B. ter Haar Romeny, L. Florack, J. Koenderink, M. Viergever (Eds.),
Scale-Space Theory in Computer Vision, Lecture Notes in Computer Science,
Vol. 1252, Springer, Berlin, pp. 3–28, 1997. Invited paper.

A Review of Nonlinear Diffusion Filtering

Joachim Weickert*

Image Sciences Institute,
Utrecht University Hospital,
E.01.334, Heidelberglaan 100,
3584 CX Utrecht, The Netherlands.
E-mail: Joachim.Weickert@cv.ruu.nl

Abstract. This paper gives an overview of scale-space and image enhancement techniques which are based on parabolic partial differential equations in divergence form. In the nonlinear setting this filter class allows to integrate a-priori knowledge into the evolution. We sketch basic ideas behind the different filter models, discuss their theoretical foundations and scale-space properties, discrete aspects, suitable algorithms, generalizations, and applications.

1 Introduction

During the last decade nonlinear diffusion filters have become a powerful and well-founded tool in multiscale image analysis. These models allow to include a-priori knowledge into the scale-space evolution, and they lead to an image simplification which simultaneously preserves or even enhances semantically important information such as edges, lines, or flow-like structures.
Many papers have appeared proposing different models, investigating their theoretical foundations, and describing interesting applications. For a non-expert in this field, however, it is often difficult to appreciate the differences and specific features of each of these filters, or to find the suitable literature for a given problem. The goal of the present paper is to ease these difficulties by giving an overview of the state-of-the art in nonlinear diffusion filtering.
We focus on approaches in divergence form. Together with reflecting or periodic boundary conditions they guarantee that the filtering does not alter the average grey value of an image. This property is useful for scale-space based segmentation algorithm such as the hyperstack [50], and for all areas where the grey value is proportional to some physical quantity, for instance in medical imaging. The restriction to diffusion models means that we will

* Current address: DIKU, Dept. of Computer Science, University of Copenhagen, Universitetsparken 1, 2100 Copenhagen East, Denmark.

not discuss related models based on curve evolution, which cannot be written in divergence form. Also total variation denoising methods and systems of reaction–diffusion equations shall not be treated here. Overviews of these interesting methods can be found in [5,20,33,80].
The paper is organized as follows: Section 2 reviews the physical principles behind diffusion processes. This helps us to understand the differences between linear and nonlinear, and isotropic and anisotropic processes. In Section 3 we discuss well-posedness and scale-space properties of linear diffusion filtering and describe an inhomogeneous variant which combines nonlinear adaptation with linear diffusion filtering. Nonlinear isotropic diffusion filtering is considered in Section 4. We shall study well-posed models with monotone flux functions, discuss ill-posedness aspects of the classical nonlinear diffusion filter of Perona and Malik, and review both spatial and temporal regularizations in order to make diffusion filters with a nonmonotone flux function well-posed. Section 5 is devoted to the study of different nonlinear anisotropic models and their well-posedness and scale-space properties. In Section 6 we sketch generalizations with nontrivial steady-states and extensions to dimensions ≥ 3 and to vector-valued images. The important issue of discrete nonlinear scale-spaces and efficient numerical schemes is addressed in Section 7. Section 8 gives an overview of different applications of nonlinear diffusion filtering, and the paper is concluded with a summary in Section 9.

2 The Physical Background of Diffusion

Most people have an intuitive impression of diffusion as a physical process that equilibrates concentration differences without creating or destroying mass. This physical observation can be easily cast in a mathematical formulation.
The equilibration property is expressed by *Fick's law:*

$$j = -D \cdot \nabla u. \tag{1}$$

This equation states that a concentration gradient ∇u causes a flux j which aims to compensate for this gradient. The relation between ∇u and j is described by the *diffusion tensor* D, a positive definite symmetric matrix. The case where j and ∇u are parallel is called *isotropic*. Then we may replace the diffusion tensor by a positive scalar-valued *diffusivity* g. In the general *anisotropic* case, j and ∇u are not parallel.
The observation that diffusion does only transport mass without destroying it or creating new mass is expressed by the *continuity equation*

$$\partial_t u = -\text{div } j \tag{2}$$

where t denotes the time.

