Technical Report DIKU-TR-98/23
Department of Computer Science
University of Copenhagen
Universitetsparken 1
2100 Copenhagen
Denmark

submitted for publication
October 1998

Relations Between Regularization and Diffusion Filtering †

Otmar Scherzer
Institut für Industriemathematik
Johannes–Kepler–Universität
A–4040 Linz, Austria

and

Joachim Weickert
Department of Computer Science
University of Copenhagen
Universitetsparken 1
DK–2100 Copenhagen, Denmark

CR Subject Classification: I.4.3, I.4.4, G.1.8, G.1.6.

Keywords: Regularization, diffusion filtering, image restoration, inverse problems, total variation denoising.

†The work of Otmar Scherzer is supported by the Fonds zur Förderung der Wissenschaftlichen Forschung (Austria) , SFB F1310, as well as the Fonds zur Förderung der Gewerblichen Forschung (FFF), Project 200354. Joachim Weickert is supported by the EU-TMR project VIRGO.

Abstract

Regularization may be regarded as diffusion filtering with an implicit time discretization where one single step is used. Thus, iterated regularization with small regularization parameters approximates a diffusion process. The goal of this paper is to analyse relations between noniterated and iterated regularization and diffusion filtering in image processing. In the linear setting, we show that with iterated Tikhonov regularization noise can be better handled than with noniterated. In the nonlinear framework, two filtering strategies are considered: total variation regularization and the diffusion filter of Perona and Malik. It is established that the Perona-Malik equation decreases the total variation during its evolution. While noniterated and iterated total variation regularization is well-posed, one cannot expect to find a minimizing sequence which converges to a minimizer of the corresponding energy functional for the Perona-Malik filter. To address this shortcoming, a novel regularization of the Perona-Malik process is presented which allows to construct a weakly lower semi-continuous energy functional. In analogy to recently established results for a well-posed class of regularized Perona–Malik filters, we introduce Lyapunov functionals and convergence results for regularization methods. Experiments on real-world images illustrate that iterated linear regularization performs better than noniterated, while no significant differences between noniterated and iterated total variation regularization have been observed.

1 Introduction

Image restoration is among other topics such as optic flow, stereo, and shape-from-shading one of the classical inverse problems in image processing and computer vision [4]. The inverse problem of image restoration consists in recovering information about the original image from incomplete or degraded data. Diffusion filtering has become a popular and well-founded tool for restoration in the image processing community [25, 50], while mathematicians have unified most techniques to treat inverse problems under the theory of regularization methods [14, 19, 30, 44]. Therefore it is natural to investigate relations between both approaches, as this may lead to a deeper understanding and a synthesis of these techniques. This is the goal of the present paper.

We can base our research on several previous results. In the linear setting, Torre and Poggio [45] 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. [16, 29, 42] and the references therein. In a very nice work, Nielsen et al. [31] derived linear diffusion filtering axiomatically from Tikhonov regularization, where the stabilizer consists of a sum of squared derivatives up to infinite order.

In the nonlinear diffusion framework, 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 [34, 41, 9]. The regularization parameter and the diffusion time can be identified if one regards regularization as time-discrete diffusion filtering with a single implicit time step [43, 39]. A popular specific energy functional arises from unconstrained total variation denoising [1, 8, 6]. Constrained total variation also leads to a nonlinear diffusion process with a bias term using a time-dependent Lagrange multiplier [38].

In spite of these numerous relations, several topics have not been addressed so far in the literature:

- *A comparison of the restoration properties of both approaches:* Since regularization corresponds to time-discrete diffusion filtering with a single time step, it follows that iterated regularization with a small regularization parameter gives a better approximation to diffusion filtering. An investigation whether iterated regularization is better than noniterated leads therefore to a comparison between regularization and diffusion filtering.

- *Energy formulations for stabilized Perona–Malik processes:* The Perona–Malik filter is the oldest nonlinear diffusion filter [36]. Its ill-posedness has triggered many researchers to introduce regularizations which have shown their use for image restoration. However, no regularization has been found which can be linked to the minimization of an appropriate energy functional.

- *Lyapunov functionals for regularization:* The smoothing and information-reducing properties of diffusion filters can be described by Lyapunov functionals such as decreasing L^p norms, decreasing even central moments, or increasing entropy [50]. They constitute important properties for regarding diffusion filters as scale-spaces. A corresponding scale-space interpretation of regularization methods where the regularization parameter serves as scale parameter has been missing so far.

These topics will be discussed in the present paper. It is organized as follows. Section 2 explains the relations between variational formulations of diffusion processes and regularization strategies. In Section 3 we first discuss the noise propagation for noniterated and iterated Tikhonov regularization for linear problems. In the nonlinear framework, well-posedness results for total variation regularization are reviewed and it is explained why one cannot expect to establish well-posedness for the Perona–Malik filter. We will argue that, if the Perona–Malik filter admits a smooth solution, however, then it will be total variation reducing. A novel regularization will be introduced which allows to construct a corresponding energy functional. Section 4 establishes Lyapunov functionals for regularization methods which are in accordance with those for diffusion filtering. This leads to a scale-space interpretation for linear and nonlinear regularization. In Section 5 we shall present some experiments with noisy real-world images, which compare the restoration properties of noniterated and iterated regularization in the linear setting and in the nonlinear total variation framework. Moreover, the novel Perona–Malik regularization is juxtaposed to the regularization by Catté et al. [5]. The paper will be concluded with a summary in Section 6.

2 Variational formulations of diffusion processes and the connection to regularization methods

We consider a general *diffusion process* of the form

$$\begin{cases} \partial_t u(x,t) = \nabla.\left(g(|\nabla u|^2)\nabla u\right)(x,t) & \text{on } \Omega \times (0,\infty[\\ \partial_n u = 0 & \text{on } \Gamma \times (0,\infty[\\ u(x,0) = f_\delta(x) & \text{on } \Omega\,. \end{cases} \tag{2.1}$$

Here g is a smooth function satisfying certain properties which will be explained in the course of the paper; $\Omega \subseteq \mathbb{R}^d$ is a bounded domain with piecewise Lipschitzian boundary Γ with unit normal vector n, and f_δ is a degraded version of the original image $f := f_0 : \Omega \to \mathbb{R}$.

For the numerical solution of (2.1) one can use explicit or implicit or semi-implicit difference schemes with respect to t.

The implicit scheme reads as follows

$$\begin{cases} \frac{u(x,t+h)-u(x,t)}{h} = \nabla.\left(g(|\nabla u|^2)\nabla u\right)(x,t+h) \\ u(x,0) = f_\delta(x)\,. \end{cases} \tag{2.2}$$

Here $h > 0$ denotes the step-size in t–direction of the implicit difference scheme.

In the following we assume that g is measurable on $[0,\infty[$ and there exists a differentiable function $\hat{g}$ on $[0,\infty)$ which satisfies $\hat{g}' = g$. Then the minimizer of the functional (for given $u(x,t)$)

$$T(u) := \|u - u(x,t)\|^2 + h \int_\Omega \hat{g}(|\nabla u|^2)\,, \tag{2.3}$$

satisfies (2.2) at time $t+h$. If the functional T is convex, then a minimizer of T is uniquely characterized by the solution of the equation (2.2) with homogeneous Neumann boundary conditions.

$T(u)$ is a typical regularization functional consisting of the approximation functional $\|u - u(x,t)\|^2$ and the stabilizing functional $\int_\Omega \hat{g}(|\nabla u|^2)$. The weight h is called *regularization parameter*. The case $\hat{g}(x) = x$ is called *Tikhonov regularization*.

In the next section we summarize some results on regularization and diffusion filtering and compare the theoretical results developed in both theories.

3 A survey on diffusion filtering and regularization

We have seen that each time step for the solution of the diffusion process (2.1) with an implicit, t-discrete scheme is equivalent to the calculation of the minimizer of the regularization functional (2.3). The numerical solution of the diffusion process with an implicit, t-discrete iteration scheme is therefore equivalent to *iterated regularization* where on has to minimize iteratively the set of functionals

$$T_n(u) := \|u - u_{n-1}\|^2 + h_n \int_\Omega \hat{g}(|\nabla u|^2). \tag{3.1}$$

Here u_n is a minimizer of the functional $T_n, n = 1, 2, ...,$ and $u_0 := f_\delta$. If the functionals T_n are convex, then the minimizer of (3.1) denoted by u_n is the approximation of the solution of the diffusion process with an implicit, t-discrete method at time $t_1, ..., t_n$ where $t_k = \sum_{j=1}^k h_j$.

In the following we refer to iterated regularization if $h_n = h$ for all n. That corresponds to the solution of the diffusion process with an implicit, t-discrete method using a fixed time step size $h = h_n$.

If the regularization parameters h_n are adaptively chosen (this corresponds to the situation that the time discretization in the diffusion process is changed adaptively), then the method is called *nonstationary regularization*. For some recent results on nonstationary Tikhonov regularization we refer to Hanke and Groetsch [24]; however, their results do not fit directly into the framework of this paper. They deal with regularization methods for the stable solution of operator equations

$$Iu = y\,, \tag{3.2}$$

where I is a linear bounded operator from a Hilbert space X into a Hilbert space Y, and they use nonstationary Tikhonov regularization

$$\min(\|Iu - y\|^2 + h_n \|u - u_{n-1}\|^2)$$

for the stable solution of the operator equation (3.2).

3.1 Error propagation of Tikhonov regularization with linear unbounded operators

In this subsection we consider the problem of computing values of an unbounded operator L. We will always denote by $L : \mathcal{D}(L) \subseteq H_1 \to H_2$ a closed, densely defined unbounded linear operator between two Hilbert spaces H_1 and H_2. A typical example is $Lu = \nabla u$. The problem of computing values $y = Lf_0$, for $f_0 \in D(L)$ is then ill–posed in the sense that small perturbations in f_0 may lead to data f_δ satisfying

$$\|f_0 - f_\delta\| \leq \delta, \tag{3.3}$$

but $f_\delta \notin \mathcal{D}(L)$, or even if $f_\delta \in \mathcal{D}(L)$, it may happen that $Lf_\delta \not\to Lf_0$ as $\delta \to 0$, since the operator L is unbounded. Morozov has studied a stable method for approximating the value Lf_0 when only approximate data f_δ is available [30]. This method takes as an approximation to $y = Lf_0$ the vector $y_h^\delta = Lu_h^\delta$, where u_h^δ minimizes the functional

$$T_{\text{TIK}}(u) := \|u - f_\delta\|^2 + h\|Lu\|^2 \quad (h > 0) \tag{3.4}$$

over $D(L)$.

