Nonlinear Anisotropic Diffusion Filtering for Multiscale Edge Enhancement

Stephen L. Keeling* and Rudolf Stollberger†

Abstract. Nonlinear anisotropic diffusion filtering is a procedure based on nonlinear evolution partial differential equations which seeks to improve images qualitatively by removing noise while preserving details and even enhancing edges. However, well known implementations are sensitive to parameters which are necessarily tuned to sharpen a narrow range of edge slopes; otherwise, edges are either blurred or staircased. In this work, nonlinear anisotropic diffusion filters have been developed which sharpen edges over a wide range of slope scales and which reduce noise conservatively with dissipation purely along feature boundaries. Specifically, the range of sharpened edge slopes is widened as backward diffusion normal to level sets is balanced with forward diffusion tangent to level sets. Also, noise is reduced by selectively altering the balance toward diminishing normal backward diffusion and particularly toward total variation filtering. The theoretical motivation for the proposed filters is presented together with computational results comparing them with other nonlinear anisotropic diffusion filters on both synthetic images and magnetic resonance images.

1 Introduction

Nonlinear anisotropic diffusion filtering is a procedure based on nonlinear evolution partial differential equations which seeks to improve images qualitatively by removing noise while preserving details and even enhancing edges. In particular, Perona and Malik [1] proposed the following nonlinear diffusion approach to image filtering:

$$\partial_t u = \nabla \cdot (g(|\nabla u|)\nabla u), \qquad u(0) = u_0 \tag{1.1}$$

to avoid the excessive smoothing associated with linear diffusion techniques. Here, u_0 is the unfiltered image, and after some diffusion time t, $u(t)$ is the filtered image. The nonlinear diffusivity g is a bell-shaped function reducing diffusion to zero as the gradient becomes ever larger, as near object edges. Also, certain bell shapes create forward diffusion where $|\nabla u|$ is sufficiently small, and backward diffusion where $|\nabla u|$ is sufficiently large. Thus, intra-region diffusion is favored over inter-region diffusion. Moreover, the diffusion is direction dependent and is therefore said to be *anisotropic*. On the other hand, diffusion processes are sometimes called anisotropic when g is a diffusion tensor [2] and then even the flux $g\nabla u$ may not be parallel to ∇u. In this paper, only scalar diffusivities are considered.

The first concern about the Perona-Malik approach was that the backward diffusion raises well-posedness issues [3], [4]. These bring into question such practical matters as the undesirable enhancement of image noise [1], and the divergence of solutions starting from very close initial images [5]. Paradoxically, in spite of the apparent ill-posed nature of the differential equation, discretizations have been found to be stable [3]. The primary instability observed numerically is the *staircasing effect*, where a blurred edge evolves into a collection of linear segments separated by jump discontinuities [2]. Numerical studies of this staircasing have demonstrated that discretization implicitly regularizes the Perona-Malik process [4], [6], [7], [8], [9], [10], [11]. On the other hand, regularization has been introduced explicitly into the Perona-Malik equation to avoid its dependence upon a specific numerical implementation [2], [10], [12].

Detailed comparisons of Perona-Malik diffusivities have been performed with respect to several parameters [7]. The general conclusions are that diffusivities with narrower bell shapes and

---

[*] Institut für Mathematik, Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria; email: Stephen.Keeling@kfunigraz.ac.at; tel: +43–316–380–5156; fax: +43–316–380–9815.

[†] Gemeinsame Einrichtung Magnetresonanz (MR), Karl-Franzens-Universität Graz, Auenbruggerplatz 9, 8036 Graz, Austria; email: Rudolf.Stollberger@kfunigraz.ac.at; tel: +43–316–385–3529; fax: +43–316–385–3164.

greater decay rates give rise to sharper edges which persist over longer time intervals. However, these aspects must be carefully coordinated to avoid staircasing. Regularization naturally impedes the staircasing process but also makes a filter more insensitive to smaller structures. It has been proposed further to decrease the regularization [13] and the width of the bell-shaped diffusivity [14] as t increases. The sensitivity of the Perona-Malik approach to such parameters is of primary concern in the present paper.

Indeed, this paper is motivated by the observation that well known implementations of the Perona-Malik approach are sensitive to parameters which are necessarily tuned to enhance a narrow range of edge slopes; otherwise, edges are either blurred or staircased [2]. The purpose of this work is to develop nonlinear anisotropic diffusion filters which sharpen edges over a wide range of slope scales and which reduce noise while conserving feature boundaries. To this end, it has been found that while a greater diffusivity decay rate will create sharper edges in a narrower range of edge slopes, a more gradual diffusivity decay rate will sharpen edges over a wider range of edge slopes. In general, on the basis of computational experiments and related analysis, it is proposed here that multiscale edge enhancement can best be achieved as backward diffusion normal to level sets is balanced with forward diffusion tangent to level sets. It is further proposed that noise be reduced by selectively altering the balance toward diminishing normal backward diffusion and particularly toward *total variation* (TV) filtering discussed below. Filters so contructed are herein labeled *balanced forward backward* (BFB) filters. For perhaps the earliest combination of forward and backwark diffusion along and normal to level sets, respectively, see [15], [16].

Other approaches to nonlinear diffusion filtering include the minimization of functionals of the form:

$$J(u) = \int_P \phi(|\nabla u|)dx + \frac{\mu}{2}\int_P |u_0 - u|^2 dx \tag{1.2}$$

for which the necessary optimality condition is expressed in the steady state for the following descent minimization method:

$$\left\{ \begin{array}{rcll} \partial_t u &=& \nabla \cdot \left(\phi'(|\nabla u|) \dfrac{\nabla u}{|\nabla u|} \right) + \mu(u_0 - u), & P \times [0, \infty) \\ u_n &=& 0, & \partial P \times [0, \infty) \\ u(0) &=& u_0, & P \end{array} \right. \tag{1.3}$$

where P is the image domain and u_n is the normal derivative of u at the boundary ∂P. Here, ϕ can be chosen (non-convex) to create a bell-shaped diffusivity as discussed above [17], or it can be chosen (linear) to create TV regularization [18]. However, this variational formulation is more suited to image restoration problems where the reaction term, $\mu(u_0 - u)$, holds the filtered image near the initial image instead of allowing u to diffuse away from u_0 toward a qualitatively preferred state. Note that the reaction term can actually contribute to staircasing [7].

Another approach to edge enhancement has been developed based upon *shock filtering* [19]:

$$\partial_t u = -|\nabla u| F(\mathcal{L}(u)), \qquad u(0) = u_0. \tag{1.4}$$

Here, $\mathrm{sgn}(F(s)) = \mathrm{sgn}(s)$, and $\mathcal{L}(u)$ is a second order elliptic operator (such as $u_{\nu\nu}$ below) designed to define edges at $\mathcal{L}(u) = 0$. According to this filter, level sets propagate in the direction of ∇u with speed $F(\mathcal{L}(u))$ and thereby accumulate at edges where shocks are formed. Unfortunately, shocks can also be created by noise and so regularizations to shock filtering have been developed [20].

In this paper, Perona-Malik filters are compared with BFB filters, first on synthetic images containing edge gradient profiles varying greatly with respect to both blurredness and edge height. Then these filters are compared on a high resolution magnetic resonance image with a rich variety of features and edge types but also with a significant noise level. The calculations demonstrate that the BFB filters perform quite well in relation to the others. Specifically, the new methods sharpen edges over a broader range of slope scales and they create sufficient tangential smoothing to reduce noise more conservatively.

2 Perona-Malik Filters

Table 1 lists several established Perona-Malik diffusivities along with references and related functions defined in this paper. In each case, the flux function $\Phi(s) = sg(s)$ is increasing for $0 < s < \lambda$ and decreasing for $s > \lambda$. The previously mentioned forward-backward diffusion resulting from such bell-shaped diffusivities can be revealed by decomposing the diffusion in (1.1) as follows [22]:

$$\partial_t u = \Phi'(|\nabla u|)u_{\nu\nu} + g(|\nabla u|)(\nabla^2 u - u_{\nu\nu}) \tag{2.1}$$

Table 1: Perona-Malik diffusivities $g(s)$, their associated variational penalty functions $\phi(s)$, and corresponding flux function derivatives $\Phi'(s)$. These diffusivities will be denoted by PM1, GR, PM2, and W, respectively.