If we plug in Fick's law into the continuity equation we end up with the *diffusion equation*

$$\partial_t u = \operatorname{div}(D \cdot \nabla u). \tag{3}$$

This equation appears in many physical transport processes. In the context of heat transfer it is called *heat equation*.
In image processing we may identify the concentration with the grey value at a certain location. If the diffusion tensor is constant over the whole image domain, one speaks of *homogeneous* diffusion, and a space-dependent filtering is called *inhomogeneous*. Often the diffusion tensor is a function of the differential structure of the evolving image itself. Such a feedback leads to *nonlinear diffusion filters*.
Sometimes the computer vision literature deviates from the preceding notations: It can happen that homogeneous filtering is named isotropic, and inhomogeneous blurring is called anisotropic, even if it uses a scalar-valued diffusivity instead of a diffusion tensor.

3 Linear Diffusion Filtering

3.1 Basic Idea and Well-Posedness

Let us consider a two-dimensional (scalar-valued) image which is given by a continuous bounded mapping $f : \mathbb{R}^2 \to \mathbb{R}$. One of the most widely used methods for smoothing f is to regard it as the initial state of a homogeneous linear diffusion process:

$$\partial_t u = \Delta u, \tag{4}$$

$$u(x,0) = f(x). \tag{5}$$

Its solution is given by the convolution integral

$$u(x,t) = \begin{cases} f(x) & (t = 0) \\ (K_{\sqrt{2t}} * f)(x) & (t > 0) \end{cases} \tag{6}$$

where K_σ denotes a Gaussian with standard deviation σ:

$$K_\sigma(x) := \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{|x|^2}{2\sigma^2}\right). \tag{7}$$

This solution is unique, provided we restrict ourselves to functions satisfying

$$|u(x,t)| \;\leq\; M \cdot \exp(a|x|^2) \qquad (M, a > 0). \tag{8}$$

It depends continuously on the initial image f with respect to $\|\,.\,\|_{L^\infty(\mathbb{R}^2)}$, and it fulfils the maximum–minimum principle

$$\inf_{\mathbb{R}^2} f \;\leq\; u(x,t) \;\leq\; \sup_{\mathbb{R}^2} f \quad \text{on} \quad \mathbb{R}^2 \times [0, \infty). \tag{9}$$

3.2 Scale-Space Properties

Linear diffusion filtering is the oldest and best-studied representative of a *scale-space*. In scale-space theory one embeds an image f into a continuous family $\{T_t f \mid t \geq 0\}$ of gradually smoother versions of it. The original image corresponds to the scale $t = 0$ and increasing the scale should simplify the image without creating spurious structures. Since a scale-space introduces a hierarchy of the image features, it constitutes an important step from a pixel-related image representation to a semantical image description.

Alvarez, Guichard, Lions and Morel [5] have shown that scale-spaces are naturally governed by partial differential equations (PDEs) with the original image as initial condition. The diversity of scale-space approaches has induced people to investigate which of these equations can be distinguished in a unique way from others, because they can be derived from first principles (axioms). Interestingly, imposing linearity restricts the scale-space idea to essentially one representative: convolution with Gaussians of increasing width.

Usually a 1983 paper by Witkin [88] is regarded as the first reference to the linear scale-space idea. Recent work by Weickert, Ishikawa and Imiya [83], however, shows that scale-space is more than 20 years older: An axiomatic derivation of 1-D Gaussian scale-space has already been presented by Taizo Iijima in a Japanese journal paper from 1962 [36]. Iijima derived Gaussian scale-space under five axioms: linearity, translation invariance, scale invariance, semigroup property, and preservation of positivity. This has been the starting point of an entire world of linear scale-space research in Japan, which is basically unknown in the western world.