The functional is strictly convex and therefore if $D(L)$ is nonempty and convex there exists a unique minimizer of the functional $T_{\text{TIK}}(u)$. Thus the method is well–defined. For more background on the stable evaluation of unbounded operators we refer to [20].

Let $u_0 := f_\delta$. If $L = \nabla$, then the sequence $\{u_n\}_{n\geq 1}$ of minimizers of the family of optimization problems

$$T_{\text{TIK}}^n := \|u - u_{n-1}\|^2 + h\|\nabla u\|^2 \quad n \geq 1 \tag{3.5}$$

are identical to the semi-discrete approximations of the differential equation (2.1) at time nh $(n \geq 1)$ where $g(x) = x$.

This shows

> Methods for evaluating unbounded operators can be used for diffusion filtering and vice versa. However the motivations differ: For evaluating unbounded operators we solve the optimization and evaluate in a further step the unbounded operator. In diffusion filtering we "only" have to solve the optimization problem.

In the following we compare the error propagation in Tikhonov regularization with regularization parameter h and the error propagation in iterated Tikhonov regularization of order N with regularization parameter h/N. This corresponds to making an implicit, t-discrete ansatz for a diffusion process with one step h and an implicit, t-discrete ansatz with N steps of step h/N, respectively.

Tikhonov regularization with regularization parameter h reads as follows

$$\hat{u} = (I + hL^*L)^{-1} f_\delta$$

where L^* is the adjoint operator to L (see e.g. [47] for more details). Tikhonov regularization of order N with regularization parameter h/N reads as follows

$$u_N = \left(I + \frac{h}{N} L^* L\right)^{-N} f_\delta \, .$$

Let L^*L be an unbounded operator with spectral values

$$0 < \lambda_1 \leq ..... \leq \lambda_n ....$$

such that $\lambda_n \to \infty$ as $n \to \infty$. Then

$$\begin{aligned} u_N &= \left(I + \tfrac{h}{N} L^*L\right)^{-N} f_\delta \\ &= \left(I + \tfrac{h}{N} L^*L\right)^{-N} (f_\delta - f_0) + \left(I + \tfrac{h}{N} L^*L\right)^{-N} f_0 \ . \end{aligned}$$

$\left(I + \tfrac{h}{N} L^*L\right)^{-N} (f_\delta - f_0)$ denotes the propagated error of the initial data f_δ, which remains in u_N – this corresponds to the error propagation in diffusion filtering with an implicit, t-discrete method.

Let E_λ be the spectral family according to the operator L^*L. Then it follows that [47]

$$\left(I + \frac{h}{N} L^*L\right)^{-N} (f_\delta - f_0) = \int_0^\infty \left(1 + \frac{h}{N}\lambda\right)^{-N} E_\lambda (f_\delta - f_0) \ .$$

Using

$$\left(1 + \frac{h}{N}\lambda\right)^{N} = 1 + \frac{h}{N}\lambda + \binom{N}{2}\left(\frac{h}{N}\right)^2 \lambda^2 + ... + \left(\frac{h}{N}\right)^N \lambda^N$$

we get that

$$\begin{aligned} &\left\| \left(I + \tfrac{h}{N} L^*L\right)^{-N} (f_\delta - f_0) \right\|^2 \\ &= \int_0^\infty \left(1 + \tfrac{h}{N}\lambda + \tbinom{N}{2}\left(\tfrac{h}{N}\right)^2 \lambda^2 + ... + \left(\tfrac{h}{N}\right)^N \lambda^N\right)^{-2} E_\lambda^2 \|f_\delta - f_0\|^2 \ . \end{aligned} \tag{3.6}$$

In noniterated Tikhonov regularization the error propagation is

$$\left\| (I + hL^*L)^{-1} (f_\delta - f_0) \right\|^2 = \int_0^\infty (1 + h\lambda)^{-2} E_\lambda^2 \|f_\delta - f_0\|^2 \ . \tag{3.7}$$

For large values of λ (i.e., for highly oscillating noise) the term $(1 + h\lambda)^{-2}$ in (3.7) is significantly larger than the term $\left(1 + \tfrac{h}{N}\lambda + \tbinom{N}{2}\left(\tfrac{h}{N}\right)^2 \lambda^2 + ... + \left(\tfrac{h}{N}\right)^N \lambda^N\right)^{-2}$ in (3.6). This shows that noise propagation is handled more efficiently by iterated Tikhonov regularization than by Tikhonov regularization.

Above we analyzed the error of the (iterated) Tikhonov regularized solutions and not the error in evaluating L at the Tikhonov regularized solutions. We emphasize that the less noise is contained in a data set the better the operator L can be evaluated. Therefore we conclude that the operator L can be evaluated more accurately with the method of iterated Tikhonov regularization than with noniterated Tikhonov regularization. This will be confirmed by the experiments in Section 5.

3.2 Well-posedness of regularization with nonlinear unbounded operators

In this subsection we discuss some theoretical results on regularization with nonlinear unbounded operators.

3.2.1 Well-posedness and convergence for total variation regularization

Total variation regularization goes back to Rudin, Osher and Fatemi [38] and has been further analysed by many others, e.g. [1, 7, 6, 8, 12, 13, 27, 28, 43, 40, 46]. In the unconstrained formulation of this method the data f_0 is approximated by the minimizer of the functional over $\mathrm{TV}(\Omega)$, the space of all functions with finite total variation norm

$$T_{\mathrm{TV}}(u) := \|u - f_\delta\|^2 + h\mathrm{TV}(u)\,, \tag{3.8}$$

where $\mathrm{TV}(u) := \int_\Omega |\nabla u|$ and

$$\int_\Omega |\nabla u| := \sup\{-\int_\Omega u\nabla.\rho : \rho \in C_0^1(\Omega, \mathbb{R}^d)\,, |\rho| \leq 1\}\,.$$

This expression extends the usual definition of the total variation for smooth functions to functions with jumps [22].

It is easy to see that a smooth minimizer of the functional T_{TV} satisfies

$$u - f_\delta = \frac{h}{2}\nabla.\left(\frac{1}{|\nabla u|}\nabla u\right)\,. \tag{3.9}$$

Acar and Vogel [1] proved the following results concerning existence of a minimizer of (3.8) and concerning stability and convergence of the minimizers:

Theorem 3.1 (Existence of a minimizer) *Let $f_\delta \in L^2(\Omega)$, then for fixed $h > 0$ a minimizer $u_h \in TV(\Omega)$ of (3.8) exists and is unique.*

Theorem 3.2 (Stability) *Let $f_\delta \in L^2(\Omega)$ and $f_0 \in TV(\Omega)$. Then for $\delta \to 0$*

$$u_h(f_\delta) \to u_h(f_0)$$

with respect to the L^p-norm ($1 \leq p < \frac{d}{d-1}$). Here $u_h(f_\delta)$ is the minimizer of (3.8) and $u_h(f_0)$ is the minimizer of (3.8) where f^δ is replaced by f_0.

Theorem 3.3 (Convergence) *Let $f_\delta \in L^2(\Omega)$ and $f_0 \in TV(\Omega)$ satisfy*

$$\|f_\delta - f_0\|_{L^2(\Omega)} \leq \delta\,.$$

Then for $h := h(\delta)$ satisfying $\frac{\delta^2}{h} \to 0$ as $\delta \to 0$

$$u_h \to f_0$$

with respect to the L^p-norm ($1 \leq p < \frac{d}{d-1}$).

It is evident that analogous results to Theorem 3.1, Theorem 3.2 and 3.3 also hold for the minimizers of the *iterated total variation regularization* which consists of minimizing a sequence of functionals

$$T^N_{\mathrm{TV}}(u) := \|u - u_{n-1}\|^2 + h \int_\Omega |\nabla u| \,, \tag{3.10}$$

where u_{N-1} denotes the minimizer of the functional T^{N-1}_{TV} and $u_0 = f_\delta$.

This regularization technique corresponds to the implicit, t-discrete approximation of the diffusion process (2.1) with $\hat{g}(x) = \sqrt{x}$.

3.2.2 The Perona-Malik filter

In the Perona-Malik filter [36] we have $g(s) = \frac{1}{1+s}$ and $\hat{g}(s) = \ln(1+s)$. Iterated Perona-Malik regularization minimizes the family of functionals

$$T^n_{\mathrm{PM}}(u) := \|u - u_{n-1}(x)\|^2 + h \int_\Omega \ln(1 + |\nabla u|^2) \,. \tag{3.11}$$

The functionals T^n_{PM} are not convex and therefore we cannot conclude that the minimizer of (3.11) (it it exists) satisfies the first order optimality condition

$$\frac{1}{h}(u - u_{n-1})(x) = \nabla . \left(\frac{\nabla u}{1 + |\nabla u|^2} \right) , \tag{3.12}$$

with homogeneous Neumann boundary data.

In the following we comment on some aspects of the Perona-Malik regularization technique. For the definitions of the Sobolev spaces $W^{l,p}$ and the notion of weak lower semi-continuity we refer to [2].

1. *Neumann boundary conditions:* Let Ω be a domain with smooth boundary $\partial\Omega$. Using trace theorems (see e.g. [2]) it follows that the Neumann boundary data are well–defined in $L^2(\partial\Omega)$ for any function in $W^{\frac{3}{2},2}(\Omega)$. Suppose we could prove that there exists a minimizer of the functional T^n_{PM}, then this minimizer must satisfy

   $$\int_\Omega \ln(1 + |\nabla u|^2) \, ds < \infty \,. \tag{3.13}$$

   Elementary calculations show that any function $u \in W^{1,p}(\Omega)$ $(p > 1)$ satisfies (3.13). Therefore we cannot deduce from (3.13) that the minimizer is in any Sobolev space $W^{1,p}(\Omega)(p > 1)$. Consequently, there exists no theoretical result that the Neumann boundary conditions are well-defined.

2. *Existence of a minimizer of the functional* T^n_{PM}: The function $\ln(1+s^2)$ is not convex, and therefore the functional $T^n_{\mathrm{PM}}(u)$ is not weakly lower semi-continuous on $W^{1,p}(\Omega)$ for any $1 < p < \infty$ (see [11, p. 66] and also [10])).

