Computer Vision, Graphics, and Pattern Recognition Group
Department of Mathematics and Computer Science
University of Mannheim
D-68131 Mannheim, Germany

Reihe Informatik
12 ~~8~~/1999

Scale-Space Properties of Nonstationary Iterative Regularization Methods

Esther Radmoser, Otmar Scherzer, and Joachim Weickert

Technical Report 8/1999
Computer Science Series
October 1999

The technical reports of the CVGPR Group are listed under
http://www.ti.uni-mannheim/~bmg/Publications-e.html

Scale-Space Properties of Nonstationary Iterative Regularization Methods

Esther Radmoser[1], **Otmar Scherzer**[1], **and Joachim Weickert**[2]

[1]Industrial Mathematics Institute, University of Linz,
Altenberger Str. 69, A-4040 Linz, Austria
E-Mail: {radmoser, scherzer}@indmath.uni-linz.ac.at
Web Page: http://imagewww.indmath.uni-linz.ac.at

[2]Computer Vision, Graphics, and Pattern Recognition Group,
Department of Mathematics and Computer Science,
University of Mannheim, D-68131 Mannheim, Germany
E-Mail: Joachim.Weickert@ti.uni-mannheim.de
Web Page: http://www.ti.uni-mannheim/~bmg/weickert

Abstract

Most scale-space concepts have been expressed as parabolic or hyperbolic partial differential equations (PDEs). In this paper we extend our work on scale-space properties of elliptic PDEs arising from regularization methods: we study linear and nonlinear regularization methods that are applied iteratively and with different regularization parameters. For these so-called nonstationary iterative regularization techniques we clarify their relations to both isotropic diffusion filters with a scalar-valued diffusivity and anisotropic diffusion filters with a diffusion tensor. We establish scale-space properties for iterative regularization methods that are in complete accordance with those for diffusion filtering. In particular, we show that nonstationary iterative regularization satisfies a causality property in terms of a maximum–minimum principle, possesses a large class of Lyapunov functionals, and converges to a constant image as the regularization parameters tend to infinity. We also establish continuous dependence of the result with respect to the sequence of regularization parameters. Numerical experiments in two and three space dimensions are presented that illustrate the scale-space behavior of regularization methods.

Keywords: regularization methods, diffusion filtering, scale-spaces, Lyapunov functionals

1 Introduction

Decades after Iijima's pioneering axiomatic work in the sixties [27, 52], scale-spaces have become widely-used tools in image processing and computer vision [23, 33]. Alvarez *et al.* [3] have shown that imposing a reasonable set of architectural, invariance and simplification properties automatically leads to scale-spaces that can be described in terms of partial differential equations.

Partial differential equations may be classified into three main types: parabolic equations behaving in a diffusion-like manner, hyperbolic processes with wave-like character, and elliptic PDEs that can be related to variational problems. For more details on PDEs we refer to Colton [14].

Examples for PDE-based scale-spaces incude parabolic PDEs such as linear and nonlinear diffusion scale-spaces [27, 37, 49], but also curvature scale-spaces like mean-curvature motion [4, 30] and affine morphological scale space [3, 40]. Hyperbolic PDEs with scale-space properties are given by the dilation and erosion equations arising from continuous-scale morphology [3, 6, 7, 8, 29].

Recently, Scherzer and Weickert [42] showed that a large class of regularization methods reveals the same scale-space properties as diffusion filtering, if one regards the regularization parameter of these elliptic PDEs as a scale parameter. This class includes the linear Tikhonov regularization as well as many nonlinear regularization methods that can be regarded as modified total variation (TV) denoising strategies.

The goal of the present paper is to extend this theory to regularization methods that are applied iteratively and with different regularization parameters. This framework is important since one can show that iterating regularization methods improves the restoration results in some cases [42], and varying the regularization parameter can be useful for accelerating the filtering procedure. We also extend our work by clarifying relations between diffusion filtering with nonmonotone fluxes or anisotropic diffusion filtering with a diffusion tensor on one hand, and iterated convex regularization methods on the other hand.

Our paper is organized as follows: In Section 2 and 3 we survey scale-space properties of diffusion filtering and noniterated regularization, respectively. Afterwards, this framework is extended to iterated nonstationary regularization in Section 4, where detailed proofs are presented. Section 5 gives an interpretation of diffusion filtering with nonmonotone fluxes or diffusion tensors in terms of iterated convex regularization methods. In Section 6 our theory is illustrated by experiments with 2D MR images and 3D ultrasound data.

Related work. Often there have been fruitful interactions between linear scale-space techniques and regularization methods. Torre and Poggio [47] emphasized that differentiation is ill-posed in the sense of Hadamard, and applying suitable regularization strategies approximates linear diffusion filtering or – equivalently – Gaussian convolution. Much of the linear scale-space literature is based on the regularization properties of convolutions with Gaussians. In particular, differential geometric image analysis is performed by replacing derivatives by Gaussian-smoothed derivatives; see e.g. [18, 31, 44] and the references therein. In a very interesting work, Nielsen *et al.* [32] derived linear diffusion filtering axiomatically from Tikhonov regularization, where the stabilizer consists of a sum of squared derivatives up to infinite order.

Nonlinear diffusion filtering can be regarded both as a restoration method and a scale-space technique [37, 49]. When considering the restoration properties, natural relations between *biased* diffusion and regularization theory exist via the Euler equation for the regularization functional. This Euler equation can be regarded as the steady-state of a suitable nonlinear diffusion process with a bias term [13, 36, 43]. A popular specific energy functional arises from unconstrained total variation denoising [1, 10, 11]. Constrained total variation also leads to a nonlinear diffusion process with a bias term using a time-dependent Lagrange multiplier [39].

Strong and Chan [46] proposed to regard the regularization parameter of total variation denoising as a scale parameter. The present paper extends and completes our recent work on scale-space properties for noniterated regularization [42]. Following [26, 41, 46] we interpret the regularization parameter as a diffusion time by considering regularization as time-discrete diffusion filtering with a single implicit time step. Numerical implications of this relation are discussed in [51], and a shorter preliminary version of the present manuscript has been presented at the *Second International Conference on Scale-Space Theories in Computer Vision* [38].

2 Diffusion Filtering

In this section we review essential scale-space properties of nonlinear diffusion filtering. The presented results can also be extended to a broader class of methods including regularized filters with nonmonotone flux functions and anisotropic filters with a diffusion tensor. More details and proofs can be found in [49].

We consider a *diffusion process* of the form[1]

$$\begin{cases} \partial_t u(x,t) &= \nabla . \left(g(|\nabla u|^2)\nabla u\right)(x,t) & \text{on } \Omega \times [0,\infty) \\ \partial_n u(x,t) &= 0 & \text{on } \Gamma \times [0,\infty) \\ u(x,0) &= f(x) & \text{on } \Omega\,. \end{cases} \tag{1}$$

The image domain $\Omega \subseteq \mathbb{R}^d$ is assumed to be bounded with piecewise Lipschitzian boundary Γ with unit normal vector n, and $f \in L^\infty(\Omega)$ is a degraded original image with $a := \operatorname{ess\,inf}_\Omega f$ and $b := \operatorname{ess\,sup}_\Omega f$.

The diffusivity g satisfies the following properties:

1. Smoothness: $g \in C^\infty([0,\infty))$
2. The flux $g(s^2)s$ is monotonically increasing in s.
3. Positivity: $g(s) > 0$ for all $s \geq 0$.

Under these assumptions there exists a unique solution $u(x,t)$ of (1), such that $\|u(t)\|_{L^2(\Omega)}$ is continuous for $t \geq 0$. Here and in the following we use the abbreviation $u(t)$ for $u(.,t)$. It should be noted that this continuity property is necessary for relating structures over scales and for retrieving the original image for $t \to 0$. It is one of the fundamental architectural ingredients of scale-space theory. Furthermore, it is possible to show that $u(x,t) \in C^\infty(\bar{\Omega} \times (0,\infty))$.

[1]We denote by (a,b) the open interval with startpoint a and endpoint b, $(a,b]$ denotes the interval which is open at a and closed at b, and $[a,b]$ denotes the closed interval.

Diffusion processes with reflecting boundary conditions preserve the average grey level:

$$\frac{1}{|\Omega|}\int_{\Omega} u(x,t)dx = Mf \qquad \text{for all } t > 0\,,$$

with

$$Mf := \frac{1}{|\Omega|}\int_{\Omega} f(x)\,dx\;.$$

A constant average grey level is essential for scale-space segmentation algorithms such as the *hyperstack* [34]. It is also a desirable quality in medical imaging where grey values measure physical quantities of the depicted object, for instance proton densities in MR images.