$g(s) = \phi'(s)/s$	$\phi(s)$	$\Phi'(s) = \phi''(s)$
$\left[1 + \left(\frac{s}{\lambda}\right)^2\right]^{-1}$ [1]	$\frac{\lambda^2}{2} \log\left[1 + \left(\frac{s}{\lambda}\right)^2\right]$	$\dfrac{1 - \left(\frac{s}{\lambda}\right)^2}{\left[1 + \left(\frac{s}{\lambda}\right)^2\right]^2}$
$\left[1 + \frac{1}{3}\left(\frac{s}{\lambda}\right)^2\right]^{-2}$ [21]	$\frac{s^2}{2}\left[1 + \frac{1}{3}\left(\frac{s}{\lambda}\right)^2\right]^{-1}$	$\dfrac{1 - \left(\frac{s}{\lambda}\right)^2}{\left[1 + \frac{1}{3}\left(\frac{s}{\lambda}\right)^2\right]^3}$
$\exp\left[-\frac{1}{2}\left(\frac{s}{\lambda}\right)^2\right]$ [1]	$\lambda^2\left(1 - \exp\left[-\frac{1}{2}\left(\frac{s}{\lambda}\right)^2\right]\right)$	$\dfrac{1 - \left(\frac{s}{\lambda}\right)^2}{\exp\left[\frac{1}{2}\left(\frac{s}{\lambda}\right)^2\right]}$
$1 - \exp\left[-\gamma\left(\frac{s}{\lambda}\right)^{-8}\right]$ $e^\gamma = 1 + 8\gamma$ [2]	$\frac{s^2}{2}\left(1 - \exp\left[-\gamma\left(\frac{s}{\lambda}\right)^{-8}\right]\right)$ $+ \left[\frac{\gamma}{16}\left(\frac{s}{\lambda}\right)^{-8}\right]^{\frac{1}{4}} \Gamma\left[\frac{3}{4}, \gamma\left(\frac{s}{\lambda}\right)^{-8}\right]$	$1 - \dfrac{1 + 8\gamma\left(\frac{s}{\lambda}\right)^{-8}}{\exp\left[\gamma\left(\frac{s}{\lambda}\right)^{-8}\right]}$

where $u_{\nu\nu}$ is the second derivative of u in the direction of ∇u. Thus, with respect to level sets of u, the first term captures normal diffusion and the second captures tangential diffusion. Therefore, according to the shape of Φ, diffusion normal to the level sets of u is forward where $|\nabla u| < \lambda$ and it is backward where $|\nabla u| > \lambda$. Given this significance of the (contrast) parameter λ, note that the diffusivities in Table 1 are listed in the order of increasing decay rates, $\max_{s \geq \lambda} |g'(s)|$, in the backward diffusion regime. Furthermore, the order reflects increasing flux function decay rates, $|\Phi'(\chi)|$, where χ is the flux function inflection point discussed below.

As stated earlier, rather than implicitly regularize the differential equation with a particular discretization, regularization has been introduced explicitly, as in the following spatial regularization approach [12]:

$$\partial_t u = \nabla \cdot (g(|\nabla u_\sigma|)\nabla u), \qquad u(0) = u_0 \tag{2.2}$$

where $u_\sigma = K_\sigma * u$ is a smoothing of u obtained by convolving it with a Gaussian kernel K_σ of width σ. Equivalently [2], u_σ is $v(\sigma^2/2)$ where $v(s)$ satisfies: $\partial_s v = \nabla^2 v$ with $v(0) = u$.

3 Proposed Multiscale Edge Enhancement Filters

Figure 1a shows a plot of a function u_0 with several edges varying greatly with respect to both blurredness and edge height. Specifically, u_0 is normalized to have $\min\{u_0\} = 0$ and $\max\{u_0\} = 1$, and it is a sum of six functions of the form $a(1 + bx^2)^{-c}$ chosen to create ever weaker edges. Treating u_0 with any of the filters presented in the previous section leads to the conclusion that these edges cannot all be sharpened simultaneously. However, investigation naturally raises the prospect of creating a presumed ideal diffusivity, as shown in figure 1d, either by locally scaling $|u_0'|$ in figure 1b, or by locally adapting the parameters of a typical

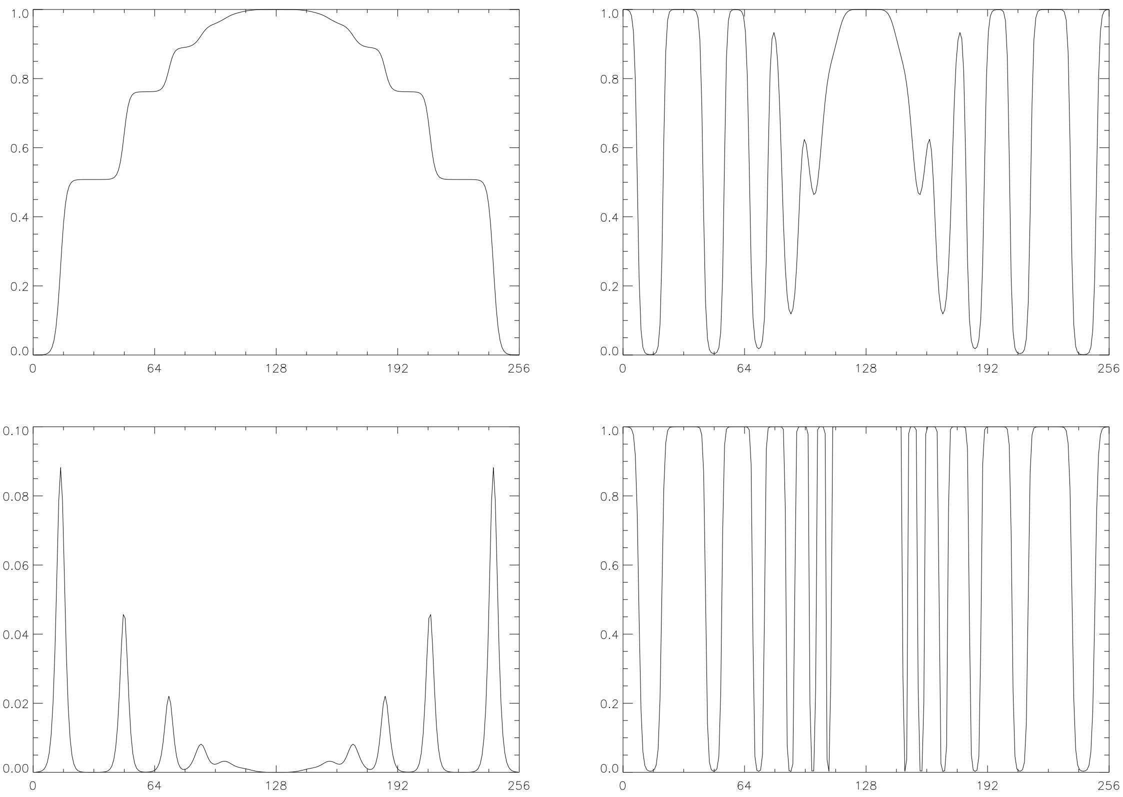

Figure 1: (a) TOP LEFT: a function u_0 with varying edge types, (b) BOTTOM LEFT: its absolute derivative $|u_0'|$, (c) TOP RIGHT: the Table 1 diffusivity, PM1, calculated from $|u_0'|$, and (d) BOTTOM RIGHT: an idealized diffusivity for u_0.

diffusivity shown in figure 1c. Indeed, some success was achieved with such approaches for 1D images. However, results for 2D images were not sufficiently robust.