   Therefore, there exists a sequence $u_k \in W^{1,p}(\Omega)$ with $u_k \rightharpoonup u$ in $W^{1,p}(\Omega)$, but

   $$\liminf T^n_{\mathrm{PM}}(u_k) < T^n_{\mathrm{PM}}(u) \, .$$

   Consequently, we cannot expect that a minimizing sequence converges (in $W^{1,p}(\Omega)$) to a minimizer of the functional T^n_{PM}. Thus the solution of the Perona-Malik regularization technique is ill–posed on $W^{1,p}(\Omega)$!

The diffusion process associated with the Perona-Malik regularization technique is

$$\partial_t u = \nabla . \left(\frac{\nabla u}{1+|\nabla u|^2} \right) \, . \tag{3.14}$$

The Perona-Malik diffusion filtering technique can be split up in a natural way into a forward and a backward diffusion process:

$$\begin{aligned} \partial_t u &= \nabla . \left(\tfrac{1}{1+|\nabla u|^2} \nabla u \right) \\ &= \tfrac{(1-|\nabla u|^2)\Delta u}{(1+|\nabla u|^2)^2} \\ &=: (a(|\nabla u|) - b(|\nabla u|))\Delta u \, . \end{aligned}$$

Here

$$a(|\nabla u|) := \frac{1}{(1+|\nabla u|^2)^2} \, , \quad b(|\nabla u|) := \frac{|\nabla u|^2}{(1+|\nabla u|^2)^2} \, .$$

Both functions a and b are non-negative. In general the solution of a backward diffusion equation is severely ill–posed (see e.g. [14]). We argue below that this nonlinear backward diffusion is well–posed with respect to appropriate norms. In fact we argue that the backward diffusion equation

$$v_t = -b(|\nabla v|)\Delta v \tag{3.15}$$

satisfies

$$TV(v(.,t)) = TV(v(.,0)) =: TV(v_0(.)) \, . \tag{3.16}$$

The intuitive reason for the validity of this is the following: Let $v \in C^2(\Omega \times [0,T])$ then

$$\partial_t |\nabla v| \approx \partial_t \sqrt{|\nabla v|^2 + \beta^2} = \frac{\nabla v \nabla \partial_t v}{\sqrt{|\nabla v|^2 + \beta^2}} \, . \tag{3.17}$$

Using (3.17), (3.15), and integration by parts it follows that

$$\begin{aligned}
\textstyle\int_\Omega \partial_t |\nabla v| &\approx \textstyle\int_\Omega \frac{\nabla v}{\sqrt{|\nabla v|^2+\beta^2}} \nabla \partial_t v \\
&= \textstyle\int_\Omega \nabla . \left(\frac{\nabla v}{\sqrt{|\nabla v|^2+\beta^2}} \right) \frac{|\nabla v|^2}{(1+|\nabla v|^2)^2} \Delta v \\
&= \beta^2 \textstyle\int_\Omega \frac{|\nabla v|^2}{|\nabla v|^2+\beta^2} \frac{|\Delta v|^2}{(1+|\nabla v|^2)^2} \;.
\end{aligned}$$

If $v \in C^2(\Omega \times [0,T])$ then the right hand side tends to zero as $\beta \to 0$. These arguments indicate that

$$\partial_t \int_\Omega |\nabla v| = 0 \;.$$

Consequently the total variation of $v(.,t)$ does not change in the course of the evolutionary process (3.17). Indeed, (3.15) may be regarded as a total variation preserving shock filter in the sense of Osher and Rudin [35].

The diffusion process

$$\partial_t u = a(|\nabla u|) \Delta u$$

is a forward diffusion process which decreases the total variation during the evolution. In summary we have argued that the Perona-Malik diffusion equation decreases the total variation during the evolutionary process.

3.2.3 A regularized Perona-Malik filter

Although the ill-posedness of the Perona-Malik filter can be handled by applying regularizing finite difference discretizations [51], it would be desirable to have a regularization which does not depend on discretization effects. In this subsection we study a regularized Perona-Malik filter

$$T^n_{\text{R-PM}}(u) := \|u - u_{n-1}(x)\|^2 + h \int_\Omega \ln(1 + |\nabla L_\gamma u|^2) \,, \tag{3.18}$$

where L_γ is linear and compact from $L^2(\Omega)$ into $C^1(\overline{\Omega})$. The applications which we have in mind include the case that L_γ is a convolution operator with a smooth kernel.

In the following we prove that the functional $T^n_{\text{R-PM}}$ attains a minimium:

Theorem 3.4 *The functional $T^n_{R\text{-}PM}$ is weakly lower semi continuous on $L^2(\Omega)$.*

Proof: Let $\{u_s : s \in \mathbb{N}\}$ be a sequence in $L^2(\Omega)$ which satisfies

$$T^n_{\text{R-PM}}(u_s) \to \min\{T^n_{\text{R-PM}}(u) : u \in L^2(\Omega)\} \;.$$

Then $\{u_s\}$ has a weakly convergent subsequence (which is again denoted by $\{u_s\}$) with weak limit u. Since L_γ is compact from $L^2(\Omega)$ into $C^1(\overline{\Omega})$ the sequence $\ln(1 + |\nabla L_\gamma u_s|^2)$ converges uniformly to $\ln(1 + |\nabla L_\gamma u|^2)$. In particular, we have

$$\int_\Omega \ln\left(1 + |\nabla L_\gamma u_s|^2\right) \to \int_\Omega \ln\left(1 + |\nabla L_\gamma u|^2\right) \;.$$

Using the weak lower semi-continuity of the norm $\|.\|_{L^2(\Omega)}$ it follows that the functional $T^n_{\text{R-PM}}$ is weakly lower semi-continuous. q.e.d.

The minimizer of the regularized Perona-Malik functional satisfies

$$u - u_{n-1}(x) = hL_\gamma^* \nabla. \left(\frac{\nabla L_\gamma u}{1 + |\nabla L_\gamma u|^2} \right) . \tag{3.19}$$

The corresponding nonlinear diffusion process associated with this regularization technique is

$$\partial_t u(x) = L_\gamma^* \nabla. \left(\frac{\nabla L_\gamma u}{1 + |\nabla L_\gamma u|^2} \right) . \tag{3.20}$$

Regularized Perona-Malik filters have been considered in the literature before [3, 5, 32, 48, 50]. Catté et al. [5] for instance investigated the nonlinear diffusion process

$$\partial_t u(x) = \nabla. \left(\frac{\nabla u}{1 + |\nabla L_\gamma u|^2} \right) . \tag{3.21}$$

This technique (as well as other previous regularizations) does not have a corresponding formulation as an optimization problem. The differences between (3.20) and (3.21) will be explained in Section 5.

4 Lyapunov functionals for regularization methods

Lyapunov functionals play an important role in continuous diffusion filtering (see [49, 50]). In order to introduce Lyapunov functionals of regularization methods, we first give a survey on Lyapunov functionals in diffusion filtering. We consider the diffusion process (here and in the following Ω will be a domain with piecewise smooth boundary)

$$\begin{array}{rcllll} \partial_t u(x,t) &=& \nabla. \left(g(L_\gamma(\nabla u))\nabla u\right) & \text{on} & \Omega \times (0,T) \\ u(x,0) &=& f(x) & \text{on} & \Omega \\ \partial_n u &=& 0 & \text{on} & \Gamma \times (0,T) \end{array} \tag{4.1}$$

We assume that the following assumptions hold:

1. $f \in L^\infty(\Omega)$, with $a := \text{ess inf}_{x\in\Omega} f$ and $b := \text{ess sup}_{x\in\Omega} f$.
2. L_γ is a compact operator from $L^2(\Omega)$ into $C^p(\overline{\Omega})$ for any $p \in \mathbb{N}$.
3. $T > 0$.
4. For all $w \in L^\infty(\Omega, \mathbb{R}^d)$ with $|w(x)| \leq K$ on $\overline{\Omega}$, there exists a positive lower bound $\nu(K)$ for g.

The regularizing operator L_γ may be skipped in (4.1), if one assumes that $\hat{g}(|.|^2)$ is convex from $\mathbb{R}^d$ to $\mathbb{R}$. Moreover, it is also possible to generalize (4.1) to the anisotropic case where the diffusivity g is replaced by a diffusion tensor [50].

Under the preceding assumptions it can be shown that (4.1) is well-posed (see [5, 50]):

Theorem 4.1 *The equation (4.1) has a unique solution $u(x,t)$ which satisfies*

$$u \in C([0,T]; L^2(\Omega)) \cap L^2([0,T]; H^1(\Omega)) \tag{4.2}$$

$$\partial_t u \in L^2([0,T]; H^1(\Omega)). \tag{4.3}$$

Moreover,

$$u \in C^\infty(\overline{\Omega} \times [0,T]) \ .$$

The solution fulfills the extremum principle

$$a \le u(x,t) \le b \ on \ \Omega \times (0,T]. \tag{4.4}$$

For fixed t the solution depends continuously on f with respect to $\|.\|_{L^2(\Omega)}$.

This diffusion process leads to the following class of Lyapunov functionals [50]:

Theorem 4.2 *Suppose that u is a solution of (4.1) and that assumptions 1 – 4 are satisfied. Then the following properties hold*

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

$$V(t) := \phi(u(t)) := \int_\Omega r(u(x,t))\, dx \tag{4.5}$$

is a Lyapunov functional:

1. *$\phi(u(t)) \ge \phi(Mf)$ for all $t \ge 0$; here*

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

2. *$V \in C[0,\infty) \cap C^1(0,\infty)$ and $V'(t) \le 0$ for all $t > 0$.*

Moreover, if $r'' > 0$ on $[a,b]$, then $V(t) = \phi(u(t))$ is a strict Lyapunov functional:

3. *$\phi(u(t)) = \phi(Mf)$ if and only if $u(t) = Mf$ on $\overline{\Omega}$ for $t > 0$ and $u(t) = Mf$ a.e. on $\overline{\Omega}$ for $t = 0$.*
4. *If $t > 0$, then $V'(t) = 0$ if and only if $u(t) = Mf$ on $\overline{\Omega}$.*
5. *$V(0) = V(T)$ for $T > 0$ if and only if $f = Mf$ a.e. on Ω and $u(t) = Mf$ a.e. on $\overline{\Omega} \times (0,T]$.*

(b) (Convergence)

1. $\lim_{t\to\infty} \|u(t) - Mf\|_{L^p(\Omega)} = 0$ *for* $p \in [1,\infty)$.