Until today, more than 10 different axiomatics for Gaussian scale-space exist in the literature; see [83] for an overview. Each of these axiomatics confirms and enhances the evidence that the others give: that Gaussian scale-space is unique within a *linear* framework.

A detailed treatment of the various aspects of Gaussian scale-space theory can be found in [45,24,68] and the references therein.

3.3 Limitations

Figure 1 (a) gives an impression of the temporal evolution under linear diffusion filtering. It depicts an MR slice of the human head. We observe that the image gets more and more simplified and noise and small-scale structures vanish very well. On the other hand, we also observe two typical disadvantages of Gaussian smoothing:

(a) Gaussian smoothing does not only reduce noise, but also blurs important features such as edges and, thus, makes them harder to identify. Since Gaussian scale-space is *designed* to be completely uncommitted, it cannot take into account any a-priori information on structures which are worth being preserved (or even enhanced).

(b) Linear diffusion filtering dislocates edges when moving from finer to coarser scales. So structures which are identified at a coarse scale do not give the right location and have to be traced back to the original image [88]. In practice, relating dislocated information obtained at different scales is difficult and bifurcations may give rise to instabilities. These coarse-to-fine tracking difficulties are generally denoted as the *correspondence problem.*

Due to the uniqueness of the Gaussian scale-space within a linear framework we know that any modification in order to overcome these problems will either renounce linearity or some other scale-space properties. In the following we shall see that a natural way to reduce both problems is to make the diffusion inhomogeneous by steering the process by the geometry of the image itself.

3.4 Inhomogeneous linear diffusion

One of the simplest models for including a-priori knowledge and for reducing the correspondence problem is inhomogeneous linear diffusion filtering.[1]
Suppose we are interested in a smoothing process which reduces smoothing at edges in order to preserve their contrast and location in a better way than Gaussian scale-space. We can use $|\nabla f|$ as a fuzzy edge detector: locations with large $|\nabla f|$ have a higher likelihood to be an edge. Hence, one can reduce the diffusivity for larger values of $|\nabla f|$, for instance by setting [13]

$$g(|\nabla f|^2) := \frac{1}{\sqrt{1+|\nabla f|^2/\lambda^2}} \qquad (\lambda > 0). \tag{10}$$

Although this adaptation is nonlinear, the diffusion equation remains linear:

$$\partial_t u = \operatorname{div}\,(g(|\nabla f|^2)\ \nabla u). \tag{11}$$

Such an adaptation of the diffusion process to the original image has been studied by Fritsch [27]. Figure 1 (b) shows the behaviour of inhomogeneous linear diffusion filtering when being applied to the MR image. Compared with homogeneous linear diffusion, edges remain better localized and their blurring is reduced. On the other hand, for large t the filtered image reveals some artefacts which reflect the differential structure of the initial image.

[1] Another inhomogeneous linear smoothing technique uses Gaussian convolution where the kernel size and/or shape is adapted to the underlying image structure; see e.g. [51]. Such a shape-adapted Gaussian smoothing is no longer equivalent to an inhomogeneous diffusion process, and it does not preserve the average grey level. If one wants to relate shape-adapted Gaussian smoothing to a PDE, one has to carry out sophisticated scaling limits [51].

4 Nonlinear Isotropic Diffusion Filtering

4.1 Models with Monotone Flux Function

Basic Idea. A natural idea to reduce the before mentioned artefacts of inhomogeneous linear diffusion filtering would be to introduce a feedback in the process by adapting the diffusivity g to the gradient of the actual image $u(x,t)$ instead of the original image $f(x)$. This leads to the nonlinear diffusion equation [54]

$$\partial_t u = \operatorname{div}\,(g(|\nabla u|^2)\ \nabla u). \tag{12}$$

Figure 1 (c) shows how such a nonlinear feedback is useful to increase the edge localization in a significant way: Structures remain well-localized as long as they can be recognized. Also blurring at edges is reduced very much. The absolute contrast at edges, however, becomes smaller.