Continuum level analysis of smooth images (cf. [1], [7], and [13]) shows first that the diffusivity should be chosen so that u_ν increases at edges, i.e., so that $\partial_t u_\nu > 0$. Furthermore, to avoid staircasing requires that $\partial_t u_\nu$ be locally maximal in the normal direction, i.e., that $\partial_t u_{\nu\nu} = 0$ and $\partial_t u_{\nu\nu\nu} < 0$ at edges. Assuming that normal derivatives of even order vanish at edges leads from (2.1) to the following explicit formulation of these conditions for $P \subset \boldsymbol{R}^1$:

$$\begin{array}{rcl} \partial_t u_\nu & = & \Phi' u_{\nu\nu\nu} \quad > \quad 0 \\ \partial_t u_{\nu\nu} & = & 0 \\ \partial_t u_{\nu\nu\nu} & = & 3\Phi'' u_{\nu\nu\nu}^2 + \Phi' u_{\nu\nu\nu\nu\nu} \quad < \quad 0. \end{array} \tag{3.1}$$

Assuming that normal derivatives of odd order alternate in sign implies that these conditions can be met with $\Phi' < 0$ and $\Phi'' < 0$. This means that only a narrow range of edge slopes are sharpened by Perona-Malik diffusion. Specifically, edge slopes with $|u_x| < \lambda$ are blurred, edge slopes with $\lambda < |u_x| < \chi$ are heightened at a locally maximal rate which leads to sharpening, and edge slopes with $|u_x| > \chi^+$ (i.e., $\Phi'' > \frac{1}{3}|\Phi'|u_{\nu\nu\nu\nu\nu}u_{\nu\nu\nu}^{-2}$) are heightened at a locally minimal rate which leads to staircasing. In higher dimensions, the situation is more complex since tangential diffusion can contribute to edge development by flattening homogeneous regions, and it can also contribute to edge degradation by drawing edge mass tangentially away.

Aside from this continuum level analysis, it is found in practice that separate diffusivities with the same values for λ and χ can still provide very different edge enhancement results. In particular, computations demonstrate that the best sharpening effect results when a diffusivity is well suited to the image gradient profile near an edge. This edge gradient profile in turn is influenced not only by the image blurredness, but also by the fact that discretization alone creates different slopes for edges of different heights.

Necessarily constrained by the findings above, computational experiments performed for this work led nevertheless to the speculation that the most uniformly effective diffusivity would be that for which the normal diffusion coefficient Φ' in (2.1) would be as uniformly negative as possible. For example, if M is the maximum possible gradient magnitude on a given grid, then defining $g(s) = (M - s)/s$ gives $\Phi'(s) = -1$. Again, some success was achieved applying this diffusivity to 1D images. However, since the tangential diffusion coefficient g is so much larger in magnitude than the normal diffusion coefficient Φ', 2D results exhibited excessive tangential smoothing. Such tangential smoothing can be explicitly subtracted away, as seen in (1.4) with $\mathcal{L}(u) = u_{\nu\nu}$ and $F(s) = s$, but a dissipative mechanism is desirable here at least for noise reduction. These considerations suggested to balance smoothing in the tangential direction with sharpening in the normal direction as follows:

$$-1 = \frac{\Phi'(s)}{g(s)} = \frac{sg'(s) + g(s)}{g(s)} \quad \text{to obtain} \quad g(s) = \frac{1}{s^2}. \tag{3.2}$$

Computational results in the next section demonstrate that this *balanced forward-backward* (BFB) diffusivity can sharpen edges over a wide range of slopes. On the other hand, to control the enhancement of noise, the balance can be introduced gradually as follows:

$$-\frac{s}{\kappa + s} = \frac{\Phi'(s)}{g(s)} = \frac{sg'(s) + g(s)}{g(s)} \quad \text{to obtain} \quad g(s) = \frac{1}{s(\kappa + s)}. \tag{3.3}$$

This *parameterized BFB-κ* diffusivity is shown with the BFB diffusivity in Table 2, followed by Gaussian [23] and TV [18] diffusivities in Table 3, to demonstrate the respective thresholds met concerning the relative amounts of normal and tangential diffusion. Observe that the BFB-κ diffusivity follows the TV model s^{-1} for gradients that are small in relation to κ, and then it passes to the BFB model s^{-2} for gradients that are large in relation to κ. By comparison, Perona-Malik diffusivities follow the Gaussian model for small gradients where $\Phi' > 0$ and $\Phi''(0) = 0$; however, only the PM1 diffusivity follows the BFB model for gradients that are large in relation to λ. Also, since Perona-Malik flux functions are first increasing and then decreasing, their variational penalty functions are neither purely convex nor purely concave. On the other hand, the flux functions for the Table 2 diffusivities are globally non-increasing and their variational penalty functions are globally concave ($\Phi' = \phi'' \leq 0$). Since such *concave filters* never follow the Gaussian model, they do not introduce normal smoothing for any gradient magnitudes. Paradoxically, in spite of such extremely limited dissipation in concave filters, they are seen below to be remarkably stable and effective for image enhancement.

Table 2: Concave BFB and BFB-κ filter diffusivities $g(s)$, along with their associated variational penalty functions $\phi(s)$, corresponding flux function derivatives $\Phi'(s)$, and quotients of normal and tangential diffusion coefficients.

$g(s) = \phi'(s)/s$	$\phi(s)$	$\Phi'(s) = \phi''(s)$	$\Phi'(s)/g(s)$
s^{-2}	$\log(s)$	$-s^{-2}$	-1
$[s(\kappa + s)]^{-1}$	$\log(\kappa + s)$	$-(s + \kappa)^{-2}$	$-s/(\kappa + s)$

Table 3: Non-concave Gaussian and TV filter diffusivities $g(s)$, along with their associated variational penalty functions $\phi(s)$, corresponding flux function derivatives $\Phi'(s)$, and quotients of normal and tangential diffusion coefficients.

$g(s) = \phi'(s)/s$	$\phi(s)$	$\Phi'(s) = \phi''(s)$	$\Phi'(s)/g(s)$
1 [23]	$\frac{1}{2}s^2$	1	1
s^{-1} [18]	s	0	0

4 Numerical Methods

Filters based upon the diffusivities in tables 1-3 and the spatial regularizations of them were implemented numerically for this study as follows. Since most implementations are based upon finite differences and single-step time stepping [2], this practice is followed here as well; however, because of the requirements of the unbounded diffusivities, implicit time stepping is implemented uniformly instead of using the more usual explicit methods where they might be applicable. All cells are assumed to have unit aspect ratios and width h. For planar images, the diffusion operator is discretized diagonally at the cell centroid $(x_i, y_j) = ((i-\frac{1}{2})h, (j-\frac{1}{2})h)$, where $1 \leq i, j \leq N$, as follows:

$$\begin{aligned} & 2h^2[\nabla \cdot (g(|\nabla V|)\nabla U)]_{i,j} \\ & \quad = \; g(|\nabla V|_{i+\frac{1}{2},j+\frac{1}{2}})\,[U_{i+1,j+1} - U_{i,j}] - g(|\nabla V|_{i-\frac{1}{2},j-\frac{1}{2}})\,[U_{i,j} - U_{i-1,j-1}] \qquad (4.1) \\ & \quad + \; g(|\nabla V|_{i-\frac{1}{2},j+\frac{1}{2}})\,[U_{i-1,j+1} - U_{i,j}] - g(|\nabla V|_{i+\frac{1}{2},j-\frac{1}{2}})\,[U_{i,j} - U_{i+1,j-1}]\,. \end{aligned}$$

Here, the notation $U(x_i, y_j, t) = U_{i,j}(t)$ is used for an approximation to $u(x_i, y_j, t)$, and $U^n_{i,j}$ is used to denote an approximation to $u(x_i, y_j, t^n)$ at the nth time level $t^n = n\tau$. Also, $\bar{U}(t)$ and $\bar{U}^n$ represent the vectors $\{U_{i,j}(t)\}$ and $\{U^n_{i,j}\}$. For diffusivity computations, V is U or a smoothing thereof, and the gradient terms $|\nabla V|$ are computed at cell corners as follows:

$$2[h|\nabla V|_{i+\frac{k}{2},j+\frac{l}{2}}]^2 = |V_{i+k,j+l} - V_{i,j}|^2 + |V_{i+k,j} - V_{i,j+l}|^2, \quad k, l = \pm 1. \qquad (4.2)$$

Here, values of U and V with out-of-range indices are understood equal to their reflections across the image boundary, and values of g at cell corners on the boundary are understood to be zero. As a result of computing the discrete diffusion operator at the boundary in this way, natural no-flux boundary conditions (from the variational formulation in (1.2)) are implemented implicitly. Also, the discrete diffusion can be written as the product $h^{-2}G(\bar{V})\bar{U}$ for a symmetric matrix G. These discretizations are defined similarly in other spatial dimensions.

Note that, as discussed in [17], the diagonal discretization in (4.1) involves interleaved (red-black) diagonal grids coupled through the diffusivity calculated from the gradient in (4.2). In some cases, the coupling can be weak enough that diagonal artifacts may result; however, for images with mildly well behaved statistics, this coupling is sufficient. On the other hand, horizontal-vertical discretizations necessarily involve a broader stencil for the gradient calculation; consequently, they are more dissipative and conspicuously so where the gradient is diagonally oriented.