2. If $\Omega \subseteq \mathbb{R}$*, then the convergence* $\lim_{t\to\infty} u(x,t) = Mf$ *is uniform.*

In the sequel we introduce Lyapunov functionals of regularization methods.

In the beginning of this section we discuss existence and uniqueness of the minimizer of the regularization functional in $H^1(\Omega)$

$$I(u) := \|u - f_\delta\|^2_{L^2(\Omega)} + h \int_\Omega \hat{g}(|\nabla u|^2) \ . \tag{4.6}$$

Lemma 4.3 *Let* $\Omega \subseteq \mathbb{R}^d$*,* $d \geq 1$*. Moreover, let* $\hat{g}$ *satisfy:*

$$\begin{gathered} \hat{g}(.) \text{ is in } C^0(K) \text{ for any compact } K \subseteq [0,\infty[\\ \hat{g}(0) = \min\{\hat{g}(x) : x \in [0,\infty[\} \end{gathered} \tag{4.7}$$

$$\hat{g}(|.|^2) \text{ is convex from } \mathbb{R}^d \text{ to } \mathbb{R}\ . \tag{4.8}$$

Moreover, we assume that there exists a constant $c > 0$ *such that*

$$\hat{g}(s) \geq cs \tag{4.9}$$

Then the minimizer of (4.6) exists and is unique in $H^1(\Omega)$*.*

Proof: By virtue of (4.9) it follows that

$$\|u - f_\delta\|^2_{L^2(\Omega)} + h \int_\Omega \hat{g}(|\nabla u|^2) \geq \|u - f_\delta\|^2_{L^2(\Omega)} + h \int_\Omega c|\nabla u|^2 \ . \tag{4.10}$$

Suppose now that u_n is a sequence such that $I(u_n)$ converges to the minimum of the functional $I(.)$ in $H^1(\Omega)$. From (4.10) it follows that u_n has a weakly convergent subsequence in $H^1(\Omega)$, which we also denote by u_n; the weak limit will be denoted by u_*. Since $\hat{g}(|.|^2)$ is convex, the functional $\int_\Omega \hat{g}(|\nabla u|^2)$ is weakly lower semi continuous in $H^1(\Omega)$ (see [11, 10]), and thus

$$\int_\Omega \hat{g}(|\nabla u_*|^2) \leq \liminf_{n\in\mathbb{N}} \int_\Omega \hat{g}(|\nabla u_n|^2) \ .$$

Thanks to the the Sobolev embedding theorem (see [2]) it follows that the functional $\|u - f_\delta\|^2_{L^2(\Omega)}$ is weakly lower semi continuous on $H^1(\Omega)$. Consequently

$$I(u_*) \leq \liminf_{n\in\mathbb{N}} I(u_n)$$

and thus u_* is a minimizer of I in $H^1(\Omega)$. Suppose now that u_1 and u_2 are two minimizers of the functional I. Then, from the optimality condition it follows that

$$\langle u_1 - f_\delta, u_2 - u_1\rangle + h\,\langle g(|\nabla u_1|^2)\nabla u_1, \nabla(u_2 - u_1)\rangle = 0 \tag{4.11}$$

$$\langle u_2 - f_\delta, u_2 - u_1\rangle + h\,\langle g(|\nabla u_2|^2)\nabla u_2, \nabla(u_2 - u_1)\rangle = 0. \tag{4.12}$$

Consequently

$$\|u_2 - u_1\|^2 + h \left\langle g(|\nabla u_2|^2)\nabla u_2 - g(|\nabla u_1|^2)\nabla u_1, \nabla(u_2 - u_1)\right\rangle = 0 \,.$$

And thus the minimizer of I is unique. q.e.d

The minimizer of (4.6) will be denoted by u_h in the remaining of this paper.

In the following we establish the average grey level invariance of regularization methods.

Theorem 4.4 *Let (4.7), (4.8), (4.9) hold. Then for different values of h the minimizers of (4.6) are grey-level invariant, i.e., for $h > 0$*

$$\int_\Omega u_h = \int_\Omega f_\delta .$$

Proof: Elementary calculations show that the minimizer of (4.6) satisfies for all $v \in H^1(\Omega)$

$$\langle u_h - f_\delta, v\rangle + h \left\langle g(|\nabla u_h|^2)\nabla u_h, \nabla v \right\rangle = 0 \,. \tag{4.13}$$

Taking $v = 1$ the second term vanishes and the assertion follows. q.e.d.

In the following we establish some basic results on regularization techniques. As we will show the proofs of the following results can be carried out following the ideas of the corresponding results in the book of Morozov [30]. However Morozov's results can not be applied directly since they are only applicable in the case that $\hat{g}(|x|^2) = |x|^2$, which is not sufficient for the presentation of this paper. Later these results are used to establish a family of Lyapunov functionals for regularization methods.

Lemma 4.5 *Let (4.7), (4.8), (4.9) hold. Then for any $h > 0$*

$$\|u_{h+t} - u_h\|_{L^2(\Omega)} \to 0 \qquad \textit{for } t \to 0$$

and for $h = 0$

$$\|u_t - f_\delta\|_{L^2(\Omega)} \to 0 \qquad \textit{for } t \to 0^+ \,.$$

Proof: If $\hat{g}(|.|^2)$ is convex, then $g(|s|^2)s$ is monotone (see e.g. [11]), i.e., for all $s, t \in \mathbb{R}^d$

$$\langle g(|s|^2)s - g(|t|^2)t, s - t\rangle_{\mathbb{R}^d} \geq 0.$$

1. First we consider the case $h > 0$: from (4.13) it follows by using the notation

   $$h_1 := h, \quad h_2 := h + t, \quad u_1 := u_h, \quad u_2 := u_{h+t}$$

   that

   $$\begin{aligned} \langle u_1 - f_\delta, u_2 - u_1\rangle + h_1 \left\langle g(|u_1|^2)\nabla u_1, \nabla(u_2 - u_1)\right\rangle &= 0. \\ \langle u_2 - f_\delta, u_2 - u_1\rangle + h_2 \left\langle g(|u_2|^2)\nabla u_2, \nabla(u_2 - u_1)\right\rangle &= 0. \end{aligned}$$

Consequently

$$\|u_2 - u_1\|^2_{L^2(\Omega)} + h_1 \langle g(|u_2|^2)\nabla u_2 - g(|u_1|^2)\nabla u_1, \nabla(u_2-u_1)\rangle \tag{4.14}$$
$$= (h_1 - h_2)\langle g(|u_2|^2)\nabla u_2, \nabla(u_2-u_1)\rangle.$$

Thus using the Cauchy–Schwarz inequality and the identity (4.14) it follows that

$$\|u_2 - u_1\|^2_{L^2(\Omega)} \le |h_2 - h_1| \frac{\|u_2 - f_\delta\|_{L^2(\Omega)} \|u_2 - u_1\|_{L^2(\Omega)}}{h_2}$$

which shows the continuity of u_h.

2. If $h = 0$: There exists a sequence $f_n \in H^1(\Omega)$ with $f_n \to f_\delta$ in $L^2(\Omega)$. Consequently for any $h > 0$ it follows from the definition of a minimum of the Tikhonov-like functional it follows that

$$\|u_h - f_\delta\|^2_{L^2(\Omega)} \le \|f_n - f_\delta\|^2_{L^2(\Omega)} + h \int_\Omega \hat{g}(|\nabla f_n|^2)\ .$$

Consequently by taking the limit $h \to 0$ it follows that for any $n \in \mathbb{N}$

$$\lim_{h\to 0} \|u_h - f_\delta\|^2_{L^2(\Omega)} \le \|f_n - f_\delta\|^2_{L^2(\Omega)}\ ,$$

which shows the assertion.

q.e.d.

In the following we present some monotonicity results for the regularized solutions.

Lemma 4.5 implies that we can set $u_0 = f_\delta$ without causing any confusion.

Lemma 4.6 *Let (4.7), (4.8), (4.9) hold. Then* $\int_\Omega \hat{g}(|\nabla u_h|^2)$ *is monotonically decreasing in h and* $\|u_h - f_\delta\|^2$ *is monotonically increasing in h.*

Proof: Using the definition of the regularized solution it follows

$$\begin{aligned}
&\|u_h - f_\delta\|^2_{L^2(\Omega)} + h \int_\Omega \hat{g}(|\nabla u_h|^2)\\
&\le \|u_{h+t} - f_\delta\|^2_{L^2(\Omega)} + (h+t) \int_\Omega \hat{g}(|\nabla u_{h+t}|^2) - t \int_\Omega \hat{g}(|\nabla u_{h+t}|^2)\\
&\le \|u_h - f_\delta\|^2_{L^2(\Omega)} + h \int_\Omega \hat{g}(|\nabla u_h|^2) + t\left(\int_\Omega \hat{g}(|\nabla u_h|^2) - \int_\Omega \hat{g}(|\nabla u_{h+t}|^2)\right)
\end{aligned}$$

and therefore, for $t > 0$,

$$\int_\Omega \hat{g}(|\nabla u_{h+t}|^2) - \int_\Omega \hat{g}(|\nabla u_h|^2) \le 0.$$

This shows the monotonicity of the functional $\int_\Omega \hat{g}(|\nabla u_h|^2)$. Using very similar arguments it can be shown that $\|u_h - f_\delta\|^2_{L^2(\Omega)}$ is monotonically increasing in h. q.e.d.

In the following we analyze the behaviour of the functionals $\int_\Omega \hat{g}(|\nabla u_h|^2)$ and $\|u_h - f_\delta\|^2_{L^2(\Omega)}$ for $h \to \infty$.

Lemma 4.7 *Let (4.7), (4.8), (4.9) hold. Then, for $h \to \infty$ the regularized solution converges (with respect to the L^2-norm) to the solution of the optimization problem*

$$\|u - f_\delta\|^2_{L^2(\Omega)} = \min$$

under the constraint

$$\int_\Omega \hat{g}(|\nabla u|^2) = 0 \; .$$

Proof: The proof is similar to the proof in the book of Morozov [30] (p.35) and thus omitted. q.e.d.

In the following lemma we establish the boundedness of the regularized solution. For the proof of this result we utilize Stampacchia's Lemma (see [23]).

Lemma 4.8 *Let B be an open domain, u a function in $H^1(B)$ and a a real number. Then $u^a = \max\{a, u\} \in H^1(B)$ and*

$$\int_B |\nabla u^a|^2 \leq \int_B |\nabla u|^2 \; .$$

We are using this result to prove that each regularized solution lies between the minimal and maximal value of the data f.

