert. Coherence-Enhancing Diffusion of Color Images, *Image and Vision Computing*, **17**:201–212, 1999.

[We98] J. Weickert. Anisotropic Diffusion in Image Processing, *ECMI Series*, Teubner-Verlag, Stuttgart, 1998.

[We97] J. Weickert. A Review of Nonlinear Diffusion Filtering, *Proceedings of the First International Conference on Scale-Space Theory in Computer Vision*, **1252**:3–28, 1997.

[We96b] J. Weickert. Theoretical Foundations of Anisotropic Diffusion in Image Processing, *Computing Suplement*, **11**:221−236, 1996.

[We96a] J. Weickert. Anisotropic Diffusion in Image Processing, *Doktor der Naturwissenschaften enehmigte Dissertation*, Universität Kaiserslautern, Kaiserslautern, 1996.

[We94] J. Weickert. Anisotropic Diffusion Filters for Image Processing Based Quality Control, A. Fasano, M. Primicerio (Eds.), *Proceedings of the Seventh European Conference on Mathematics in Industry*, 355−361, 1994.

[Wi84] A.P. Witkin. Scale-Space Filtering: A New Approach to Multi-Scale Description, *Proceedings of the IEEE International Conference on Acoustic, Speech & Signal Processing*, **9**:150−153, 1984.

[Wi83] A.P. Witkin. Scale-Space Filtering, *Proceedings of the Eight International Conference on Artificial Intelligence*, **2**:1019−1022, 1983.

[WP93] R.T. Whitaker, S.M. Pizer. A Multi-Scale Approach to Non-Uniform Diffusion, *Computer Vision, Graphics, and Image Processing: Image Understanding*, **57**(1):99−110, 1993.

[WRV98] J. Weickert, B.M.t.H. Romeny, M.A. Viergever. Efficient and Reliable Schemes for Nonlinear Diffusion Filtering, *IEEE Transactions on Image Processing*, **7**:398–410, 1998.

[WS00] J. Weickert, C. Schnörr. PDE-Based Preprocessing of Medical Images, *Kunstliche Intelligenz*, **3**:5–10, 2000.

[Xu06] J. Xu. Iterative Regularization and Nonlinear Inverse Scale Space Methods in Image Restoration, *PhD Thesis, University of California Los Angeles*, Los Angeles, 2006.