Assume that the unbounded diffusivities are computed numerically as $g(\max\{s, \varepsilon\})$ to avoid division by zero. Here, ε is not viewed as an additional parameter as it is simply chosen near machine zero or small enough for results to be insensitive to its variation.[1] As with TV filtering, other regularizations can be used such as $g(\sqrt{s^2+\varepsilon^2})$ [24], but the decay rates of the original diffusivities are given special attention here.

Now, with constructions analogous to those established here, define U_σ as the solution to the discrete heat equation ($g = 1$) up to time $\sigma^2/2$ starting from U, and let $V = U_\sigma$ for a given $\sigma \geq 0$. Finally, with $\bar{U}_0$ denoting the initial image values, the semidiscrete formulation of the filter is $\bar{U}' = h^{-2}G(\bar{U}_\sigma)\bar{U}$ with $\bar{U}(0) = \bar{U}_0$, and a stable fully discrete formulation is obtained with the following implicit time stepping:

$$[I - \rho G(\bar{U}^n_\sigma)]\,\bar{U}^{n+1} = \bar{U}^n, \qquad \bar{U}^0 = \bar{U}_0 \qquad (4.3)$$

where $\rho = \tau h^{-2}$. These time stepping equations are solved only approximately, with little extra computational expense, using just a few iterations of a Jacobi preconditioned conjugate gradient

---

[1] The details of this insensitivity to variations in a sufficiently small ε will be addressed elsewhere in a treatment of well-posedness issues for concave filters.

method [25]. Other forms of preconditioning, such as Cholesky incomplete factorization [25], were not as rapid or as accurate for this system. Even the presently adopted preconditioning required here for accuracy is a modification $\tilde{D}$ of the diagonal D of $\left[I - \rho G(\bar{U}^n_\sigma)\right]$. Specifically, the definition of the discrete planar diffusion operator in (4.1) gives:

$$D_{i,j} = 1 + \frac{\rho}{2}\left(g_{i+\frac{1}{2},j+\frac{1}{2}} + g_{i-\frac{1}{2},j-\frac{1}{2}} + g_{i-\frac{1}{2},j+\frac{1}{2}} + g_{i+\frac{1}{2},j-\frac{1}{2}}\right) \tag{4.4}$$

where values of g at cell corners on the image boundary are understood to be zero. For $\tilde{D}$, these boundary values of g are the same as at the nearest cell corners off the boundary.

5 Computational Results

All the computations reported in this section were performed using the IDL[2] (Interactive Data Language) system. In every example, $N = 256$ and $h = 1$; hence, for diffusivity evaluations $\lambda h = \lambda$, and for time stepping $\rho = \tau$. For every calculation involving spatial regularization, a single explicit time step was used to calculate $\bar{U}^n_\sigma = \left[I + \frac{\sigma^2}{2h^2}G(\bar{U}^n)\right]\bar{U}^n$ with $\sigma = \sqrt{0.8}$. Except where otherwise stated, (4.3) was treated with two conjugate gradient iterations after the initial gradient step.

Figure 2 illustrates the results of filtering the function u_0 of figure 1a with the Table 1 diffusivities and their spatial regularizations. To avoid losing the weakest edges, the contrast parameter used for all these diffusivities was $\lambda = 0.003$. The course of image evolution is similar for each diffusion process in the sense that weaker edges are sharpened sooner than stronger edges. Also, each sharpened edge eventually disintegrates and becomes blurred in relation to a stronger neighbor until the whole image reaches a constant state after a very long time. Moreover, the separation between these events for the respective edge strengths, as well as the total duration of the evolution process, are very much longer with each increase in decay rate in Table 1.

The plots displayed in figure 2 were obtained by choosing $\tau = 1.0$ and taking $n = 1000$ time steps to reach $t = 1000$, a time by which each solution was relatively steady. The final-time relative L_1-norm measure of differences between successive iterates ($\|U^n - U^{n-1}\|_{L_1}/\|U^{n-1}\|_{L_1}$) was, for figures 2a – 2h, 6×10^{-6}, 2×10^{-6}, 7×10^{-8}, 3×10^{-8}, 3×10^{-5}, 3×10^{-6}, 3×10^{-7}, and 4×10^{-7}, respectively. Note that in no case were all the edges sharpened equally well; however, the PM1 filter used for figure 2a appears to provide the most uniformly effective sharpening. This performance can be understood from the relation to the BFB model, but more immediately from the fact that, among the diffusivities in Table 1, PM1 has the smallest $|\Phi'(\chi)|$ and so the most uniformly negative Φ' in the backward diffusion regime. Even though the continuum analysis accompanying (3.1) predicts sharpening in figure 2a only for slopes between λ and $\chi = \lambda\sqrt{3}$, the numerical results are more favorable. Nevertheless, only one accurately enhanced edge pair appears third from the boundary in figure 2a where the inflection points of the solution agree with one of the six pairs of u_0. Similarly, one accurately enhanced edge pair appears remaining innermost in figures 2b – 2d. Otherwise, various degrees of staircasing are evident in all figures 2a – 2d, and greater staircasing can be induced by choosing ever smaller time steps. The staircasing appearing in figures 2a – 2d occurs because most edge slopes are large in relation to λ, so they turn too slowly in relation to neighboring slopes. A larger λ can be chosen to sharpen stronger edges better, but then weaker edge slopes are diminished. Adding spatial regularization rounds and reduces the number of staircase steps until sufficient regularization prevents the development of the steps as seen in figures 2e – 2h where spatial regularization was performed. However, the smoothing from spatial regularization can evidently flatten small-slope regions and retard edge sharpening.

On the other hand, figure 3 shows the results of filtering u_0 with the Table 2 diffusivities and their spatial regularizations. For figures 3b and 3d, $\kappa = 0.003$ as with λ for figure 2.

---

[2] See `http://www.rsinc.com/idl/index.cfm`.