Lemma 4.9 *Let (4.7), (4.8), (4.9) hold. Moreover, let*

$$\hat{g} \text{ be monotone in } [0, \infty[. \tag{4.15}$$

If $f \in L^\infty(\Omega)$, then for any $h > 0$ the regularized solution satisfies

$$a := \mathit{ess\ inf}\{f(x) : x \in \Omega\} \leq u_h \leq \mathit{ess\ sup}\{f(x) : x \in \Omega\} =: b \; . \tag{4.16}$$

Proof: We verify that the maximum of u_h is less than b. The corresponding assertion for the minimum values can be proven analogously. Let $u_h^b = \min\{b, u_h\}$, then from Lemma 4.8 and the assumption (4.15) it follows that

$$\int_\Omega \hat{g}(|\nabla u_h|^2) \geq \int_\Omega \hat{g}(|\nabla u_h^b|^2) \; .$$

Since

$$\|u_h - f^\delta\|^2_{L^2(\Omega)} \geq \|u_h^b - f^\delta\|^2_{L^2(\Omega)}$$

it follows from the definition of a regularized solution that $u_h(x) \leq b$. q.e.d.

Next we establish the announced family of Lyapunov functionals.

Theorem 4.10 *Let $\Omega \subseteq \mathbb{R}^d$, $d = 1, 2, 3$ and let a, b be as in (4.16). Morover, let (4.7), (4.8), (4.9), and (4.15) be satisfied. Suppose that u_h is a solution of (4.6). Then the following properties hold* *(a) (Lyapunov functionals for regularization methods) For all $r \in C^2[a,b]$ with $r'' \geq 0$, the function*

$$V(h) := \phi(u_h) := \int_\Omega r(u_h(x))\,dx \tag{4.17}$$

is a Lyapunov functional for a regularization method: Let

$$Mf_\delta := \frac{1}{|\Omega|}\int_\Omega f_\delta\,dx\ .$$

Then

1. *$\phi(u_h) \geq \phi(Mf_\delta)$ for all $h \geq 0$.*
2. *$V \in C[0,\infty)$, $DV(h) := \int_\Omega r'(u_h)(u_h - u_0) \leq 0$, $V(h) - V(0) \leq 0$ for all $h \geq 0$.*

Moreover, if $r'' > 0$ on $[a,b]$, then $V(h) = \phi(u_h)$ is a strict Lyapunov functional:

3. *$\phi(u_h) = \phi(Mf_\delta)$ if and only if $u_h = Mf_\delta$ on $\overline{\Omega}$ for $h > 0$ and $u_0 = Mf_\delta$ a.e. on $\overline{\Omega}$.*
4. *if $h > 0$, then $DV(h) = 0$ if and only if $u_h = Mf_\delta$ on $\overline{\Omega}$.*
5. *$V(H) = V(0)$ for $H > 0$ if and only if $f = Mf$ a.e. on Ω and $u_h = Mf$ on $\overline{\Omega} \times (0,H]$.*

(b) (Convergence)

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

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

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

Proof:

(a) 1. Since $r \in C^2[a,b]$ with $r'' \geq 0$ on $[a,b]$, we know that r is convex on $[a,b]$. Using the gray level invariance and Jensen's inequality it follows

$$\begin{aligned}
\phi(Mf_\delta) &= \int_\Omega r\left(\tfrac{1}{|\Omega|}\int_\Omega u_h(x)\,dx\right) dy \\
&\leq \int_\Omega \tfrac{1}{|\Omega|}\left(r\int_\Omega u_h(x)\,dx\right) dy \\
&= \int_\Omega r(u_h(x))\,dx \\
&= \phi(u_h)\ .
\end{aligned} \tag{4.18}$$

2. From Lemma 4.5 it follows that $V \in C[0,\infty[$. Setting $v = r'(u_h)$ it follows from (4.13) and (4.8) that

$$\langle u_h - u_0, r'(u_h)\rangle = -h\left\langle g(|\nabla u_h|^2)\nabla u_h, r''(u_h)\nabla u_h\right\rangle . \tag{4.19}$$

The right hand side is negative since r is convex.
We represent $V(h) - V(0)$ in the following way

$$\begin{aligned}
&V(h) - V(0)\\
&\quad= \phi(u_h) - \phi(u_0)\\
&\quad= \textstyle\int_\Omega r(u_h(x)) - r(u_0(x))\,dx\\
&\quad= \textstyle\int_\Omega \int_0^1 r'(u_0(x) + t(u_h(x) - u_0(x)))\,dt(u_h(x) - u_0(x))\,dx\\
&\quad= \textstyle\int_\Omega r'(u_h(x))(u_h(x) - u_0(x))\,dx\\
&\qquad+ \textstyle\int_\Omega \int_0^1 (r'(u_0(x) + t(u_h(x) - u_0(x))) - r'(u_h(x)))\,dt\\
&\qquad\quad (u_h(x) - u_0(x))\,dx\\
&\quad= \textstyle\int_\Omega r'(u_h(x))(u_h(x) - u_0(x))\,dx\\
&\qquad- \textstyle\int_\Omega \int_0^1 \int_0^1 r''(u_h(x) - \tau(1-t)(u_h(x) - u_0(x))\,d\tau\\
&\qquad\quad (1-t)(u_h(x) - u_0(x))^2\,dt\,dx\ .
\end{aligned}$$

From (4.19) and the convexity of r it follows that the last two terms in the above chain of inequalities are negative. Thus the assertion is proved.

3. Let $\phi(u_h) = \phi(Mf^\delta)$. Let us now show that the estimate (4.18) implies that $u_h = \text{const}$ on $\overline{\Omega}$. Suppose that $u_h \neq c$ Since $u_h \in H^1(\Omega)$, there exists a partition $\Omega = \Omega_1 \cup \Omega_2$ with $|\Omega_1|, |\Omega_2| \in (0, |\Omega|)$ and

$$\alpha := \frac{1}{|\Omega_1|}\int_{\Omega_1} u_h\,dx \neq \frac{1}{|\Omega_2|}\int_{\Omega_2} u_h\,dx =: \beta\ .$$

This assertion follows from the Poincare inequality for functions in Sobolev spaces [15]. From the strict convexity of r it follows that

$$\begin{aligned}
r\left(\tfrac{1}{|\Omega|}\textstyle\int_\Omega u_h\,dx\right) &= r\left(\tfrac{|\Omega_1|}{|\Omega|}\alpha + \tfrac{|\Omega_2|}{|\Omega|}\beta\right)\\
&< \tfrac{|\Omega_1|}{|\Omega|} r(\alpha) + \tfrac{|\Omega_2|}{|\Omega|} r(\beta)\\
&\leq \tfrac{1}{|\Omega|}\textstyle\int_{\Omega_1} r(u_h)\,dx + \tfrac{1}{|\Omega|}\int_{\Omega_2} r(u_h)\,dx\\
&= \tfrac{1}{|\Omega|}\textstyle\int_\Omega r(u_h)\,dx
\end{aligned}$$

If we utilize this result in (4.18) we observe that for $h > 0$ $\phi(u_h) = \phi(Mf_\delta)$ implies that $u_h = \text{const}$ on $\overline{\Omega}$. Thanks to the average grey value invariance we finally obtain $u_h = Mf_\delta$ on $\overline{\Omega}$.

We turn to the case $h = 0$. From (1.) and (2.) it follows that

$$\phi(Mf_\delta) \leq \phi(u_h) \leq \phi(u_0) \ .$$

If $\phi(u_0) = \phi(Mf_\delta)$, then for all $\theta > 0$

$$\phi(Mf_\delta) = \phi(u_\theta) \ .$$

Thus we have that for all $\theta > 0$ $u_\theta = Mf$. Using the continuity of u_θ with respect to $\theta \in [0, \infty[$ (cf. Lemma 4.5) the assertion follows.

4. The proof is analogous to the proof of the (iv)-assertion in Theorem 3 in [50].
5. Suppose that $V(H) = V(0)$, then from (2.) it follows that

$$V(h) = \text{const on } [0, H] \ .$$

Let $\varepsilon > 0$. Then for any $h \in [\varepsilon, H]$ it follows from (4.) that $u_h = Mf_\delta$. Using the continuity of u_h with respect to $h \in [0, \infty[$ (cf. Lemma 4.5) the assertion follows. The converse direction is obvious.

(b) From Lemma 4.7 and assumption 4.9 it follows that

$$\int_\Omega |\nabla u_h|^2 \to 0 \text{ and } \|u_h - Mf_\delta\|^2_{L^2(\Omega)} \to 0 \ .$$

This shows that

$$\|u_h - Mf_\delta\|_{H^1(\Omega)} \to 0 \ .$$

From the Sobolev embedding theorem it follows in particular that for $h \to \infty$

d=1: u_h converges uniformly to Mf

d=2: $\|u_h - Mf_\delta\|^2_{L^p(\Omega)} \to 0$ for any $1 \leq p < \infty$ (note that we assumed that Ω is bounded domain)

d=3: $\|u_h - Mf_\delta\|^2_{L^p(\Omega)} \to 0$ for any $1 \leq p \leq 6$ (note that we assumed that Ω is bounded domain).

q.e.d.

In Theorem 4.10 we obtained similar results as for Lyapunov functional of diffusion operators (see [50]). In (2.) of Theorem 4.10 the difference of 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 in the whole paper. It is therefore natural that the role of the time derivative in diffusion filtering is replaced by the time discrete approximation around 0.

Example 4.11 *In this example we study different regularization techniques which have been used for denoising of images:*

1. *Tikhonov regularization: Here we have* $\hat{g}(|u|^2) = |u|^2$. *In this case the assumptions (4.7), (4.8), (4.9) and (4.15) are satisfied.*

2. *total variation Regularization: Here we have* $\hat{g}(|u|^2) = \sqrt{|u|^2)}$. *In this case the assumption (4.9) is not satisfied.*

   However, for the modified versions, proposed by Ito and Kunisch [27], where the functional is replaced by

   $$\hat{g}(|u|^2) = \sqrt{|u|^2} + \alpha |u|^2, \text{ with } \alpha > 0$$

   (4.7), (4.8), (4.9), and (4.15) are satisfied.