The unique solution of (1) fulfills the extremum principle

$$a \le u(x,t) \le b \text{ on } \Omega \times (0,T]. \tag{2}$$

The extremum principle is an equivalent formulation of Koenderink's causality requirement [25]. Together with the continuity it ensures that level sets can be traced back in scale.

Another important simplification property can be expressed in terms of Lyapunov functionals. For all $r \in C^2[a,b]$ with $r'' \ge 0$ on $[a,b]$, the function

$$V(t) := \phi(u(t)) := \int_{\Omega} r(u(x,t))\,dx \tag{3}$$

is a Lyapunov functional:

1. It is bounded from below: $\phi(u(t)) \ge \phi(Mf)$ for all $t \ge 0$.
2. It is smoothly decreasing:
   (a) $V \in C[0,\infty) \cap C^1(0,\infty)$
   (b) $V'(t) \le 0$ for all $t > 0$.

Lyapunov functionals show that diffusion filters create simplifying transformations: the special choices $r(s) := |s|^p$, $r(s) := (s - Mf)^{2n}$ and $r(s) = s\ln(s)$, respectively, imply that all L^p norms with $p \ge 2$ are decreasing, all even central moments are decreasing, and the entropy $S[u(t)] := -\int_{\Omega} u(x,t)\ln u(x,t)\,dx$, a measure of uncertainty and missing information, is increasing with respect to t. Lyapunov functionals have been used for scale-selection and texture analysis [45], for the synchronization of different diffusion scale-spaces [34], and for the automatic determination of stopping times [53]. Moreover, they allow to prove that the filtered image converges to a constant image as t tends to ∞: $\lim_{t\to\infty} \|u(t) - Mf\|_{L^p(\Omega)} = 0$ for $p \in [1,\infty)$. For $d = 1$ we have even uniform convergence.

3 Regularization

An interesting relation between nonlinear diffusion filtering and regularization methods becomes evident when considering an implicit time discretization [26, 41, 46]. The first step of an implicit scheme with step-size h in t–direction reads as follows.

$$\begin{cases} \frac{u(x,h)-u(x,0)}{h} &= \nabla.\left(g(|\nabla u|^2)\nabla u\right)(x,h) \\ \partial_n u(x,h) &= 0 \\ u(x,0) &= f(x)\,. \end{cases} \tag{4}$$

In the following we assume the existence of a differentiable function $\hat{g}$ on $[0,\infty)$ which satisfies $\hat{g}' = g$. Then the minimizer of the functional

$$T(u) := \|u-f\|^2_{L^2(\Omega)} + h\int_\Omega \hat{g}(|\nabla u|^2)\,dx \tag{5}$$

satisfies (4). This can be seen by calculating the formal Gateaux derivative of T in direction v, i.e.

$$(T'(u),v) = \lim_{t\to 0^+}\frac{T(u+tv)-T(u)}{t} = \int_\Omega 2(u-f)v\,dx + h\int_\Omega 2g(|\nabla u|^2)\nabla u\nabla v\,dx.$$

We remark that for the numerical solution of parabolic differential equations several numerical schemes rely on implicit time discretizations, since they are unconditionally stable, i.e., for any choice of the time discretization the solution is stable with respect to data perturbations. In our context the unconditional stability of time implicit numerical schemes for solving the parabolic differential equation could as well be derived from regularization theory.

Since a minimizer of (5) satisfies $(T'(u),v) = 0$ for all v, we can conclude that the minimizer satisfies the differential equation (4). If the functional T is convex, then a minimizer of T is uniquely characterized by the solution of equation (4).

$T(u)$ is a typical regularization functional consisting of the approximation functional $\|u-f\|^2_{L^2(\Omega)}$ and the stabilizing functional $\int_\Omega \hat{g}(|\nabla u|^2)\,dx$. The weight h is called *regularization parameter.* An extensive discussion of regularization methods can be found in [16].

Now we sketch our scale-space theory for a broad class of regularization methods. For proofs and full details we refer to [42]. Let $\hat{g}$ satisfy the following properties.

I. $\hat{g}(.)$ is continuous for any compact $K \subseteq [0,\infty)$.

II. $\hat{g}(0) = \min\{\hat{g}(x) : x \in [0,\infty)\} \geq 0$.

III. $\hat{g}(|.|^2)$ is convex from $\mathbb{R}^d$ to $\mathbb{R}$.

IV. There exists a constant $c > 0$ such that $\hat{g}(s) - \hat{g}(0) \geq cs$.

V. $\hat{g}$ is monotone in $[0, \infty)$.

These assumptions guarantee existence and uniqueness of a minimizer u_h for the regularization functional (5) in the Sobolev space[2] $H^1(\Omega)$.

III. implies that $g(|.|^2).$, where $g = \hat{g}'$ is monotone, i.e., for all $s_1, s_2 \in \mathbb{R}^d$

$$\langle g(|s_1|^2)s_1 - g(|s_2|^2)s_2, s_1 - s_2 \rangle \geq 0 \; .$$

Assumptions I.-V. are satisfied for the following regularization techniques:

1. Tikhonov regularization:
$$\hat{g}(|s|^2) = |s|^2 \; .$$

2. The modified total variation regularization of Ito and Kunisch [28]:
$$\hat{g}(|s|^2) = \sqrt{|s|^2} + \alpha |s|^2, \text{ with } \alpha > 0 \; .$$

3. The modified total variation regularization of Nashed and Scherzer [35]:
$$\hat{g}(|s|^2) = \sqrt{|s|^2 + \beta^2} + \alpha |s|^2 \; .$$

4. The regularization of Geman and Yang [22] and Chambolle and Lions [10]:
$$\hat{g}(|s|^2) = \begin{cases} \frac{1}{2\epsilon}|s|^2 & |s| \leq \epsilon \\ |s| - \frac{\epsilon}{2} & \epsilon \leq |s| \leq \frac{1}{\epsilon} \\ \frac{\epsilon}{2}|s|^2 + \frac{1}{2}\left(\frac{1}{\epsilon} - \epsilon\right) & |s| > \frac{1}{\epsilon} \; . \end{cases}$$

5. Schnörr's [43] convex nonquadratic regularization:
$$\hat{g}(|s|^2) = \begin{cases} \lambda_h^2 |s|^2 & |s| \leq c_\rho \\ \lambda_l^2 |s|^2 + (\lambda_h^2 - \lambda_l^2) c_\rho (2|s| - c_\rho) & |s| > c_\rho \; . \end{cases}$$

Assumption IV. on $\hat{g}$ is violated for the total variation regularization in its original formulation by Rudin *et al.* [39]. Note that for TV-regularization $|\nabla u|$ only exists as a measure (see [17]). Therefore, we cannot set $\hat{g}(\sqrt{\cdot}) = \sqrt{\cdot}$ to obtain TV-regularization. Consequently, we cannot derive an equivalent optimality condition for a minimizer of (5). In this case our mathematical framework cannot guarantee existence of a minimizer of (5) in $H^1(\Omega)$, and in turn we have no existence theory for the partial differential equation (4). However, this does not mean that it is impossible to establish similar results by using other mathematical tools in the proofs; see e.g. the recent existence and uniqueness results by Andreu *et al.* [5].

[2] A function f belongs to the Sobolev space $H^m(\Omega)$ if f and all its derivatives up to order m belong to $L^2(\Omega)$. For more details on Sobolev spaces we refer to Adams [2].

The functional $\|u_h\|_{L^2(\Omega)}$ can also be shown to be continuous in $h \geq 0$. Regarding spatial smoothness, the solution belongs to $H^2(\Omega)$. This result is weaker than for the diffusion case where we have C^∞ results.

In analogy to diffusion filtering, the average grey level invariance

$$\int_\Omega u_h \, dx = \int_\Omega f \, dx \quad \text{for all } h \geq 0$$

and the extremum principle

$$a \leq u_h \leq b \quad \text{for all } h \geq 0$$

can be established.

Moreover, Lyapunov functionals for regularization methods can be constructed in a similar way. For all $r \in C^2[a,b]$ with $r'' \geq 0$, the function

$$V(h) := \phi(u_h) := \int_\Omega r(u_h(x)) \, dx \tag{6}$$

is a Lyapunov functional:

1. It is bounded from below: $\phi(u_h) \geq \phi(Mf)$ for all $h \geq 0$.
2. It is continuous and decreasing with respect to the original image:
   (a) $V \in C[0,\infty)$,
   (b) $DV(h) := \int_\Omega r'(u_h)(u_h - u_0) \leq 0$, for all $h \geq 0$.
   (c) $V(h) - V(0) \leq 0$ for all $h \geq 0$.