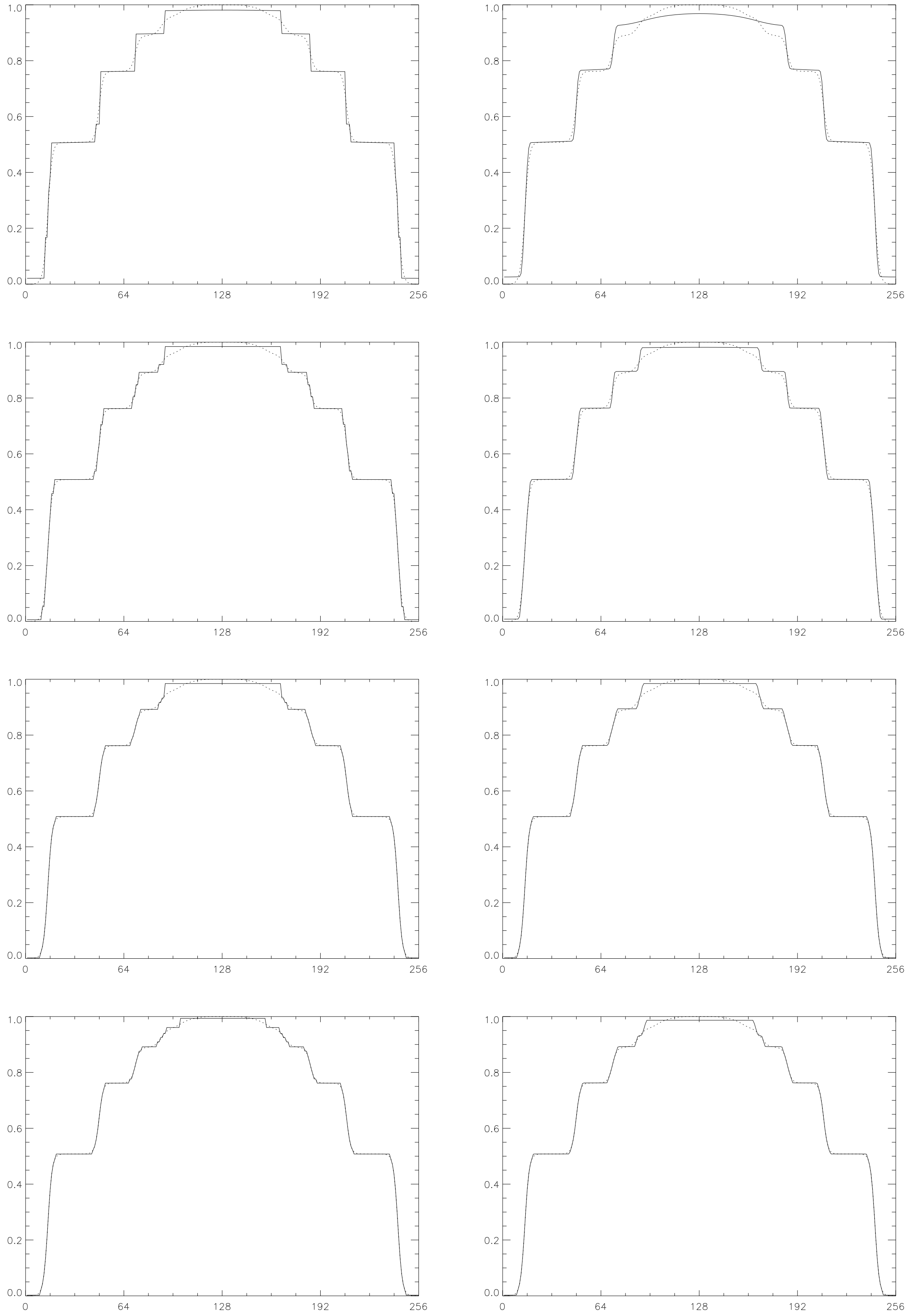

Figure 2: (a – d) LEFT COLUMN: the results of filtering u_0 in figure 1a (shown dotted here) with the Table 1 diffusivities respectively, and (e – h) RIGHT COLUMN: the corresponding results for the spatial regularizations of these diffusivities.

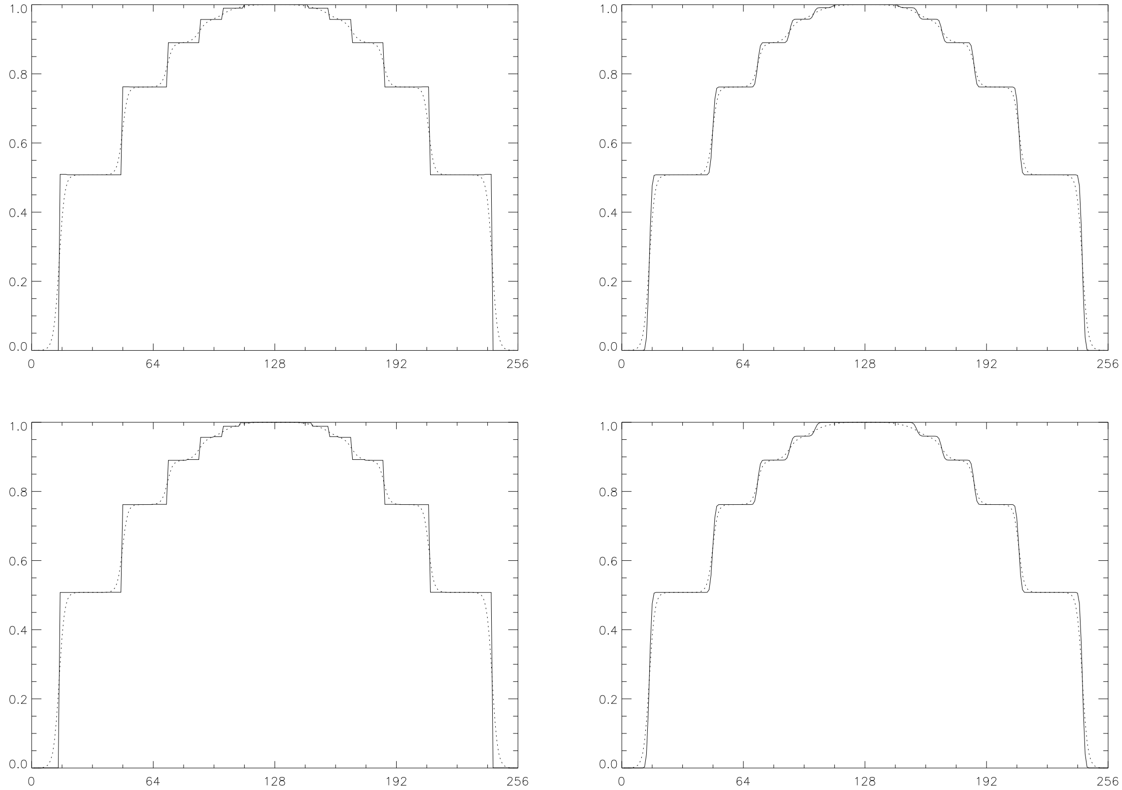

Figure 3: (a – b) LEFT COLUMN: the results of filtering u_0 in figure 1a (shown dotted here) with the Table 2 diffusivities respectively, and (c – d) RIGHT COLUMN: the corresponding results for the spatial regularizations of these diffusivities.

Here, the course of image evolution is much simpler than for figure 2. All edges are sharpened simultaneously and within very few iterations. All plots in figure 3 were obtained by choosing $\tau = 1.0$ and taking $n = 1000$ time steps to reach $t = 1000$, as for figure 2. The final-time relative L_1-norm measure of differences between successive iterates was, for figures 3a – 3d, 1×10^{-18}, 9×10^{-11}, 2×10^{-7}, and 1×10^{-6}, respectively. However, for figure 3a the difference was already 6×10^{-18} after only $n = t = 23$. For figure 3b the image was already quite stable after only $n = t = 70$ when the difference was 1×10^{-7}. Note in figures 3a and 3b that every edge is sharpened quite well. In fact, in figure 3a, the numerically significant inflection points of the solution are identical to those of u_0. In figure 3b, only the innermost inflection points of the solution differ from those of u_0 and only by one cell. The continuum analysis accompanying (3.1) predicts very ineffective sharpening here since Φ'' is always positive, but the numerical results are certainly more favorable. Here, as the time step is reduced through extremely small values, staircasing can be induced, and earlier at the more gradual slopes. Also, the number of staircase steps increases as the time step decreases. On the other hand, the number of staircase steps can be reduced with larger time steps and a sufficient number of conjugate gradient iterations, until even the most gradually sloped edges can be sharpened into a single step. As with the other filters, adding spatial regularization rounds and reduces the number of staircase steps. Eventually, sufficient smoothing from spatial regularization can flatten small-slope regions and retard edge sharpening as seen in figure 3d, but the one-step regularization has only mildly rounded the edges in figure 3c.

Now consider the planar function u_0 shown in figure 4 with several edges varying greatly with respect to both blurredness and edge height. As with figure 1, here u_0 is normalized to have $\min\{u_0\} = 0$ and $\max\{u_0\} = 1$, and it is a sum of six functions of the form $a(1 + bx^2 + cy^2)^{-d}$ chosen to create ever weaker edges. This planar function u_0 of figure 4 was filtered with each of

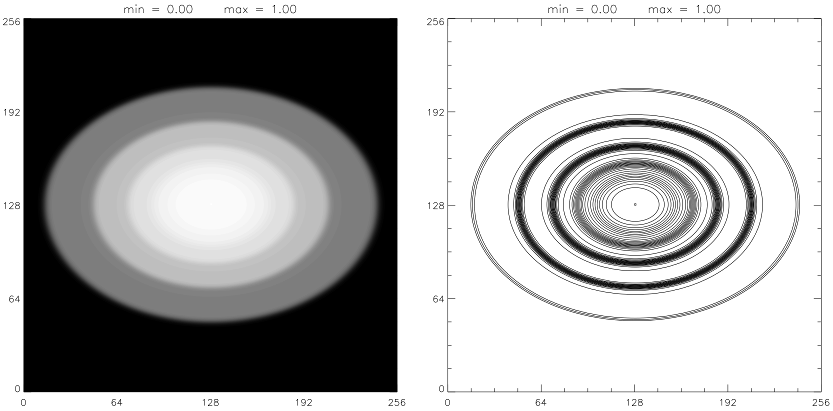

Figure 4: (a) LEFT: a region-filled contour plot with equally spaced levels for a function u_0 with varying edge types, (b) RIGHT: a lined contour plot of u_0 without region filling and with unequally spaced levels $\{(i/i_{\max})^{1/4}\}$ to accentuate small-slope features near the image center.

the Table 1 diffusivities and their spatial regularizations, and the results are shown in figure 5. The contrast parameter used here for these diffusivities was $\lambda = 0.003$ as for figure 2. The course of image evolution for the 2D diffusion processes here is very similar to that for the corresponding 1D diffusion processes except that the time scale is shorter. Here, processes are accelerated since tangential smoothing can quicken edge development and then hasten edge degradation as discussed earlier. All plots in figure 5 were obtained by choosing $\tau = 1.0$ and taking $n = 250$ time steps to reach $t = 250$, a time by which image cross sections appeared similar to their counterparts in figure 2. The final-time relative L_1-norm measure of differences between successive iterates was, for figures 5a – 5h, 4×10^{-5}, 2×10^{-5}, 3×10^{-6}, 3×10^{-6}, 7×10^{-5}, 1×10^{-5}, 3×10^{-6}, 3×10^{-6}, respectively. As with figure 2, none of the filters sharpened all the edges here equally well, but the PM1 filter used for figure 5a appears to provide the most uniformly effective sharpening since level curves are more closely accumulated at edges. Also, staircasing is manifested by the creation of spurious plateaus in all figures 5a – 5d, and not in figures 5e – 5h where spatial regularization was performed.