   For the functional [1, 9]

   $$\hat{g}(|u|^2) = \sqrt{|u|^2 + \beta^2}$$

   the assumption (4.9) is not satisfied. For the modified version

   $$\hat{g}(|u|^2) = \sqrt{|u|^2 + \beta^2} + \alpha |u|^2$$

   studied in [33], the assumptions (4.7), (4.8), (4.9), and (4.15) are satisfied.

   For the functional

   $$\hat{g}(|s|^2) \quad = \quad \begin{cases} \frac{1}{2\varepsilon}|s|^2 & |s| \leq \varepsilon \\ |s| - \frac{\varepsilon}{2} & \varepsilon \leq |s| \leq \frac{1}{\varepsilon} \\ \frac{\varepsilon}{2}|s|^2 + \frac{1}{2}\left(\frac{1}{\varepsilon} - \varepsilon\right) & |s| > \frac{1}{\varepsilon} \end{cases}$$

   the assumptions (4.7), (4.8), (4.9), and (4.15) are satisfied. This method has been proposed by Geman and Yang [17] and was studied extensively by Chambolle and Lions [8] (see also [33]).

3. *Convex Nonquadratic Regularizations: The functional used by Schnörr [41]*

   $$\hat{g}(|s|^2) \quad = \quad \begin{cases} \lambda_h^2 |s|^2 & |s| \leq c_\rho \\ \lambda_l^2 |s|^2 + (\lambda_h^2 - \lambda_l^2) c_\rho (2|s| - c_\rho) & |s| > c_\rho \end{cases}$$

   satisfies (4.7), (4.8), (4.9), and (4.15), whereas the Green functional [18]

   $$\hat{g}(|s|^2) = \ln(\cosh(|s|)^2)$$

   violates the assumption (4.9).

5 Experiments

In this section we illustrate some of the previous regularization strategies by applying them to noisy real-world images.

Regularization was implemented by using central finite differences. In the linear case this leads to a linear system of equations with a positive definite system matrix. It was solved iteratively by a Gauß–Seidel algorithm. It is not difficult to establish error bounds for its solution, since the residue can be calculated and the condition number of the matrix may be estimated using Gerschgorin's theorem. The Gauß-Seidel iterations were stopped when the relative error in the Euclidean norm was smaller than 0.0001.

Discretizing stabilized total variation regularization with

$$\hat{g}(x) = \sqrt{\beta^2 + x}$$

leads to a nonlinear system of equations. It was numerically solved for $\beta = 0.1$ by combining convergent fixed point iterations as outer iterations [13] with inner iterations using the Gauß-Seidel algorithm for solving the linear system of equations. The fixed point iteration turned out to converge quite rapidly, such that not more than 20 iterations were necessary.

Figure 5.1 shows three common test images and a noisy variant of each of them: an outdoor scene with a camera, a magnetic resonance (MR) image of a human head, and an indoor scene. Gaussian noise with zero mean has been added. Its variance was chosen to be a quarter, equal and four times the image variance, respectively, leading to signal-to-noise (SNR) ratios of 4, 1, and 0.25.

The goal of our evaluation was to find out which regularization leads to restorations which are closest to the original images. We applied linear and total variation regularization to the three noisy test images, used 1, 4, and 16 regularization steps and varied the regularization parameter until the optimal restoration was found. The distance to the original image was computed using the Euclidean norm. The results are shown in Table 1, as well as in Figs. 5.2 and 5.3. This gives rise to the following conclusions:

- In all cases, total variation regularization performed better than Tikhonov regularization. As expected, total variation regularization leads to visually sharper edges. The TV-restored images consist of piecewise almost constant patches.

- In the linear case, iterated Tikhonov regularization produced better restorations than noniterated. Visually, noniterated regularization resulted in images with more high-frequent fluctuations. This is in complete agreement with the theoretical considerations in our paper. Improvements caused by iterating the regularization were mainly seen between 1 and 4 iterations. Increasing the iteration number to 16 did hardly lead to further improvements, in one case the results were even slightly worse.

- It appears that the theoretical and experimental results in the linear setting do not carry over to the nonlinear case with total variation regularization: TV regularization was extremely robust: different iteration numbers gave similar results, and the optimal total regularization parameter did not depend much on the iteration number. Thus, in practice one should give the preference to the faster method. In our case iterated regularization was slightly more efficient, since it led to matrices with smaller condition numbers and the Gauß–Seidel algorithm converged faster. Using for instance multigrid methods, which solve the linear systems with a constant effort for all condition numbers, would make noniterated total variation regularization favourable.

In a final experiment we juxtapose the regularizations (3.20) and (3.21) of the Perona–Malik filter. Both processes have been implemented using an explicit finite difference scheme. The results using the MR image from Figure 5.1(c) are shown in Figure 5.4, where different values for γ, the standard deviation of the Gaussian, have been used. For small values of γ, both filters produce rather similar results, while larger values lead to a completely different behaviour. For (3.20), the regularization smoothes the diffusive flux, so that it becomes close to 0 everywhere, and the image remains unaltered. The regularization in (3.21), however, creates a diffusivity which gets closer to 1 for all image locations, so that the filter creates blurry results resembling linear diffusion filtering.

6 Summary

The goal of this paper was to investigate connections between regularization theory and the framework of diffusion filtering. The regularization methods we considered were Tikhonov regularization, total variation regularization, and we focused on linear diffusion filters as well as regularizations of the nonlinear diffusion filter of Perona and Malik. We have established the following results:

- We analyzed the restoration properties of iterated and noniterated regularization both theoretically and experimentally. While linear regularization can be improved by iteration, there is no clear evidence that this is also the case in the nonlinear setting.

- We introduced an alternative regularization of the Perona–Malik filter. In contrast to previous regularization, it allows a formulation as a minimizer of a suitable energy functional.

- We have established Lyapunov functionals and convergence results for regularization methods using a similar theory as for nonlinear diffusion filtering.

These results can be regarded as contributions towards a deeper understanding as well as a better justification of both paradigms. It appears interesting to investigate the following topics in the future:

Table 1: Best restoration results for the different methods and images. The total regularization parameter for N iterations with parameter h is denoted $t = Nh$, and the distance describes the average Euclidean distance per pixel between the restored and the original image without noise.

image	regularization	t	distance
camera	linear, 1 iteration	0.82	15.41
camera	linear, 4 iterations	0.54	15.06
camera	linear, 16 iterations	0.48	15.02
MR	linear, 1 iteration	2.05	23.09
MR	linear, 4 iterations	1.16	22.62
MR	linear, 16 iterations	1.02	22.64
office	linear, 1 iteration	5.7	31.76
office	linear, 4 iterations	3.3	30.47
office	linear, 16 iterations	2.9	30.45
camera	TV, 1 iteration	13.2	11.92
camera	TV, 4 iterations	12.8	12.10
camera	TV, 16 iterations	12.4	12.19
MR	TV, 1 iteration	33.75	20.39
MR	TV, 4 iterations	33.5	20.52
MR	TV, 16 iterations	33	20.65
office	TV, 1 iteration	102	28.66
office	TV, 4 iterations	104	27.99
office	TV, 16 iterations	106	28.05

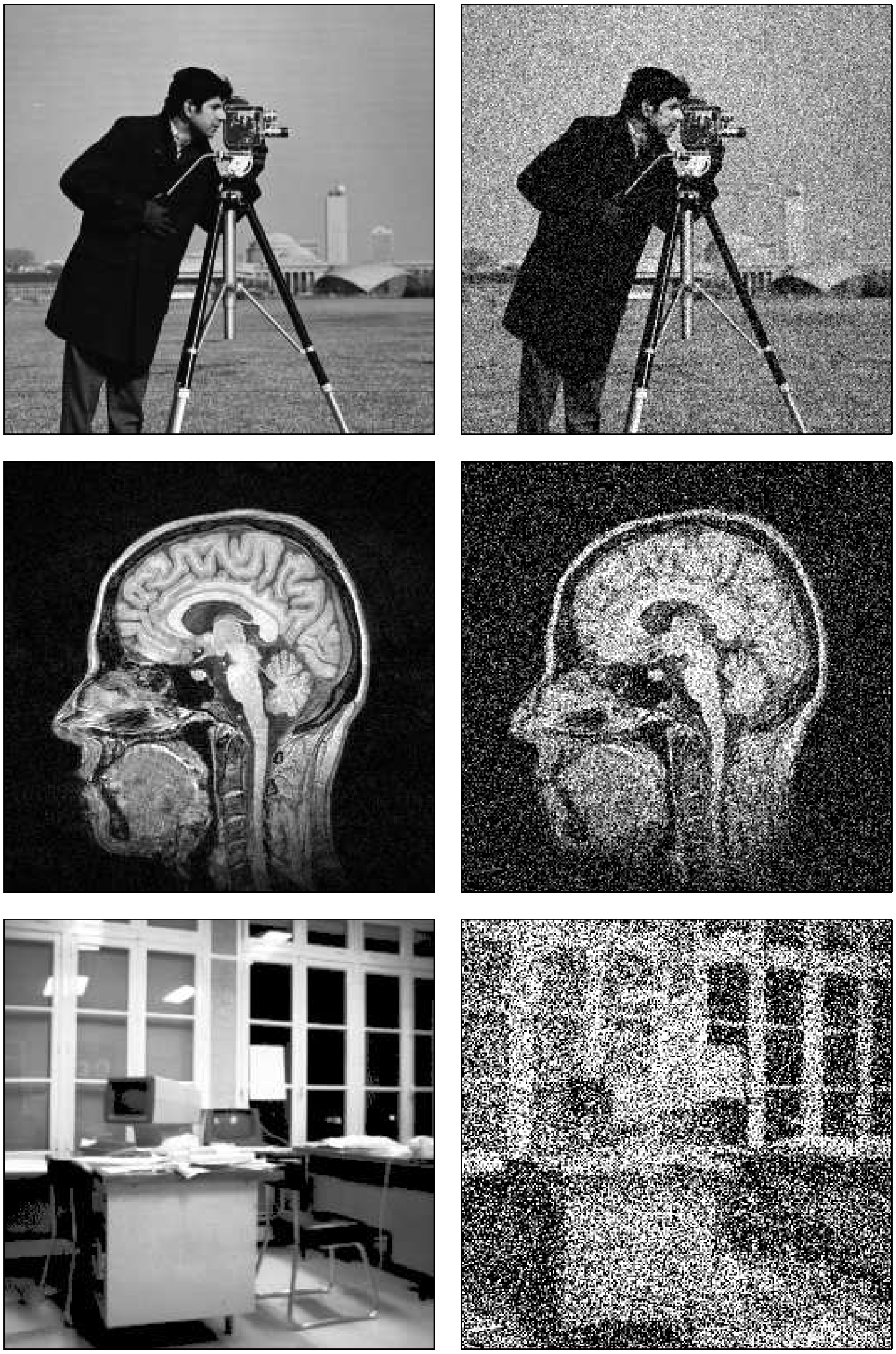

Figure 5.1: Test images, $\Omega = (0, 256)^2$. (A) Top Left: Camera scene. (B) Top Right: Gaussian noise added, SNR=4. (C) Middle Left: Magnetic resonance image. (D) Middle Right: Gaussian noise added, SNR=1. (E) Bottom Left: Office scene. (F) Bottom Right: Gaussian noise added, SNR=0.25.