Here, a difference between Lyapunov functionals for diffusion processes and regularization methods becomes evident. For Lyapunov functionals in diffusion processes we have $V'(t) \leq 0$, and in regularization processes we have $DV(h) \leq 0$. $DV(h)$ is obtained from $V'(t)$ by making a time discrete ansatz at time 0. We note that this is exactly the way we compared diffusion filtering and regularization techniques. It is therefore natural that the role of the time derivative in diffusion filtering is replaced by the time discrete approximation around 0.

Again, these Lyapunov functionals allow to prove convergence of the filtered images to a constant image as $h \to \infty$. For $d = 3$, however, the convergence result is slightly weaker than in the diffusion case.

$d = 1:$ u_h converges uniformly to Mf for $h \to \infty$

$d = 2:$ $\lim_{h\to\infty} \|u_h - Mf\|_{L^p(\Omega)} = 0$ for any $1 \leq p < \infty$

$d = 3:$ $\lim_{h\to\infty} \|u_h - Mf\|_{L^p(\Omega)} = 0$ for any $1 \leq p \leq 6$

4 Iterated Regularization

Regularization can be applied iteratively where the regularized solution of the previous step serves as initial image for the next iteration. For small regularization parameters, iterated regularization becomes therefore a good approximation to a nonlinear diffusion filter.

Let us consider an iterative regularization process with a sequence of positive regularization parameters $\mathcal{H} := (h_k)_{k\in\mathbb{N}}$. With $\mathcal{T} := (t_k)_{k\in\mathbb{N}}$ we denote the sequence of corresponding "diffusion times", i.e., $t_k := \sum_{i=1}^{k} h_i$. Note that $t_k - t_{k-1} = h_k$.

The n-th iteration of the nonstationary iterative regularization method reads as follows:

$$\begin{aligned}
\frac{u^{\mathcal{H}}(x,t) - u^{\mathcal{H}}(x,t_{n-1})}{t - t_{n-1}} &= \nabla.\left(g(|\nabla u^{\mathcal{H}}|^2)\nabla u^{\mathcal{H}}\right)(x,t) && t \in (t_{n-1}, t_n],\ x \in \Omega \\
\partial_n u^{\mathcal{H}}(x,t) &= 0 && x \in \Gamma \\
u^{\mathcal{H}}(x,0) &= f(x) && x \in \Omega
\end{aligned} \tag{7}$$

where now $t - t_{n-1}$ serves as the regularization parameter in the interval $(t_{n-1}, t_n]$. The superscript $\mathcal{H}$ at u refers to the fact that u is dependent on the discretization time in t direction. In the following we establish a scale-space theory for nonstationary iterated regularization. The terminology "iterative" refers to the fact that Tikhonov regularization is implemented iteratively. The terminology "nonstationary" refers to the fact that the parameters h_k may vary during the iterative process.

A minimizer $u \in H^1(\Omega)$ of the functional

$$T_t^{\mathcal{H}}(u) := \|u - u^{\mathcal{H}}(t_{n-1})\|^2_{L^2(\Omega)} + (t - t_{n-1}) \int_\Omega \hat{g}(|\nabla u|^2)\, dx, \tag{8}$$

satisfies (7) at time $t \in (t_{n-1}, t_n]$. If the functional $T_t^{\mathcal{H}}$ is strictly convex, then a minimizer of $T_t^{\mathcal{H}}$ is uniquely characterized by the solution of equation (7). Under these assumptions the minimizer of (8) exists and is unique in $H^1(\Omega)$ (cf. [42]).

Moreover, the spatial smoothness increases in each iteration step: a more detailed analysis using techniques from [54] shows that after n iterations the solution belongs to the Sobolev space $H^{2n}(\Omega)$ for fixed $t \in (t_{n-1}, t_n]$ (provided the diffusivity g is sufficiently smooth). This suggests that, if one uses the regularized solution for calculating derivatives of order $2n$, one should perform at least n iterations.

As for noniterated regularization, the average grey level invariance, temporal continuity in the L^2-norm and a maximum principle hold, if the function $\hat{g}$ satisfies I. – V. This can be seen, by noting that in the interval $(t_{n-1}, t_n]$ iterated regularization is noniterated regularization with initial data $u^{\mathcal{H}}(t_{n-1})$ and regularization parameter $t - t_{n-1}$. The results in [42] imply that in each interval $[t_{n-1}, t_n]$ the average grey level invariance holds, the function $u^{\mathcal{H}}(t)$ is bounded by the maximal and minimal values of $u^{\mathcal{H}}(t_{n-1})$, and the function is continuous with respect to t. The rest of the assertion follows by an inductive argument.

Using these properties we are able to establish a Ljapunov theory for iterated regularization.

Theorem 1 (Lyapunov functionals for nonstationary regularization methods) *Let $\Omega \subseteq \mathbb{R}^d$, $d = 1, 2, 3$ with smooth boundary. Let $f \in L^\infty(\Omega)$ with essential minimal value a and essential maximal value b. Moreover, let $\mathcal{H}$ be a sequence of positive numbers h_k, satisfying $\lim_{k\to\infty} h_k = \infty$. Then the following properties hold:*

(a) For all $r \in C^2[a,b]$ with $r'' \geq 0$, the function

$$V^{\mathcal{H}}(t) := \phi(u^{\mathcal{H}}(t)) := \int_\Omega r(u^{\mathcal{H}}(x,t))\,dx \tag{9}$$

is a Lyapunov functional for iterative regularization:

1. *It is bounded from below:* $\phi(u^{\mathcal{H}}(t)) \geq \phi(Mf)$ *for all* $t \geq 0$,
2. *It is continuous and decreasing with respect to the original image:*
   (a) $V^{\mathcal{H}} \in C[0,\infty)$,
   (b) $DV^{\mathcal{H}}(t) := \int_\Omega r'(u^{\mathcal{H}}(x,t)\left(u^{\mathcal{H}}(x,t) - u^{\mathcal{H}}(x,t_{n-1})\right)dx \leq 0$, *for all* $t \in (t_{n-1}, t_n]$,
   (c) $V^{\mathcal{H}}(t) - V^{\mathcal{H}}(t_{n-1}) \leq 0$ *for all* $t \in (t_{n-1}, t_n]$.

Moreover, if $r'' > 0$ on $[a,b]$, then $V^{\mathcal{H}}(t)$ is a strict Lyapunov functional:

3. $\phi(u^{\mathcal{H}}(t)) = \phi(Mf)$ *for all* $t \in [0,\infty)$ *if and only if* $u^{\mathcal{H}}(t) = Mf$ *on* $\overline{\Omega}$ *for* $t > 0$, *and* $u^{\mathcal{H}}(.,0) = Mf$ *almost everywhere on* $\overline{\Omega}$.
4. *If $t > 0$, then $DV^{\mathcal{H}}(t) = 0$ if and only if $u^{\mathcal{H}}(t) = Mf$ on $\overline{\Omega}$.*
5. $V^{\mathcal{H}}(T) = V^{\mathcal{H}}(0)$ *for* $T > 0$ *if and only if* $f = Mf$ *almost everywhere on* Ω, *and* $u^{\mathcal{H}}(t) = Mf$ *on* $\overline{\Omega} \times (0,T]$.

(b) (Convergence)

d=1: $u^{\mathcal{H}}(t_n)$ *converges uniformly to* Mf *for* $n \to \infty$

d=2: $\lim_{n\to\infty} \|u^{\mathcal{H}}(t_n) - Mf\|_{L^p(\Omega)} = 0$ *for any* $1 \leq p < \infty$

d=3: $\lim_{n\to\infty} \|u^{\mathcal{H}}(t_n) - Mf\|_{L^p(\Omega)} = 0$ *for any* $1 \leq p \leq 6$

Proof: Using the general result in [42] it follows that the assertions claimed in part (a) hold on the subintervals $(t_{n-1}, t_n]$. Moreover, from the results in [42] it follows that continuity of $u(.,t)$, with respect to t, also holds on $[t_{n-1}, t_n]$. By induction with respect to n the assertions of part (a) follow.