Figure 6 now illustrates the results of filtering the planar function u_0 shown in figure 4 with the Table 2 diffusivities and their spatial regularizations. For figures 6b and 6d, $\kappa = 0.003$, as for figure 3. All plots in figure 6 were obtained by choosing $\tau = 1.0$ and taking $n = 250$ time steps to reach $t = 250$, as for figure 5. The final-time relative L_1-norm measure of differences between successive iterates was, for figures 6a – 6d, 1×10^{-11}, 1×10^{-7}, 5×10^{-4}, and 5×10^{-4}, respectively. However, for figure 6a the difference was already 6×10^{-10} after only $n = t = 38$. For figure 6b the difference was already 9×10^{-6} after only $n = t = 21$. Note in figures 6a and 6b that, aside from some negligible fluctuations of contour lines in flat zones, every edge is sharpened very well. In fact, in figure 6a, the numerically significant inflection points of the solution along the diagonal lines through the center of the image are identical to those of u_0. Along the horizontal and vertical lines through the center, only a few inflection points differ and by at most half a cell. In figure 6b, along the diagonal, horizontal, and vertical lines through the center of the image, only a few of the inflection points of the solution differ from those of u_0 and by at most one cell. As evident in figures 6c and 6d, the smoothing from spatial regularization not only retards edge sharpening and flattens small-slope regions, but it also rounds formerly elliptical level curves into circular ones.

Results in figures 2 and 5 can be improved by taking larger time steps. In fact, as $\tau \to \infty$ and $\lambda \to 0$ while $\tau\lambda^2 = \hat{\tau}$, the PM1 filter approaches the BFB filter with the time step $\hat{\tau}$. On the other hand, to clearly distinguish these filters by reducing such temporal regularization, figure 7 shows the time-accurate results of filtering u_0 from figure 1a with the PM1 and BFB

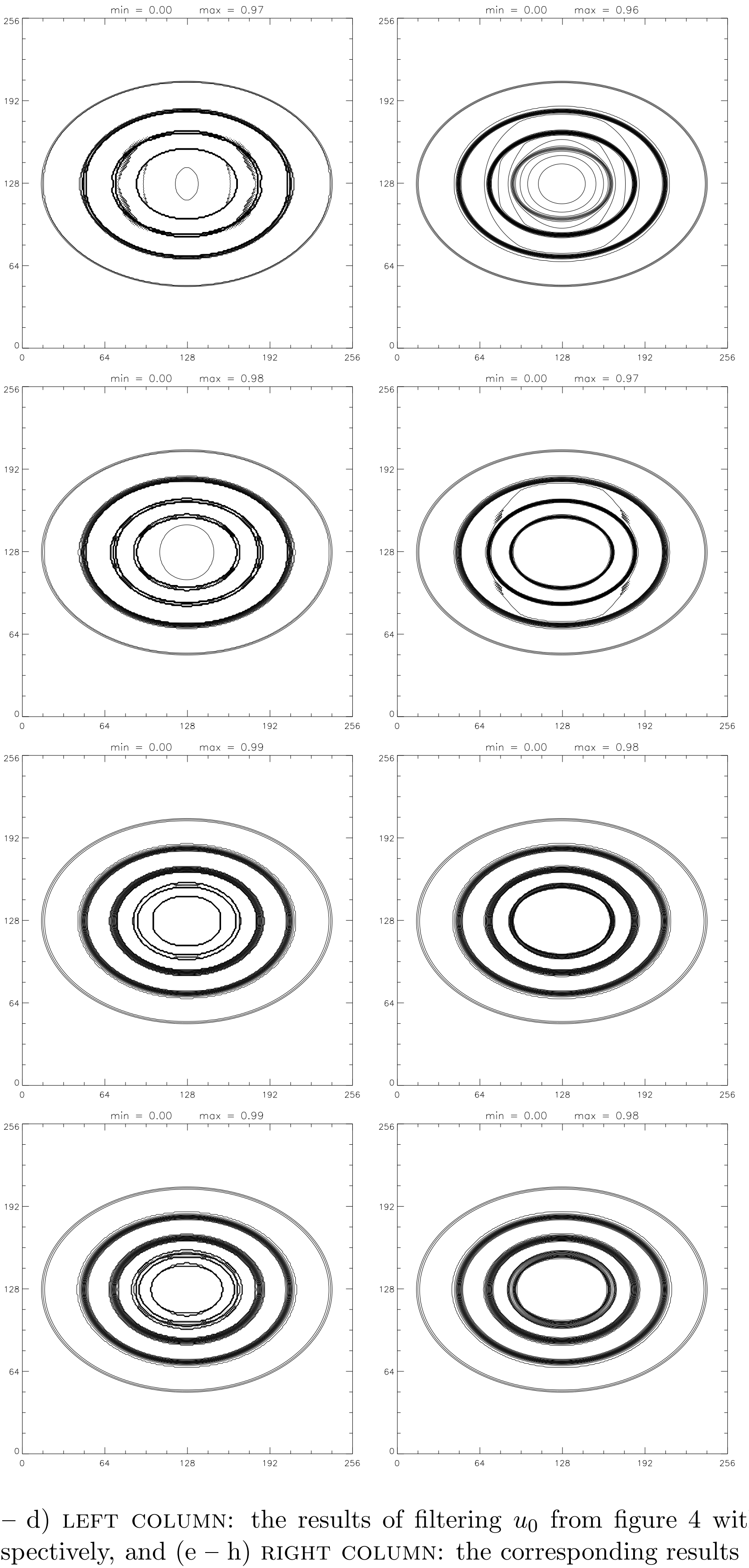

Figure 5: (a – d) LEFT COLUMN: the results of filtering u_0 from figure 4 with the Table 1 diffusivities respectively, and (e – h) RIGHT COLUMN: the corresponding results for the spatial regularizations of these diffusivities. Again, contour spacing $\{(i/i_{\max})^{1/4}\}$ is used.