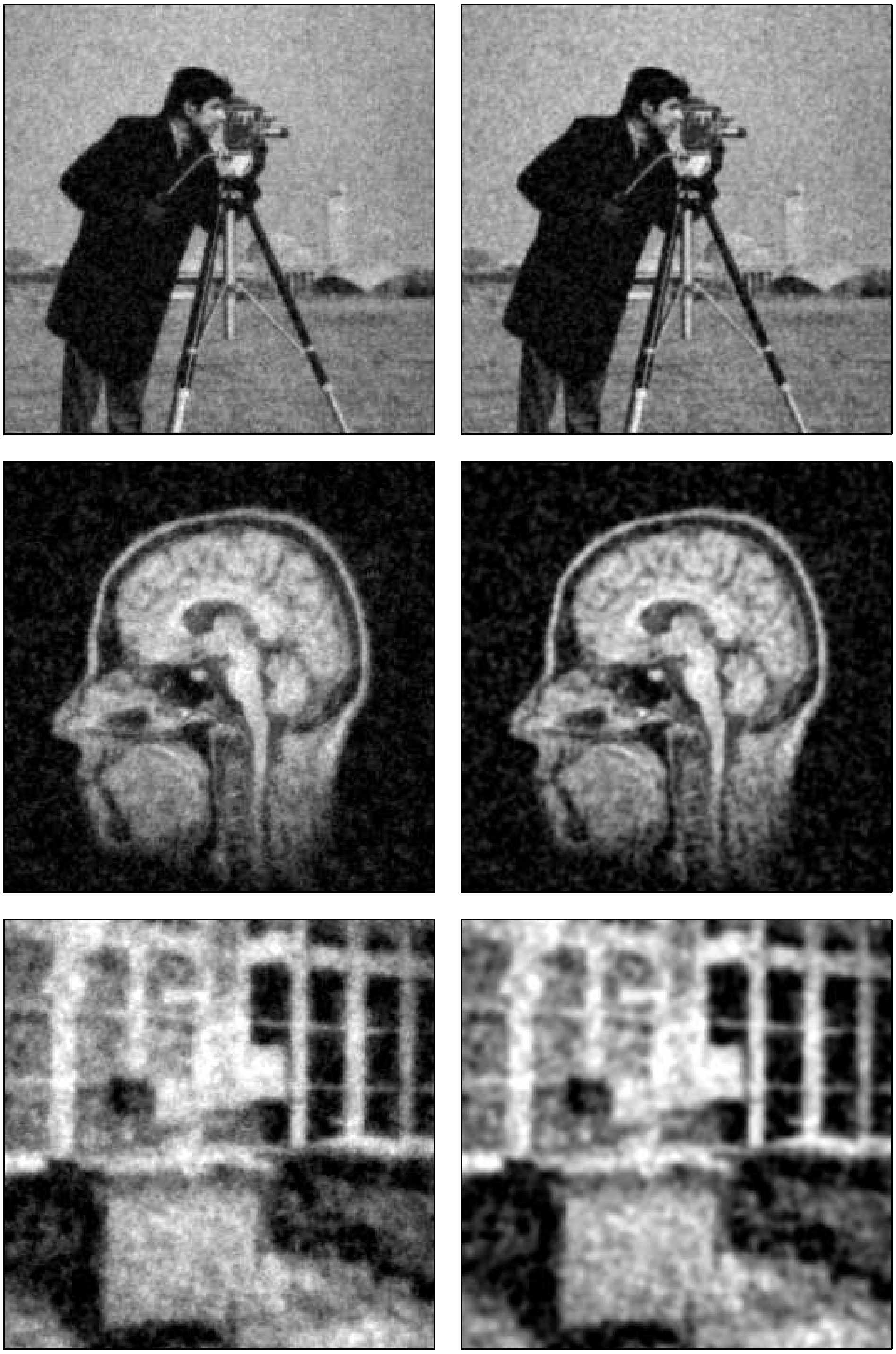

Figure 5.2: Optimal restoration results for Tikhonov regularization. (A) TOP LEFT: Camera, 1 iteration. (B) TOP RIGHT: Camera, 16 iterations. (C) MIDDLE LEFT: MR image, 1 iteration. (D) MIDDLE RIGHT: MR image, 16 iterations. (E) BOTTOM LEFT: Office, 1 iteration. (F) BOTTOM RIGHT: Office, 16 iterations.

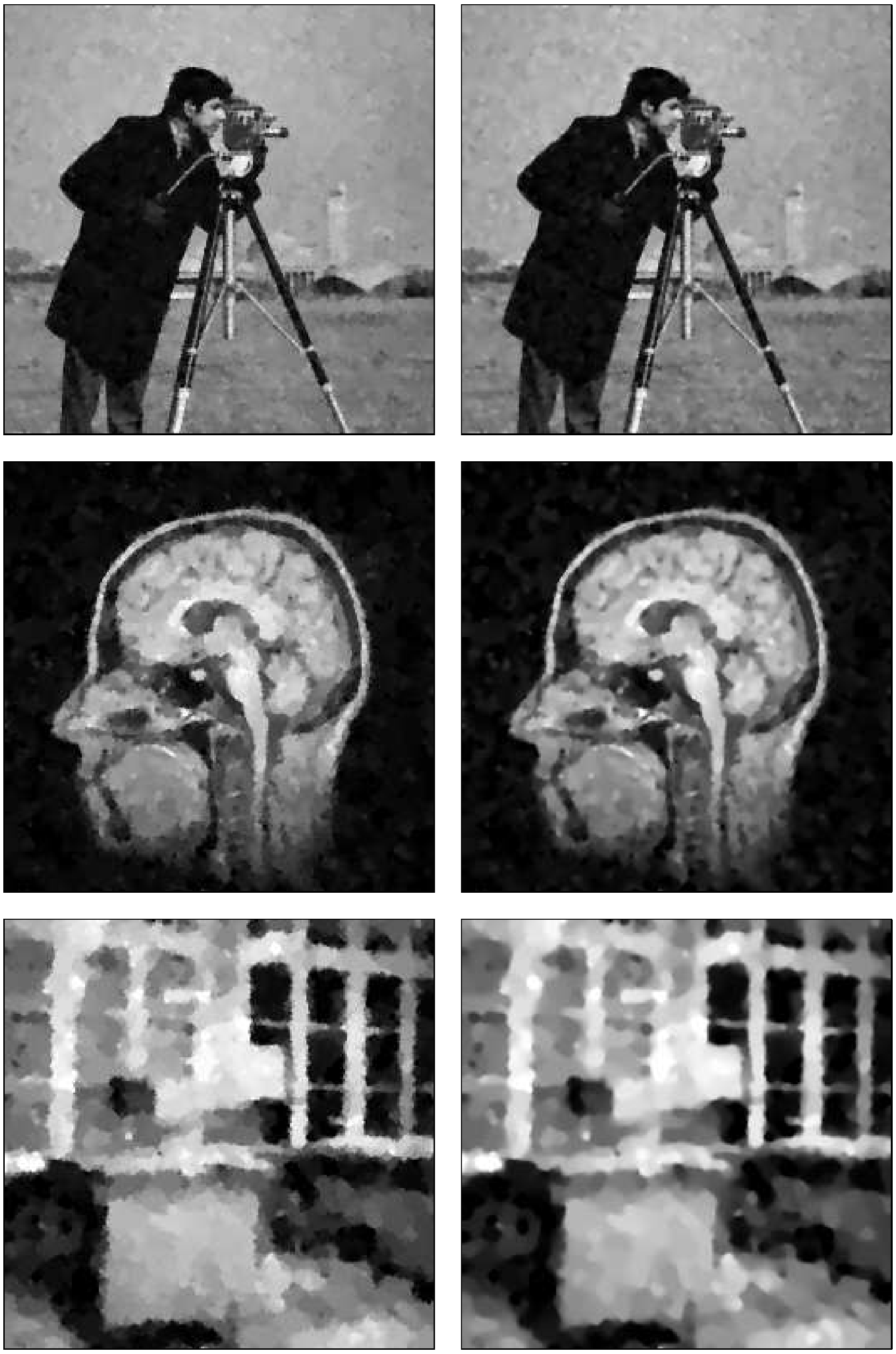

Figure 5.3: Optimal restoration results for total variation regularization. (a) Top Left: Camera, 1 iteration. (b) Top Right: Camera, 16 iterations. (c) Middle Left: MR image, 1 iteration. (d) Middle Right: MR image, 16 iterations. (e) Bottom Left: Office, 1 iteration. (f) Bottom Right: Office, 16 iterations.