We turn to a verification of the assertions of part (b): since $u^{\mathcal{H}}(t_n)$ satisfies the first order optimality condition for a minimum of the functional $T^{\mathcal{H}}_{t_n}$ we get

$$\begin{aligned}&\left\langle u^{\mathcal{H}}(t_n), v\right\rangle_{L^2(\Omega)} + (t_n - t_{n-1})\left\langle g(|\nabla u^{\mathcal{H}}(t_n)|^2)\nabla u^{\mathcal{H}}(t_n), \nabla v\right\rangle_{L^2(\Omega)}\\ =\;&\left\langle u^{\mathcal{H}}(t_{n-1}), v\right\rangle_{L^2(\Omega)}\end{aligned}$$

and taking $v = u^{\mathcal{H}}(t_n)$ shows that

$$\begin{aligned}&\|u^{\mathcal{H}}(t_n)\|^2_{L^2(\Omega)} + (t_n - t_{n-1}) \left\langle g(|\nabla u^{\mathcal{H}}(t_n)|^2)\nabla u^{\mathcal{H}}(t_n), u^{\mathcal{H}}(t_n)\right\rangle_{L^2(\Omega)} \\ &= \left\langle u^{\mathcal{H}}(t_{n-1}), u^{\mathcal{H}}(t_n)\right\rangle_{L^2(\Omega)} .\end{aligned}$$

Since $\left\langle g(|\nabla u^{\mathcal{H}}(t_n)|^2)\nabla u^{\mathcal{H}}(t_n), \nabla u^{\mathcal{H}}(t_n)\right\rangle_{L^2(\Omega)}$ is positive, which follows from the convexity of $\hat{g}$, we find that $u^{\mathcal{H}}(t_n)$ is uniformly bounded in $L^2(\Omega)$.

Since $u^{\mathcal{H}}(t_n)$ minimizes the functional $T^{\mathcal{H}}_{t_n}$ we immediately get that

$$\begin{aligned}\|u^{\mathcal{H}}(t_n) - u^{\mathcal{H}}(t_{n-1})\|^2_{L^2(\Omega)} + (t_n - t_{n-1}) \int_\Omega \hat{g}(|\nabla u^{\mathcal{H}}(x,t_n)|^2)\,dx& \\ \leq (t_n - t_{n-1}) \int_\Omega \hat{g}(|\nabla u^{\mathcal{H}}(x,t_{n-1})|^2)\,dx .&\end{aligned} \tag{10}$$

Thus, the sequence $\int_\Omega \hat{g}(|\nabla u^{\mathcal{H}}(t_n)|^2)\,dx$ is monotonically decreasing in n.
Now we show that

$$\int_\Omega \hat{g}(|\nabla u^{\mathcal{H}}(x,t_n)|^2)\,dx \xrightarrow{n\to\infty} \int_\Omega \hat{g}(0)\,dx . \tag{11}$$

Since $u^{\mathcal{H}}(t_n)$ is the minimizing element of (8) (for $t = t_n$), we have

$$(t_n - t_{n-1}) \int_\Omega \hat{g}(|\nabla u^{\mathcal{H}}(x,t_n)|^2)\,dx \leq \|Mf - u^{\mathcal{H}}(t_{n-1})\|^2_{L^2(\Omega)} + (t_n - t_{n-1}) \int_\Omega \hat{g}(0)\,dx.$$

Dividing the inequality by $h_n = t_n - t_{n-1}$ and noting that $h_n \to \infty$ shows that

$$\lim_{n\to\infty} \int_\Omega \hat{g}(|\nabla u^{\mathcal{H}}(t_n)|^2)\,dx \leq \int_\Omega \hat{g}(0)\,dx .$$

Together with II. we get

$$\lim_{n\to\infty} \int_\Omega \hat{g}(|\nabla u^{\mathcal{H}}(t_n)|^2)\,dx = \int_\Omega \hat{g}(0)\,dx . \tag{12}$$

Since $u^{\mathcal{H}}(t_n)$ is uniformly bounded in $L^2(\Omega)$, it has a weakly convergent subsequence. Using IV. it follows that the weak limit is a constant function. Since iterative regularization is grey level invariant, we find that the limit is the constant function Mf. Moreover, from (12) it follows that the H^1-seminorm is strongly convergent to 0. Thus, the sequence $\{u^{\mathcal{H}}(t_n)\}_{n\in\mathbb{N}}$ itself is strongly convergent, i.e.

$$\|u^{\mathcal{H}}(t_n) - Mf\|_{H^1(\Omega)} \to 0 .$$

By virtue of the Sobolev embedding theorem [2] it follows in particular that we obtain the following convergence results for $n \to \infty$.

d=1: $u^{\mathcal{H}}(t_n)$ converges uniformly to Mf

d=2: $\|u^{\mathcal{H}}(t_n) - Mf\|^2_{L^p(\Omega)} \to 0$ for any $1 \leq p < \infty$

d=3: $\|u^{\mathcal{H}}(t_n) - Mf\|^2_{L^p(\Omega)} \to 0$ for any $1 \leq p \leq 6$

This concludes the proof.

The assumption that the sequence of regularization parameters h_k tends to infinity does not restrict practical applications. It actually suggest that in numerical simulations a monotonically increasing step size in time is very appropriate. Such an adaptation strategy would use small time steps in the beginning when much is happening, and afterwards, when the diffusion process slows down, the time step size becomes larger.

The previous result holds independently of the sequence $\mathcal{H} = \{h_k\}_{k\in\mathbb{N}}$. For numerical realizations of nonstationary iterated regularization it is important to verify continuous dependence of $u^{\mathcal{H}}(t)$ with respect $\mathcal{H}$. In order to prove this result we first show that the functional $u^{\mathcal{H}}(t)$ is Lipschitz continuous. Let $\mathcal{H}$ be a sequence of positive regularization parameters and let $\mathcal{T}$ be the according sequence of diffusion times. Moreover, let τ_1 and τ_2 be two positive numbers in the interval $(t_{k-1}, t_k]$. Then

$$\langle u(\tau_1) - u(t_{k-1}), v\rangle_{L^2(\Omega)} + (\tau_1 - t_{k-1}) \left\langle g(|u(\tau_1)|^2)\nabla u(\tau_1), \nabla v\right\rangle_{L^2(\Omega)} = 0$$

and

$$\langle u(\tau_2) - u(t_{k-1}), v\rangle_{L^2(\Omega)} + (\tau_2 - t_{k-1}) \left\langle g(|u(\tau_2)|^2)\nabla u(\tau_2), \nabla v\right\rangle_{L^2(\Omega)} = 0\ .$$

Taking the difference of both equations gives

$$\begin{aligned} &\langle u(\tau_1) - u(\tau_2), v\rangle_{L^2(\Omega)} \\ + \quad &(\tau_1 - t_{k-1}) \langle g(|u(\tau_1)|^2)\nabla u(\tau_1) - g(|u(\tau_2)|^2)\nabla u(\tau_2), \nabla v\rangle_{L^2(\Omega)} \\ - \quad &(\tau_2 - \tau_1) \langle g(|\nabla u(\tau_2)|^2)\nabla u(\tau_2), \nabla v\rangle_{L^2(\Omega)} = 0\ . \end{aligned}$$

Taking $v = u(\tau_1) - u(\tau_2)$ and using the monotonicity of $g(|.|^2)$. gives:

$$\|u(\tau_1) - u(\tau_2)\|_{L^2(\Omega)} \leq \frac{|\tau_2 - \tau_1|}{\tau_2 - t_{k-1}} \|u(\tau_2) - u(t_{k-1})\|_{L^2(\Omega)}\ . \tag{13}$$

In particular for τ_1 and τ_2 both greater than $t_{k-1} + \epsilon$, with $\epsilon > 0$,

$$\|u(\tau_1) - u(\tau_2)\|_{L^2(\Omega)} \leq C_\epsilon |\tau_2 - \tau_1|\ ,$$

where C_ϵ is independent of the particular choice of $\mathcal{H}$ as long as t_{k-1} is an element of $\mathcal{H}$.

With this Lipschitz continuity we are able to prove continuous dependence of $u^{\mathcal{H}}(t)$ on $\mathcal{H}$.