Well-Posedness and Scale-Space Properties. If the flux function[2]

$$\Phi(s) := g(|s|^2)s \tag{13}$$

is monotonously increasing in s, then classical mathematical theories such as monotone operators [11] and differential inequalities [72] ensure well-posedness for the corresponding initial value problem with reflecting boundary conditions:

$$\begin{aligned} \partial_t u = \operatorname{div}\,(g(|\nabla u|^2)\ \nabla u) \quad &\text{on} \quad \Omega \times (0,\infty), && (14)\\ u(x,0) = f(x) \quad &\text{on} \quad \Omega, && (15)\\ \partial_n u = 0 \quad &\text{on} \quad \partial\Omega \times (0,\infty). && (16) \end{aligned}$$

Here, Ω denotes a rectangular image domain and ∂_n is the derivative in normal direction. Moreover, the solution $u(x,t)$ satisfies the extremum principle

$$\inf f \ \le\ u(x,t)\ \le\ \sup f \qquad \forall\, x, \forall\, t > 0, \tag{17}$$

and also local extrema are not enhanced. Hummel [35] has shown that under certain conditions the extremum principle is equivalent to the noncreation of new level-crossings. This *causality* property guarantees that in principle one can trace back features from coarse to fine scales by following their isoluminance surface [41]. It constitutes an important smoothing quality in scale-space theory.

[2] The "mathematical flux" Φ and the "physical flux" j have different sign.

Relations to Energy Minimization. An interesting insight into nonlinear diffusion filtering may also be gained by looking at its relation to energy minimization.[3] Let us consider a potential function $\Psi(|\nabla u|)$ whose gradient is given by the flux $\Phi(\nabla u)$:

$$\nabla\Big(\Psi(|\nabla u|)\Big) = \Phi(\nabla u) = g(|\nabla u|^2)\nabla u. \tag{18}$$

Then minimizing the energy functional

$$E(u) := \int_\Omega \Psi(|\nabla u|)\, dx \tag{19}$$

with the descent method leads to

$$\partial_t u = \operatorname{div}\left(g(|\nabla u|^2)\ \nabla u\right) \tag{20}$$

with reflecting boundary conditions. Table 1 gives an overview of some typical diffusivities and their corresponding fluxes.

Table 1. Relations between potentials and diffusivities. Adapted from [10].

method	diffusivity $g(s^2)$	potential $\Psi(s)$	$\Psi(s)$ convex for
linear diffusion [36]	1	$\frac{s^2}{2}$	all s
Charbonnier et al. [13]	$\frac{1}{\sqrt{1+s^2/\lambda^2}}$	$\sqrt{\lambda^4+\lambda^2 s^2}-\lambda^2$	all s
Perona–Malik [56]	$\frac{1}{1+s^2/\lambda^2}$	$\frac{\lambda^2}{2}\log\left(1+\left(\frac{s}{\lambda}\right)^2\right)$	$\lvert s\rvert \leq \lambda$

The potential Ψ is convex whenever the flux Φ is monotonously increasing. Convex energy functionals have exactly one minimum, and this can be found easily by gradient descent. Computationally expensive methods from nonconvex optimization are not necessary and standard finite element approximations are stable [62]. This makes the use of diffusivities leading to convex potentials very attractive. Such diffusivities have been proposed in [10,13,20,62,70]. Usually they are supplemented with an additional reaction term which will be discussed in Section 6.1.

It is not hard to see that diffusion processes with a monotonously increasing flux functions cannot enhance edges [56]. In certain applications this is an undesirable effect. So let us now draw our attention to processes which lead to nonconvex potentials or – equivalently – to nonmonotone fluxes.

[3] Besides the variational interpretation there exist other theoretical frameworks for diffusion filters such as the Markov random field and mean field annealing context [29,30,42,66], and deterministic interactive particle models [49].