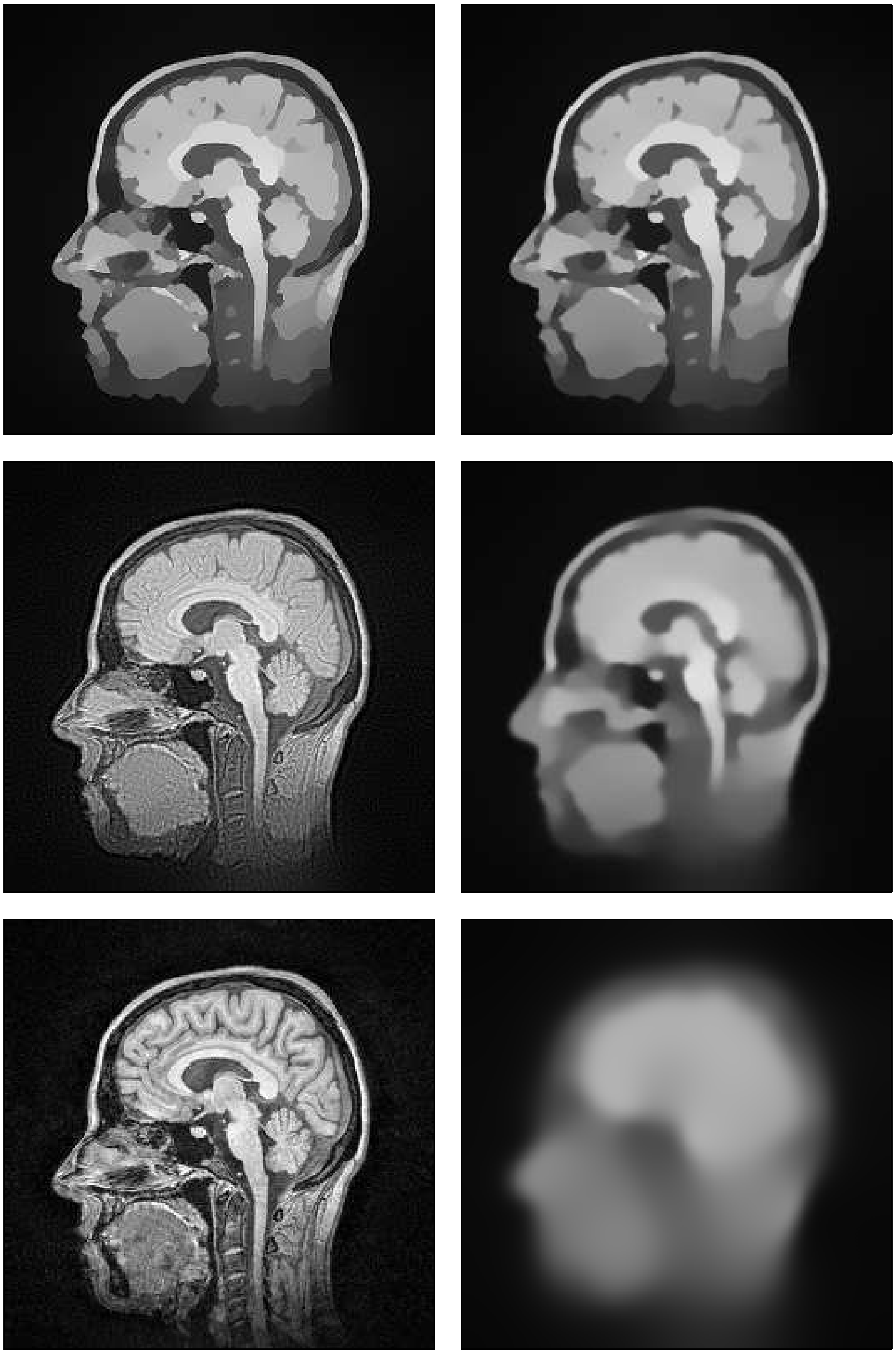

Figure 5.4: Comparison of two regularizations of the Perona–Malik filter ($t = 250$). (A) TOP LEFT: Filter (3.20), $\gamma = 0.5$. (B) TOP RIGHT: Filter (3.21), $\gamma = 0.5$. (C) MIDDLE LEFT: Filter (3.20), $\gamma = 2$. (D) MIDDLE RIGHT: Filter (3.21), $\gamma = 2$. (E) BOTTOM LEFT: Filter (3.20), $\gamma = 8$. (F) BOTTOM RIGHT: Filter (3.21), $\gamma = 8$.

- Regularization scale-spaces. So far, scale-space theory was mainly expressed in terms of parabolic and hyperbolic partial differential equations. Since scale-space methods have contributed to various interesting computer vision applications, it seems promising to investigate similar applications for regularization methods.
- Fully implicit methods for nonlinear diffusion filters using a single time step. This is equivalent to regularization and may be highly useful, if fast numerical techniques for solving the arising nonlinear systems of equations are applied.

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] Barenblatt, G.I.; Bertsch, M.; Dal Passo, R.; Ughi, M.: *A degenerate pseudoparabolic regularization of a nonlinear forward–backward heat equation arising in the theory of heat and mass exchange in stably stratified turbulent shear flow*, SIAM J. Math. Anal., 24, 1414–1439, 1993

[4] Bertero, M., Poggio, T.A.; Torre, V.: *Ill-posed problems in early vision*, Proc. IEEE, 76, 869–889, 1988

[5] 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

[6] 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

[7] Chan, T.F.; Wong, C.K.: *Total variation blind deconvolution*, IEEE Trans. Image Proc., 7, 370–375, 1998

[8] Chambolle, A.; Lions P.-L.: *Image recovery via total variation minimization and related problems*, Numer. Math., 76, 167 – 188, 1995

[9] Charbonnier, P.; Blanc–Féraud, L.; Aubert, G.; Barlaud, M.: *Two deterministic half-quadratic regularization algorithms for computed imaging*, Proc. IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), Vol. 2, IEEE Computer Society Press, Los Alamitos, 168–172, 1994

[10] Dacorogna, B.: *Weak Continuity and Weak Lower Semicontinuity of Non-Linear Functionals*, Lecture Notes in Mathematics, Springer Verlag, Berlin, Heidelberg, New York, 1982

[11] Dacorogna, B.: *Direct Methods in the Calculus of Variations*, Springer Verlag, Berlin, Heidelberg, New York, 1989

[12] Dobson, D.; Scherzer, O.: *Analysis of regularized total variation penalty methods for denoising*, Inverse Problems, 12, 601–617, 1996

[13] Dobson, D.C.; Vogel, C.R.: *Convergence of an iterative method for total variation denoising*, SIAM J. Numer. Anal., 34, 1779–1791, 1997

[14] Engl, H.W.; Hanke, M.; Neubauer, A.: *Regularization of Inverse Problems*, Kluwer, Dordrecht, Boston, London, 1996

[15] Evans, L.C.; Gariepy R.F.: *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton, 1992

[16] Florack, L.: *Image Structure*, Kluwer, Dordrecht, 1997

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

[18] Green, P.J.: *Bayesian reconstructions from emission tomography data using a modified EM algorithm*, IEEE Trans. Medical Imaging, 9, 84–93, 1990.

[19] Groetsch, C.W.: *The Theory of Tikhonov regularization for Fredholm Equations of the First Kind*, Pitman, Boston (1984)

[20] Groetsch, C.W.: *Spectral methods for linear inverse problems with unbounded operators*, J. Approx. Th., 70, 16-28, 1992

[21] Groetsch, C.W.; Scherzer O.: *Optimal order of convergence for stable evaluation of differential operators*, Electronic Journal of Differential Equations (`http://ejde.math.unt.edu`), No. 4: 1-10, 1993

[22] Giusti, E.: *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, Basel, 1984.

[23] Gilbarg, D., Trudinger, N.: *Elliptic partial differential equations of second order 2nd ed.*, Grundlehren der Mathematischen Wissenschaften. 224. Springer, Berlin, 1994

[24] Hanke M.; Groetsch C. W.: *Nonstationary iterated Tikhonov regularization* J. Optim. Theory and Applications, to appear

[25] TER HAAR ROMENY, B.M. (ED.): *Geometry-Driven Diffusion in Computer Vision*, Kluwer, Dordrecht, 1994

[26] HELMBERG, G.: *Introduction to Spectral Theory in Hilbert Space*; Amsterdam, London, 1969

[27] ITO, K.; KUNISCH, K.: *An active set strategy for image restoration based on the augmented Lagrangian formulation*, Preprint No. 410/1994, Fachbereich Mathematik (3), Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, 1994, submitted to SIAM J. Optimization

[28] LI, Y.; SANTOSA, F.: *A computational algorithm for minimizing total variation in image enhancement*, IEEE Trans. Image Proc., 5, 987–995, 1996

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

[30] MOROZOV, V.A.: *Methods for Solving Incorrectly Posed Problems*, Springer Verlag, New York, Berlin, Heidelberg, 1984

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

[32] NITZBERG, M.; SHIOTA, T.: *Nonlinear image filtering with edge and corner enhancement*, IEEE Trans. Pattern Anal. Mach. Intell., 14, 826–833, 1992

[33] NASHED, M.Z.; SCHERZER O.: *Least squares and bounded variation regularization with nondifferentiable functionals*, Numer. Funct. Anal. and Optimiz., 19, 873-901, 1998

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

[35] OSHER, S.; RUDIN L.I.: *Feature-oriented image enhencement using shock filters*, SIAM J. Numer. Anal., 27, 919 - 940, 1990

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

[37] RIESZ, F.; SZ-NAGY, B.: *Functional Analysis*, Ungar, New York, 1955

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

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

[40] Scherzer, O.: *Denoising with higher order derivatives of bounded variation and an application to parameter estimation*, Computing, 60, 1–27, 1998

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

[42] Sporring, J.; Nielsen, M.; Florack, L.; Johansen, P. (Eds.): *Gaussian Scale-Space Theory*, Kluwer, Dordrecht, 1997

[43] 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

[44] Tikhonov, A. N.; Arsenin, V. Y.: *Solutions of Ill–Posed Problems*, John Wiley & Sons, Washington, D.C., 1977, Translation editor Fritz John (English)

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

[46] Vogel, C.R.; Oman, M.E.: *Fast, robust total variation-based reconstruction of noisy, blurred images*, IEEE Trans. Image Proc., 7, 813–824, 1998

[47] Weidmann, J.: *Lineare Operatoren in Hilberträumen*, Teubner, Stuttgart, 1976

[48] Weickert, J.: *Anisotropic diffusion filters for image processing based quality control*, in Fasano, A.; Primicerio, M. (Eds.), Proc. Seventh European Conf. on Mathematics in Industry, Teubner, Stuttgart, 355–362, 1994

[49] Weickert, J.: *A review of nonlinear diffusion filtering*, in ter Haar Romeny, B.; Florack, L.; Koenderink, J.; Viergever, M. (Eds.), Scale-Space Theory in Computer Vision, Lecture Notes in Computer Science, 1252, Springer, Berlin, 3–28, 1997

[50] Weickert, J.: *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, 1998

[51] Weickert, J; Benhamouda, B.: *A semidiscrete nonlinear scale-space theory and its relation to the Perona–Malik paradox*, in Solina, F.; Kropatsch, W.G.; Klette, R.; Bajcsy, R. (Eds.), Advances in Computer Vision, Springer, Wien, 1–10, 1997

[52] Wloka, J: *Partielle Differentialgleichungen*, B.G. Teubner, Stuttgart, 1982