Lemma 1 *Let $\mathcal{H}^n$, $n \in \mathbb{N}$ and $\mathcal{H}$ be sequences of positive regularization parameters, where each sequence converges to infinity. Let $\mathcal{T}^n$, $n \in \mathbb{N}$ and $\mathcal{T}$ be the according sequences of diffusion times. Let*

$$t_k^n \to t_k \text{ for } n \to \infty, \text{ uniformly in } k\ .$$

Then

$$\|u^{\mathcal{H}}(t) - u^{\tilde{\mathcal{H}}}(t)\|_{L^2(\Omega)} \to 0 \text{ for any } t \in [0, \infty)\ .$$

Proof: By means of the assumption on uniform convergence of t_k^n to t_k we get that, for any $t \in (t_{k-1}, t_k)$, there exists a sufficiently large index $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ also $t \in (t_{k-1}^n, t_k^n)$. Using that $u^{\mathcal{H}^n}(t)$ minimizes the functional $T_t^{\mathcal{H}^n}$ and $u^{\mathcal{H}}(t)$ minimizes the functional $T_t^{\mathcal{H}}$ we get

$$\begin{aligned} 0 &= \langle u^{\mathcal{H}^n}(t) - u^{\mathcal{H}^n}(t_{k-1}^n), v\rangle_{L^2(\Omega)} \\ &+ (t - t_{k-1}^n)\langle g(|\nabla u^{\mathcal{H}^n}(t)|^2)\nabla u^{\mathcal{H}^n}(t), \nabla v\rangle_{L^2(\Omega)} \end{aligned}$$

and

$$\begin{aligned} 0 &= \langle u^{\mathcal{H}}(t) - u^{\mathcal{H}}(t_{k-1}), v\rangle_{L^2(\Omega)} \\ &+ (t - t_{k-1})\langle g(|\nabla u^{\mathcal{H}}(t)|^2)\nabla u^{\mathcal{H}}(t), \nabla v\rangle_{L^2(\Omega)} . \end{aligned} \tag{14}$$

Choosing $v := u^{\mathcal{H}}(t) - u^{\mathcal{H}^n}(t)$ and subtracting these two equations gives

$$\begin{aligned} &\|u^{\mathcal{H}^n}(t) - u^{\mathcal{H}}(t)\|_{L^2(\Omega)}^2 + \langle u^{\mathcal{H}^n}(t_{k-1}^n) - u^{\mathcal{H}}(t_{k-1}), u^{\mathcal{H}}(t) - u^{\mathcal{H}^n}(t)\rangle_{L^2(\Omega)} \\ &+ (t - t_{k-1}^n)\, \langle g(|\nabla u^{\mathcal{H}}(t)|^2)\nabla u^{\mathcal{H}}(t) - g(|\nabla u^{\mathcal{H}^n}(t)|^2)\nabla u^{\mathcal{H}^n}(t), \\ &\quad \nabla(u^{\mathcal{H}}(t) - u^{\mathcal{H}^n}(t))\rangle_{L^2(\Omega)} \\ &- (t_{k-1} - t_{k-1}^n)\, \langle g(|\nabla u^{\mathcal{H}}(t)|^2)\nabla u^{\mathcal{H}}(t), \nabla(u^{\mathcal{H}}(t) - u^{\mathcal{H}^n}(t))\rangle_{L^2(\Omega)} = 0 . \end{aligned}$$

Using that $g(|.|^2)$. is monotone, it follows from (14) that

$$\begin{aligned} \|u^{\mathcal{H}}(t) - u^{\mathcal{H}^n}(t)\|_{L^2(\Omega)} &\leq \|u^{\mathcal{H}}(t_{k-1}) - u^{\mathcal{H}^n}(t_{k-1}^n)\|_{L^2(\Omega)} \\ &+ \frac{|t_{k-1} - t_{k-1}^n|}{|t - t_{k-1}|}\|u^{\mathcal{H}}(t) - u^{\mathcal{H}}(t_{k-1})\|_{L^2(\Omega)} . \end{aligned} \tag{15}$$

Finally, we apply an inductive argument with respect to k. Let

$$\|u^{\mathcal{H}}(t_{k-1}) - u^{\mathcal{H}^n}(t_{k-1}^n)\|_{L^2(\Omega)} \to 0 \text{ for } n \to \infty .$$

For $k = 0$ this is trivially satisfied, since $u^{\mathcal{H}}(.,0) = u^{\mathcal{H}^n}(.,0) = f$. Then it follows from (15) that, for any $t \in (t_{k-1}, t_k)$,

$$\|u^{\mathcal{H}}(t) - u^{\mathcal{H}^n}(t)\|_{L^2(\Omega)} \to 0 \text{ for } n \to \infty .$$

Now, let $t = t_k$.

- If $t_k^n > t_k$, then we have $t = t_k \in (t_{k-1}^n, t_k^n)$ and analogously as above one can show that (15) holds with t replaced by t_k. Repeating the above arguments we find that

$$\|u^{\mathcal{H}}(t_k) - u^{\mathcal{H}^n}(t_k)\|_{L^2(\Omega)} \to 0 \text{ for } n \to \infty .$$

- If $t_k^n \leq t_k$, then let $t_0 \leq t_k^n$ fixed. Then it follows from triangle inequality that

$$\begin{aligned} \|u^{\mathcal{H}}(t_k) - u^{\mathcal{H}^n}(t_k^n)\|_{L^2(\Omega)} \ \leq \ & \|u^{\mathcal{H}}(t_0) - u^{\mathcal{H}^n}(t_0)\|_{L^2(\Omega)} + \\ & \|u^{\mathcal{H}}(t_k) - u^{\mathcal{H}}(t_0)\|_{L^2(\Omega)} + \|u^{\mathcal{H}^n}(t_0) - u^{\mathcal{H}^n}(t_k^n)\|_{L^2(\Omega)} \,. \end{aligned}$$

  The second and third term on the right hand side of the last inequality are of the order $\max\{|t_0 - t_k|, |t_0 - t_k^n|\}$ (cf. (13)) (independent of n) - note that t_k^n converges uniformly to t_k. Since $t_k^n \to t_k$ we can choose t_0 in such a way that $\max\{|t_0 - t_k|, |t_0 - t_k^n|\}$ becomes arbitrarily small. Moreover, the frist term tends to zero, since t_0 is an interior point of (t_{k-1}, t_k). These arguments show that

$$\|u^{\mathcal{H}}(t_k) - u^{\mathcal{H}^n}(t_k^n)\|_{L^2(\Omega)} \to 0 \,.$$

Hence, the lemma is proved.

We can use this lemma to show continuous dependence of $V^{\mathcal{H}}(t)$ on $\mathcal{H}$:

Theorem 2 *Let the assumptions of Lemma 1 hold. Then for, any $t \in [0, \infty)$,*

$$|V^{\mathcal{H}}(t) - V^{\mathcal{H}^n}(t)| \to 0 \,.$$

Proof: Using the Cauchy-Schwarz inequality it follows that

$$\begin{aligned} \left|V^{\mathcal{H}}(t) - V^{\mathcal{H}^n}(t)\right| \ \leq \ & \left(\int_\Omega \left(\frac{r(u^{\mathcal{H}}(x,t)) - r(u^{\mathcal{H}^n}(x,t))}{u^{\mathcal{H}}(x,t) - u^{\mathcal{H}^n}(x,t)}\right)^2 dx\right)^{\frac{1}{2}} \cdot \\ & \cdot \left\|u^{\mathcal{H}}(t) - u^{\mathcal{H}^n}(t)\right\|_{L^2(\Omega)} \,. \end{aligned}$$

By virtue of Lemma 1 it follows that $\|u^{\mathcal{H}}(t) - u^{\mathcal{H}^n}(t)\|_{L^2(\Omega)}$ tends to zero. Since r is continuously differentiable, the first expression on the right hand side is bounded and the proof is accomplished.

5 Extensions to the Nonconvex Case and Anisotropic Filters

The previous sections analyze relations between regularization methods and diffusion filters for the case that $\hat{g}(|s|^2)$ is convex in s. This implies that the diffusive flux is monotonously increasing in s in the sense that

$$\langle g(|s|)s - g(|t|)t, s - t\rangle \ \geq \ 0 \quad \text{for all} \quad s, t \in \mathbb{R}^d.$$

In the context of diffusion filtering, however, nonmonotone fluxes leading to forward–backward diffusion processes are used frequently. While the earliest representative of this class, the Perona–Malik filter [37] is ill-posed, several well-posed forward–backward diffusion filters have been proposed afterwards; see e.g. [9, 49]. They offer the interesting property that they can enhance features like edges or flow-line structures without renouncing smoothing properties in terms of Lyapunov functionals. It is also possible to replace the scalar-valued diffusivity g by a diffusion tensor D allowing true anisotropic behavior.

These extensions are covered by the diffusion model

$$\left\{\begin{array}{rcll} \partial_t u(x,t) &=& \nabla.\left(D(J_\rho(\nabla u_\sigma))\nabla u\right)(x,t) & \text{on } \Omega \times [0,\infty) \\ \langle D\nabla u, n\rangle &=& 0 & \text{on } \Gamma \times [0,\infty) \\ u(x,0) &=& f(x) & \text{on } \Omega \end{array}\right. \tag{16}$$

where $u_\sigma := K_\sigma * u$ denotes the convolution of u with a Gaussian K_σ of standard deviation σ, and J_ρ is the so-called structure tensor [19]

$$J_\rho(\nabla u_\sigma) = K_\rho * (\nabla u_\sigma \nabla u_\sigma^T),$$

a very useful matrix for the analysis of edges, corners and coherent structures. This model formulation comprises the regularized diffusion filter of Catté *et al.* [9] as well as edge-enhancing anisotropic diffusion filtering [48] and coherence-enhancing anisotropic diffusion filtering [50].