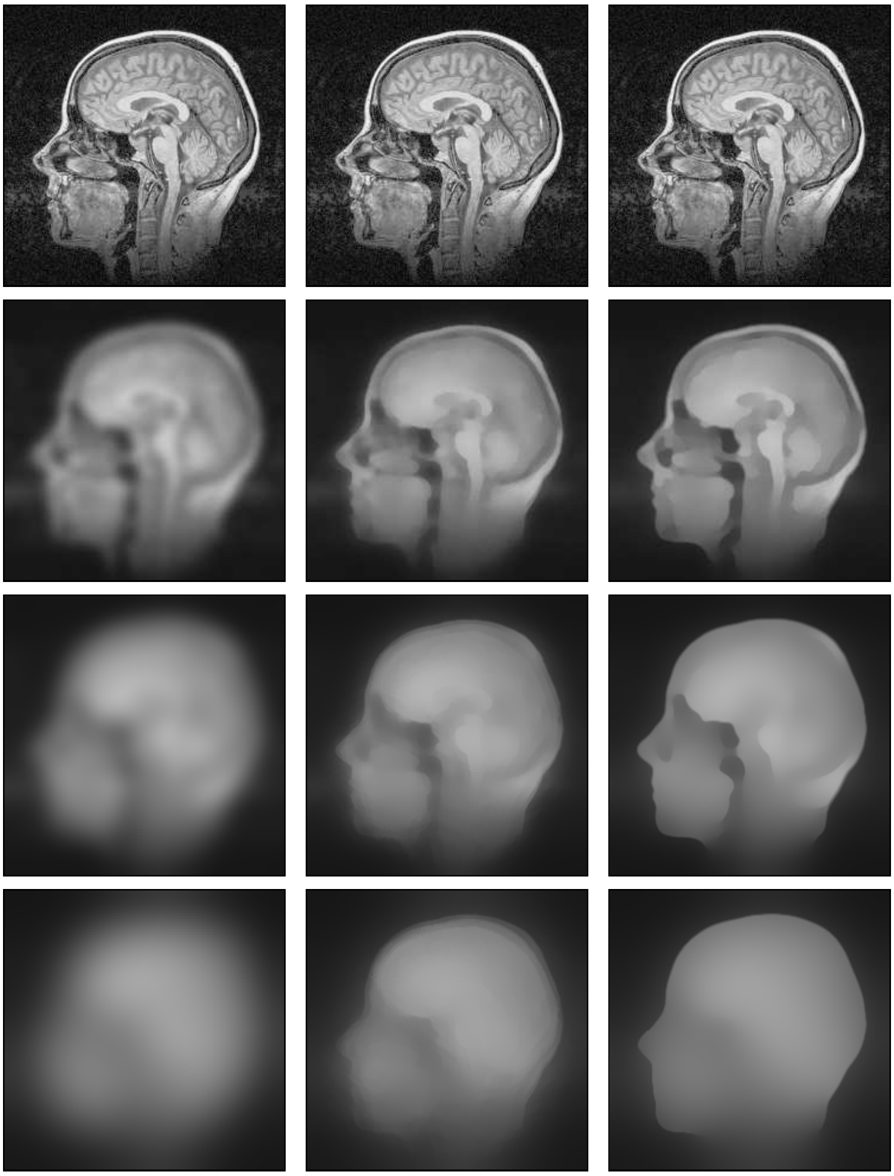

Fig. 1. Diffusion scale-spaces with a convex potential function. TOP: Original image, $\Omega = (0, 236)^2$. (A) LEFT COLUMN: Linear diffusion, top to bottom: $t = 0$, 12.5, 50, 200. (B) MIDDLE COLUMN: Inhomogeneous linear diffusion ($\lambda = 8$), $t = 0$, 70, 200, 600. (C) RIGHT COLUMN: Nonlinear isotropic diffusion with the Charbonnier diffusivity ($\lambda = 3$), $t = 0$, 70, 150, 400.