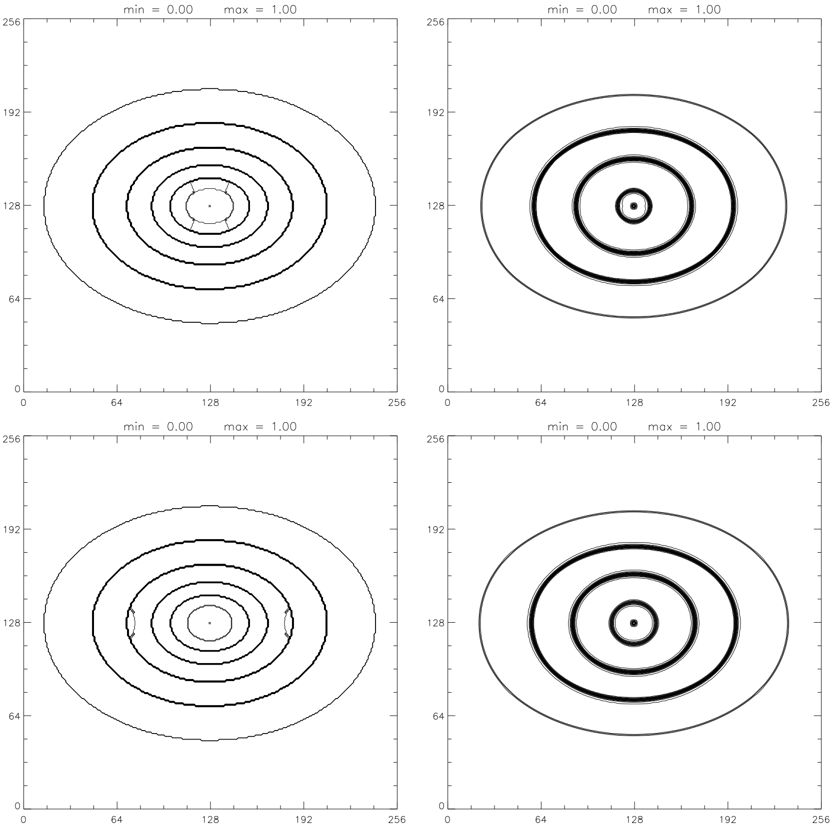

Figure 6: (a – b) LEFT COLUMN: the results of filtering u_0 from figure 4 with the Table 2 diffusivities respectively, and (c – d) RIGHT COLUMN: the corresponding results for the spatial regularizations of these diffusivities. Again, contour spacing $\{(i/i_{\max})^{1/4}\}$ is used.

filters. Specifically, these plots were obtained by choosing $\tau = 10^{-5}$ and taking $n = 10^8$ time steps with 1000 conjugate gradient iterations per time step to reach $t = 1000$, the same final time as in figures 2 and 3. The final-time relative L_1-norm measure of the difference between successive iterates for figure 7a was 2×10^{-10}, and for figure 7b was already machine zero after only $n = 904$. Note that the PM1 filtered result in figure 7a is similar to that shown in figure 2a obtained with larger time steps. In particular, the four outermost edges appear sharpened though partly staircased, while the smaller sloped edges are flattened. On the other hand, the BFB filtered result in figure 7b manifests significantly greater staircasing than that shown in figure 3a, though in neither case are smaller sloped edges flattened. The results of figure 7 reflect the nature of the respective underlying partial differential equations, which are certainly constrained by the continuum analysis outlined in connection with (3.1).

Finally, figure 8 shows a comparison between a Perona-Malik filter and a concave filter when applied to a noisy magnetic resonance image obtained from a high resolution scan of an *in vitro* atherosclerotic vessel. Specifically, figure 8a is the unfiltered image u_0 normalized so that $\min\{u_0\} = 0$ and $\max\{u_0\} = 1$, figure 8b is the result of filtering u_0 with the PM2 diffusivity used often in practice [26], and figure 8c is the result of filtering u_0 with the BFB-κ diffusivity. To reveal filter characteristics directly, no spatial regularization was used. The PM2 filter parameters were chosen just large enough to smooth the background noise without greatly changing the shape of level curves, i.e., $\lambda = 0.015$, $\tau = 0.3$, and $n = 10$ ($t = 3.0$). Similarly, the parameters chosen for the BFB-κ filter were $\kappa = 0.7$, $\tau = 0.001$ and $n = 10$ ($t = 0.01$).

To fully illustrate the differences between the two filters, magnifications of figures 8a – 8c are shown in figures 8d – 8f, respectively, where the zone of magnification is outlined in black in figure 8a. Note that the λ required for the PM2 filter to smooth the background noise also levels the signal into smoothly shaped flat zones bounded by high contrast edges to which λ happens to

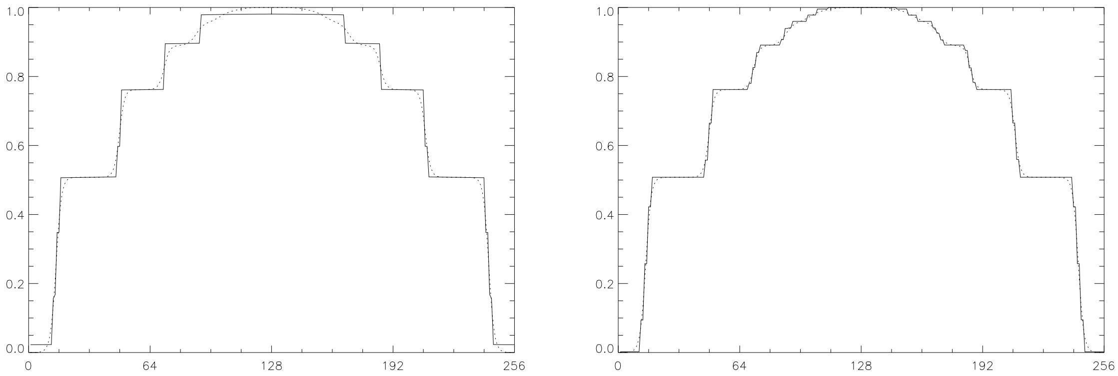

Figure 7: Time-accurate results of filtering u_0 in figure 1a (shown dotted here) with (a) LEFT: the PM1 diffusivity and (b) RIGHT: the BFB diffusivity.

be tuned. The result creates a kind of *watercolor* effect in the image. To avoid the effect, λ can be reduced, but at the cost of a conspicuous enhancement of the higher amplitude noise. On the other hand, the BFB-κ filter appears to capture image features more realistically by preserving edges while significantly reducing the noise level. Even though tangential smoothing is the only explicit dissipative mechanism for this filter, such limited smoothing performs remarkably well at noise reduction. As κ is decreased, the tangential smoothing and normal sharpening are increased simultaneously and images become more segmentation like. For sufficiently small κ, the additional tangential smoothing rounds the shape of level curves in low-gradient regions, and the additional normal sharpening enhances nonphysical edges in the noise. Nevertheless, for both filters considered here, such noise enhancement can be treated with spatial regularization, but at the cost of flattening small-slope regions, retarding edge enhancement, and rounding the shapes of level curves.

6 Conclusions

Nonlinear anisotropic diffusion filters have been developed here which, unlike those previously developed, sharpen edges over a wide range of slope scales and reduce noise conservatively with dissipation purely along feature boundaries. This work has also revealed that a more gradual Perona-Malik diffusivity decay rate will sharpen edges over a wider range of edge slopes. In fact, among the Perona-Malik diffusivities considered for this paper, the most gradually decaying, PM1, gives the most uniformly effective edge sharpening. Furthermore, in the limit of large time step and small contrast parameter, PM1 happens to agree with the BFB diffusivity derived independently here to balance backward diffusion normal to level sets with forward diffusion tangent to level sets. Also, to alter this balance toward diminishing normal backward diffusion for selected small-amplitude (noisy) gradients, the BFB-κ filter was developed to follow the TV model for smaller gradients and the BFB model for larger gradients. In contrast to Perona-Malik filters, the concave filters introduced here never smooth normal to level sets. These concave filters can be used not only to improve the visual clarity of images as explained here, but also for defining object boundaries precisely or as a preprocessor for segmentation. Forthcoming work in this direction will be reported later, as will an analysis of well-posedness issues surrounding concave filters.

Acknowledgments

The authors gratefully acknowledge support for this work from the Fonds zur Förderung der wissenschaftliche Forschung. In particular, the first author was supported under SFB F003,