By assuming that the diffusion tensor D is a symmetric matrix-valued C^∞ function of J_ρ that remains uniformly positive definite, one can prove that all theoretical results from Section 2 carry over [49]. Well-posedness is achieved in the nonconvex case by the Gaussian smoothing in u_σ; see also [9].

Because of the Gaussian convolutions there is no straightforward way to derive a diffusion filter of this type as a minimizer of some energy functional. It is, however, instructive to study a semi-implicit time discretization of such a filter: it approximates the diffusion tensor D at the old time level and the remainder of the divergence expression at the new level. Such a discretization gives

$$\frac{u^{\mathcal{H}}(x,t) - u^{\mathcal{H}}(x,t_{n-1})}{t - t_{n-1}} \;=\; \nabla.\left(D(J_\rho(\nabla u^{\mathcal{H}}(x,t_{n-1})))\nabla u^{\mathcal{H}}(x,t)\right). \tag{17}$$

It can be regarded as an iterative regularization scheme where

$$\begin{aligned} T_t^{\mathcal{H}}(u) \;&:=\; \|u - u^{\mathcal{H}}(t_{n-1})\|^2_{L^2(\Omega)} \\ &+\; (t - t_{n-1}) \int_\Omega (\nabla u)^T D(J_\rho(\nabla u_\sigma^{\mathcal{H}}(t_{n-1})))\nabla u \, dx \end{aligned} \tag{18}$$

is minimized. Now we are approximating a possibly nonconvex smoothing problem by a sequence of quadratic (and hence convex) regularization functionals.[3] As a consequence, the theoretical results for iterated regularization that we derived in Section 4 may also be extended to this case.

[3]This convexification by freezing the nonlinear part also relates our method to the adaptive linearization technique of Geman and Reynolds [21, 12] and the so-called Kačanov method from elasticity theory [20, 24, 51].

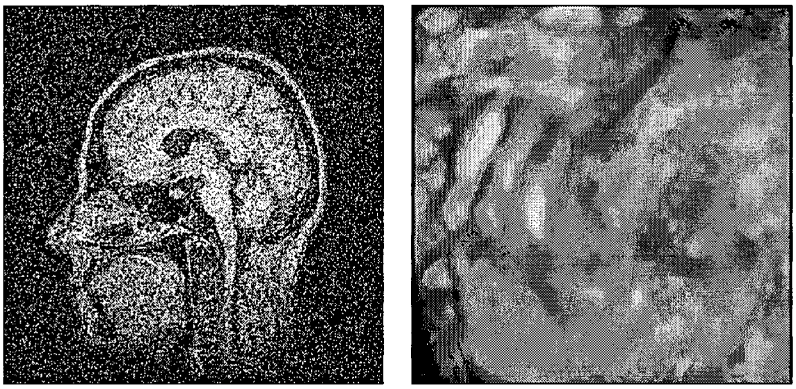


Figure 1: Test images. **(a) Left:** MR image with additive Gaussian noise ($SNR = 1$). **(b) Right:** Rendering of a three-dimensional ultrasound data set of a human fetus.

6 Experiments

The numerical experiments are performed using the software package DIFFPACK from the University of Oslo / Numerical Objects [15]. We have implemented the diffusion equation with $\hat{g}(|\nabla u|^2) = \sqrt{|\nabla u|^2 + \beta^2} + \alpha|\nabla u|^2$ which is a modified total variation regularization. For this diffusion filtering our theoretical results are applicable. The term $\alpha|\nabla u|^2$ is only of theoretical interest; in numerical realizations, the discretized version of the gradient is bounded, and there is no visible difference between using very small values of α (in which the theoretical results are applicable) and $\alpha = 0$ (where our theoretical results do not hold).

Our experiments were carried out for different sequences of time-steps and various smoothing parameters β. The influence of the parameter settings is as follows.

The impact of β on the numerical reconstruction is hardly viewable in the range from $\beta = 10^{-2}$ to 10^{-4}. Even the convergence rate is, although slower for smaller β, hardly affected.

For small values of regularization parameters h (up to approximately 5.0), there is no visible difference between iterated and noniterated regularization. The effect can only be seen for larger values of h. This is illustrated in Figure 2. It shows the result of noniterated and iterated regularization applied to the 2D MR image from Figure 1(a). The results are depicted at times $t = 10$, 30, and 100, respectively. For noniterated regularization this is achieved in one step, and for iterated regularization the regularization parameter $h = 1$ was chosen and 10, 30, or 100 iterations were performed. We observe that differences between the two methods are very small. They only become evident when subtracting one image from the other. This also indicates that even the semi-group property of regularization methods is well approximated in practice. It should be noted that the semi-group property is an ideal continuous concept which can only be approximated in time-discrete algorithms for partial differential equations.

As can be seen from the previous sections, the scale-space framework for noniterated and iterated regularization methods carries over to higher space dimensions. In the next figure we present results from a three-dimensional ultrasound data set of a fetus with $80 \times 80 \times 80$ voxels. Also in this case the differences between noniterated and iterated regularization are very small and iterated regularization appears to give slightly smoother results. This is in complete accordance with the theory in Section 4.

7 Conclusions

The novelty of our paper consists of establishing sequences of parameter dependent elliptic boundary value problems, namely nonstationary iterated regularization methods, as scale-space techniques. They satisfy the same scale-space properties as nonlinear diffusion filtering. The key ingredient for understanding this relation is the interpretation of iterated regularization methods as time-implicit or time-semi-implicit approximations to diffusion processes. In this sense, the scale-space theory of regularization methods is also a novel semi-discrete theory to diffusion filtering. This time-discrete framework completes the theory of diffusion scale-spaces where up to now only results for the continuous, the space-discrete and the fully discrete setting have been formulated [49].

The synthesis of regularization techniques and diffusion methods may lead to a deeper understanding of both fields, and it is likely that many more results can be transferred from one of these areas to the other. It would e.g. be interesting to study how results for optimal parameter selection in regularization methods can be used for diffusion filtering, or to further investigate the use of the iterated anisotropic functional (18) in the context of regularization theory. It is also promising to analyze and juxtapose efficient numerical techniques developed in both frameworks. First steps in this direction are reported in [51].

Acknowledgement

The work of E.R. and O.S. is supported by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung, SFB 1310, and J.W. received financial support from the EU-TMR project VIRGO. The authors thank Firma Kretztechnik AG, Austria, for the allowance to publish the three-dimensional ultrasound data.

References

[1] ACAR, R.; VOGEL, C.R.: *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems, 10, 1217–1229, 1994

[2] ADAMS, R.A.: *Sobolev Spaces*, Academic Press, New York, San Francisco, London, 1975

[3] ALVAREZ, L.; GUICHARD, F.; LIONS, P.L.; MOREL, J.M., *Axioms and fundamental equations of image processing*, Arch. Rat. Mech. Anal., 16, 200–257, 1993

[4] ALVAREZ, L.; LIONS, P.L.; MOREL, J.M.: *Image selective smoothing and edge detection by nonlinear diffusion. II*, SIAM J. Numer. Anal., 29, 845–866, 1992

[5] ANDREU, F.; BALLESTER, C.; CASELLES, V.; MAZÓN, J.M.: *Minimizing total variation flow*, Technical Report, Dept. of Análisis Matemático, University of Valencia, 46100 Burjassot (Valencia), Spain, 1998. Submitted to Journal of Integral and Partial Differential Equations

[6] AREHART, A.B.; VINCENT, L., KIMIA, B.B.: *Mathematical morphology: The Hamilton–Jacobi connection*, Proc. Fourth Int. Conf. on Computer Vision (ICCV '93, Berlin, May 11–14, 1993), IEEE Computer Society Press, Los Alamitos, 215–219, 1993

[7] VAN DEN BOOMGAARD, R.: *The morphological equivalent of the Gauss convolution*, Nieuw Archief Voor Wiskunde, Vierde Serie, Deel 10, 219–236, 1992

[8] BROCKETT, R.W.; MARAGOS, P.: *Evolution equations for continuous-scale morphological filtering*, IEEE Trans. Signal Processing, 42, 3377–3386, 1994

[9] CATTÉ, F.; LIONS, P.L.; MOREL J.M.; COLL T.: *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal., 32, 1895–1909, 1992

[10] CHAMBOLLE, A.; LIONS P.-L.: *Image recovery via total variation minimization and related problems*, Numer. Math., 76, 167 - 188, 1995

[11] CHAN, T.F.; GOLUB, G.H.; MULET, P.: *A nonlinear primal–dual method for total-variation based image restoration*, in Berger, M.O.; Deriche, R.; Herlin, I.; Jaffré, J.; Morel, J.M. (Eds.): ICAOS '96: Images, Wavelets and PDEs, Lecture Notes in Control and Information Sciences, 219, Springer, London, 241–252, 1996

[12] CHARBONNIER, P.; BLANC–FÉRAUD, L.; AUBERT, G.: *Deterministic edge-preserving regularization in computed imaging*, IEEE Trans. Image Proc., 6, 298–311, 1997

[13] CHARBONNIER, P.; BLANC–FÉRAUD, L.; AUBERT, G.; BARLAUD, M.: *Two deterministic half-quadratic regularization algorithms for computed imaging*, Proc. IEEE Int. Conf. Image Processing (Austin, Nov. 13–16, 1994), Vol. 2, IEEE Computer Society Press, Los Alamitos, 168–172, 1994

[14] COLTON, D.: *Partial Differential Equations*, Random House, New York, 1988

[15] DÆHLEN, M.; TVEITO, A.: *Numerical Methods and Software Tools in Industrial Mathematics*, Birkhäuser, Boston, 1997

[16] ENGL, H.W.; HANKE, M.; NEUBAUER, A.: *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996

[17] EVANS, L.C.; GARIEPY R.F.: *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton, 1992

[18] FLORACK, L.: *Image Structure*, Kluwer, Dordrecht, 1997

[19] FÖRSTNER, W.; GÜLCH, E.: *A fast operator for detection and precise location of distinct points, corners and centres of circular features*, Proc. ISPRS Intercommission Conf. on Fast Processing of Photogrammetric Data (Interlaken, June 2–4, 1987), 281–305, 1987

[20] FUČIK, S.; KRATOCHVIL, A.; NEČAS, J.: *Kačanov–Galerkin method*, Comment. Math. Univ. Carolinae, 14, 651–659, 1973

[21] GEMAN, D.; REYNOLDS, G.: *Constrained restoration and the recovery of discontinuities*, IEEE Trans. Pattern Anal. Mach. Intell., 14, 367–383, 1992

[22] GEMAN, D.; YANG, C.: *Nonlinear image recovery with half-quadratic regularization*, IEEE Transactions on Image Processing, 4, 932-945, 1995

[23] TER HAAR ROMENY, B.; FLORACK, L.; KOENDERINK, J.; VIERGEVER, M. (EDS.): *Scale-Space Theory in Computer Vision*, Lecture Notes in Computer Science, Vol. 1252, Springer, Berlin, 1997

[24] HEERS, J.; SCHNÖRR, C., STIEHL, H.S.: *Investigation of parallel and globally convergent iterative schemes performing nonlinear variational segmentation*, IEEE Trans. Image Proc., to appear

[25] HUMMEL, R.A.: *Representations based on zero crossings in scale space*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (Miami Beach, June 22 - 26, 1986), IEEE Computer Society Press, Washington, 204 – 209, 1986

[26] MOREL, J.F.; SOLIMINI S.: *Variational Methods in Image Segmentation*; Birkhäuser, Boston, 1995

[27] IIJIMA, T.: *Basic theory on normalization of pattern (in case of typical one-dimensional pattern)*, Bulletin of the Electrotechnical Laboratory, 26, 368–388, 1962 (in Japanese)

[28] ITO, K.; KUNISCH, K.: *An active set strategy based on the augmented Lagrangian formulation for image reconstruction*, RAIRO Math. Mod. and Num. Analysis, 33, 1–21, 1999

[29] JACKWAY, P.T.; DERICHE, M.: *Scale-space properties of the multiscale morphological dilation–erosion*, IEEE Trans. Pattern Anal. Mach. Intell., 18, 38–51, 1996

[30] KIMIA, B.B.; SIDDIQI, K.: *Geometric heat equation and non-linear diffusion of shapes and images*, Computer Vision and Image Understanding, 64, 305–322, 1996

[31] LINDEBERG, T.: *Scale-Space Theory in Computer Vision*, Kluwer, Boston, 1994

[32] NIELSEN, M.; FLORACK, L.; DERICHE, R.: *Regularization, scale-space and edge detection filters*, J. Math. Imag. Vision, 7, 291–307, 1997

[33] NIELSEN, M.; JOHANSEN, P.; OLSEN, O.F.; WEICKERT, J. (EDS.): *Scale-Space Theories in Computer Vision*, Lecture Notes in Computer Science, Vol. 1682, Springer, Berlin, 1999

[34] NIESSEN, W.J., VINCKEN, K.L., WEICKERT, J., VIERGEVER, M.A.: *Nonlinear multiscale representations for image segmentation*, Computer Vision and Image Understanding, 66, 233–245, 1997

[35] NASHED, M.Z.; SCHERZER O.: *Least squares and bounded variation regularization with nondifferentiable functionals*, Numerical Functional Analysis and Optimization, 19, 873–901, 1998

[36] NORDSTRÖM, N.: *Biased anisotropic diffusion – a unified regularization and diffusion approach to edge detection*, Image and Vision Computing, 8, 318–327, 1990

[37] PERONA, P.; MALIK, J.: *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell., 12, 629–639, 1990

[38] RADMOSER, E.; SCHERZER, O.; WEICKERT, J.: *Scale-space properties of regularization methods*, Nielsen, M.; Johansen, P.; Olsen, O.F.; Weickert, J. (Eds.): Scale-space theories in computer vision, Lecture Notes in Computer Science, Vol. 1682, Springer, Berlin, 1999

[39] RUDIN, L.I.; OSHER, S.; FATEMI, E.: *Nonlinear total variation based noise removal algorithms*, Physica D, 60, 259–268, 1992

[40] SAPIRO, G.; TANNENBAUM, A.: *Affine invariant scale-space*, Int. J. Comput. Vision, 11, 25–44, 1993

[41] SCHERZER, O.: *Stable evaluation of differential operators and linear and nonlinear multi-scale filtering*, Electronic Journal of Differential Equations (`http://ejde.math.unt.edu`), No. 15, 1–12, 1997

[42] SCHERZER, O.; WEICKERT, J.: *Relations between regularization and diffusion filtering*, J. Math. Imag. Vision, in press

[43] SCHNÖRR, C.: *Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals*, J. Math. Imag. Vision, 4, 189–198, 1994

[44] SPORRING, J.; NIELSEN, M.; FLORACK, L.; JOHANSEN, P. (EDS.): *Gaussian Scale-Space Theory*, Kluwer, Dordrecht, 1997

[45] SPORRING, J.; WEICKERT, J.: *Information measures in scale-spaces*, IEEE Trans. Information Theory, 45, 1051–1058, 1999

[46] STRONG, D.M.; CHAN, T.F.: *Relation of regularization parameter and scale in total variation based image denoising*, CAM Report 96-7, Dept. of Mathematics, Univ. of California, Los Angeles, CA 90024, U.S.A., 1996

[47] TORRE, V.; POGGIO, T.A.: *On edge detection*, IEEE Trans. Pattern Anal. Mach. Intell., 8, 148–163, 1986

[48] WEICKERT, J.: *Theoretical foundations of anisotropic diffusion in image processing*, Computing, Suppl. 11, 221–236, 1996

[49] WEICKERT, J.: *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, 1998

[50] WEICKERT, J.: *Coherence-enhancing diffusion filtering*, Int. J. Comput. Vision, 31, 111–127, 1999.

[51] WEICKERT, J.; HEERS, J.; SCHNÖRR, C.; ZUIDERVELD, K.J.; SCHERZER, O.; STIEHL, H.S.: *Fast parallel algorithms for a broad class of nonlinear variational diffusion approaches*, Real-Time Imaging, to appear.