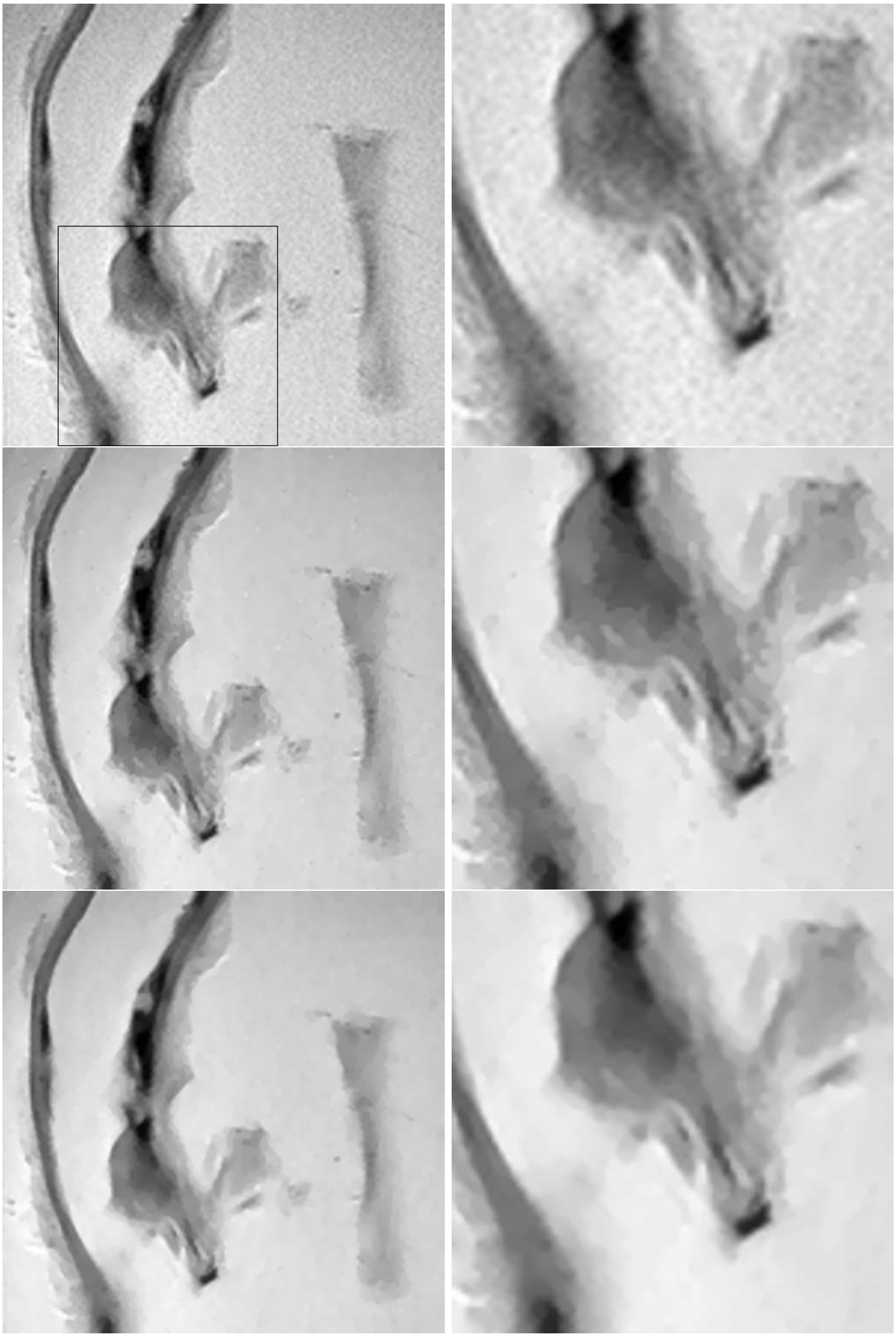

Figure 8: (a) top left: an unfiltered magnetic resonance image u_0 of an atherosclerotic vessel, (b) middle left: the result of filtering u_0 with the PM2 diffusivity, (c) bottom left: the result of filtering u_0 with the BFB-κ diffusivity, and (d – f) right column: magnifications from images (a) – (c), respectively.

“Optimierung und Kontrolle”, and the second author under P 11525-MED. The authors also wish to thank the referees for the keen observations and helpful suggestions made for this paper.

References

[1] P. Perona and J. Malik, *Scale Space and Edge Detection Using Anisotropic Diffusion*, Proc. IEEE Comp. Soc. Workshop on Computer Vision, Miami Beach, November 30 – December 2, 1987, IEEE Computer Society Press, Washington, pp. 16 – 22, 1987.

[2] J. Weickert, *Anisotropic Diffusion in Image Processing*, B. G. Teubner Stuttgart, 1998.

[3] B. Kawohl and N. Kutev, *Maximum and Comparison Principle for One-dimensional Anisotropic Diffusion*, Math. Ann., Vol. 311, pp. 107 – 123, 1998.

[4] S. Kichenassamy, *The Perona-Malik Paradox*, SIAM J. Appl. Math., Vol. 57, No. 5, pp. 1328 – 1342, October 1997.

[5] Y.-L. You, W. Xu, A. Tannenbaum, and M. Kaveh, *Behavioral Analysis of Anisotropic Diffusion in Image Processing*, IEEE Trans. Image Proc., Vol. 5, pp. 1539 – 1553, 1996.

[6] S. Acton, *Edge Enhancement of Infrared Imagery by Way of the Anisotropic Diffusion Pyramid*, Proc. IEEE Int. Conf. Image Processing, ICIP-96, Lausanne, September 16 – 19, 1996, Vol. 1, pp. 865 – 868, 1996.

[7] B. Benhamouda, *Parameter Adaptation for Nonlinear Diffusion in Image Processing*, Master’s Thesis, Dept. of Mathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1994.

[8] G. Cong and S.D. Ma, *Nonlinear Diffusion for Early Vision*, Proc. 13th Int. Conf. Pattern Recognition, ICPR 13, Vienna, August 25 – 30, 1996, Vol. A, pp. 403 – 406, 1996.

[9] J. Fröhlich and J. Weickert, *Image Processing Using a Wavelet Algorithm for Nonlinear Diffusion*, Report No. 104, Laboratory of Technomathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1994.

[10] M. Nitzberg and T. Shiota, *Nonlinear Image Filtering with Edge and Corner Enhancement*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, pp. 826 – 833, 1992.

[11] J. Weickert and B. Benhamouda, *A Semidiscrete Nonlinear Scale-Space Theory and Its Relation to the Perona-Malik Paradox*, F. Solina, W.G. Kropatsch, R. Klette, R. Bajcsy (Eds.), Theoretical Foundations of Computer Vision, TFCV 1996, Dagstuhl, March 18 – 22, 1996, Springer, Wien, pp. 1 – 10, 1997.

[12] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, *Image Selective Smoothing and Edge Detection by Nonlinear Diffusion*, SIAM J. Numer. Anal., Vol. 29, No. 1, pp. 182 – 193, February 1992.

[13] R.T. Whitaker and S.M. Pizer, *A Multi-Scale Approach to Nonuniform Diffusion*, CVGIP: Image Understanding, Vol. 57, pp. 99 – 110, 1993.

[14] X. Li and T. Chen, *Nonlinear Diffusion with Multiple Edginess Thresholds*, Pattern Recognition, Vol. 27, pp. 1029 – 1037, 1994.

[15] D. Gabor, *Information Theory in Electron Microscopy*, Laboratory Investigation, Vol. 14, pp. 801-807, 1965.

[16] M. Lindenbaum, M. Fischer and A. Bruckstein, *On Gabor's Contribution to Image Enhancement*, Pattern Recognition, Vol. 27, pp. 1-8, 1994.

[17] N. Nordström, *Biased Anisotropic Diffusion – A Unified Regularization and Diffusion Approach to Edge Detection*, Image Vision Comput., Vol. 8, pp. 318 – 327, 1990.

[18] L.I. Rudin, S. Osher, and E. Fatemi, *Total Variation Based Noise Removal Algorithms*, Physica D, Vol. 60, pp. 259 – 269, 1992.

[19] S. Osher and L.I. Rudin, *Feature-Oriented Image Enhancement Using Shock Filters*, SIAM J. Numer. Anal., Vol. 27, No. 4, pp. 919 – 940, August 1990.

[20] L. Alvarez and L. Mazorra, *Signal and Image Restoration Using Shock Filters and Anisotropic Diffusion*, SIAM J. Numer. Anal., Vol. 31, No. 2, pp. 590 – 605, April 1994.

[21] S. Geman, G. Reynolds, *Constrained Restoration and the Recovery of Discontinuities*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, pp. 367 – 383, 1992.

[22] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel, *Axioms and Fundamental Equations in Image Processing*, Arch. Rational Mech. Anal., Vol. 123, pp. 199 – 257, 1993.

[23] A.N. Tikhonov and V.Y. Arsenin, *Solutions of Ill-Posed Problems*, Wiley, New York, 1977.

[24] R. Acar and C.R. Vogel, *Analysis of Bounded Variation Penalty Methods for Ill-posed Problems*, Inverse Problems, Vol. 10, pp. 1217 – 1229, 1994.

[25] R. Barrett, M.W. Berry, T.F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, Philadelphia, 1993.

[26] G. Gerig, O. Kübler, R. Kikinis, and F.A. Jolesz, *Nonlinear Anisotropic Filtering of MRI Data*, IEEE Transactions on Medical Imaging, Vol. 11, No. 2, pp. 221 – 232, June 1992.