[52] WEICKERT, J.; ISHIKAWA, S.; IMIYA, A.: *Linear scale-space has first been proposed in Japan*, J. Math. Imag. Vision, 10, 237–252, 1999

[53] WEICKERT, J.; ZUIDERVELD, K.J.; TER HAAR ROMENY, B.M.; NIESSEN, W.J.: *Parallel implementations of AOS schemes: A fast way of nonlinear diffusion filtering*, Proc. 1997 IEEE International Conference on Image Processing (Santa Barbara, Oct. 26–29, 1997), Vol. 3, 396–399, 1997

[54] WLOKA, J.: *Partielle Differentialgleichungen*, Teubner, Stuttgart, 1982

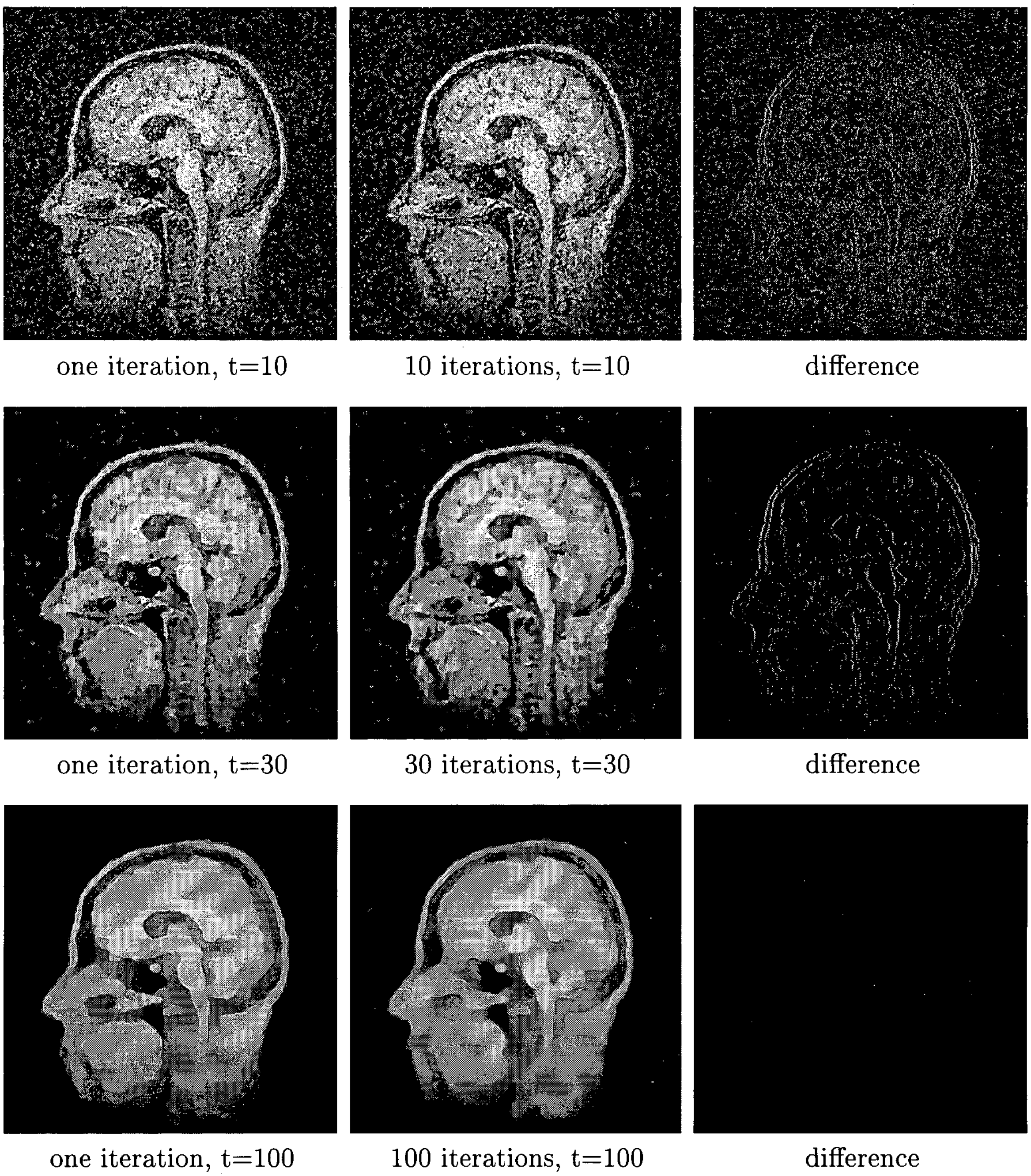
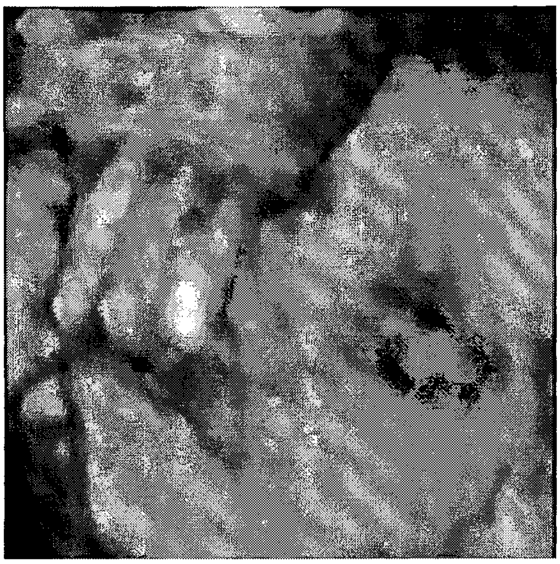
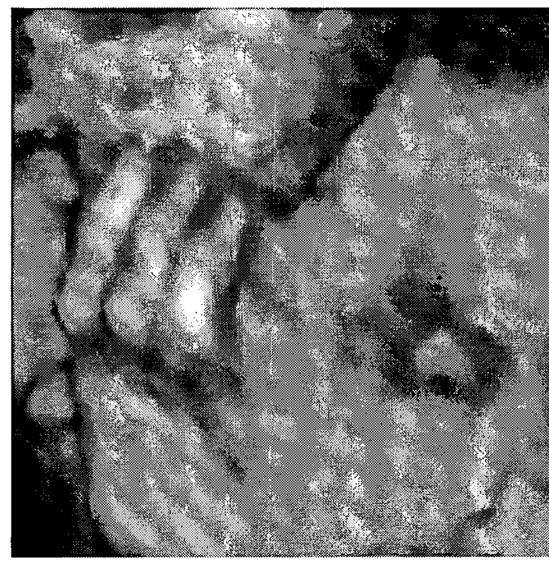


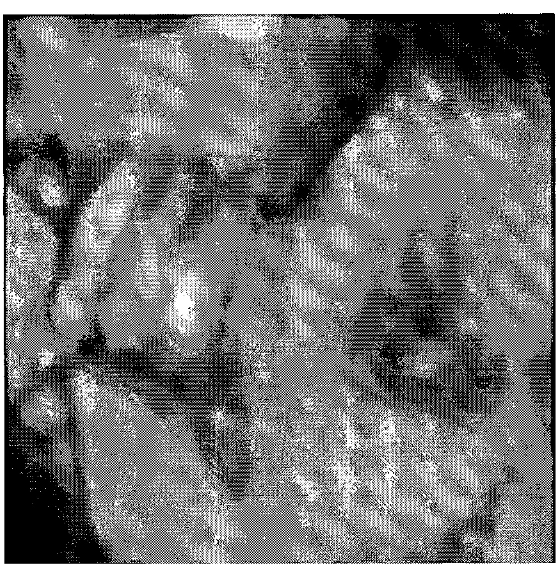
Figure 2: Results for the MR image from Figure 1(a) with noniterated and iterated regularization ($\beta = 0.001$). The left column shows the results for noniterated, the middle column for iterated regularization. The images in the right column depict the modulus of the differences between the results for the iterated and noniterated method.



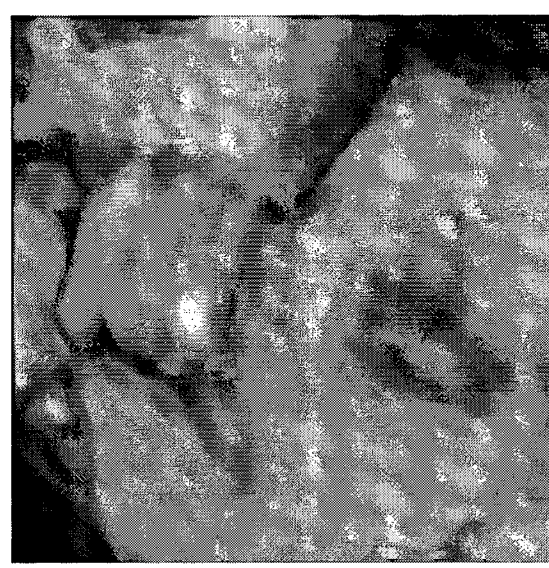
1 iteration, t=8



4 iterations, t=8



1 iteration, t=20



10 iterations, t=20

Figure 3: Results for the three-dimensional ultrasound data from Figure 1(b) with $\beta = 0.001$. The left column shows the renderings for noniterated, the right column for iterated regularization. The regularization parameter for iterated regularization was $h = 2$.