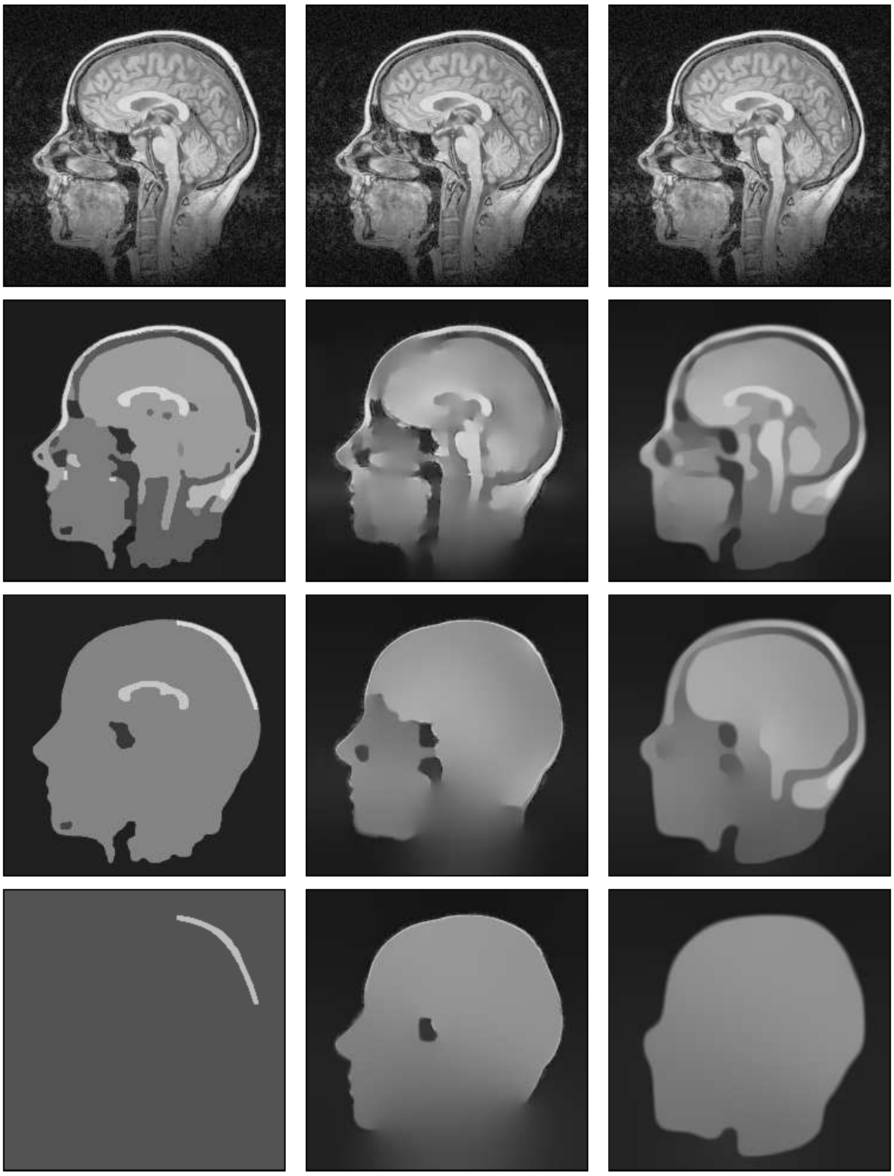

Fig. 2. Nonlinear diffusion scale-spaces with a spatial regularization. TOP: Original image, $\Omega = (0, 236)^2$. (A) LEFT COLUMN: Isotropic nonlinear diffusion ($\lambda = 3$, $\sigma = 1$), $t = 0$, 25000, 500000, 7000000. (B) MIDDLE COLUMN: Isotropic nonlinear diffusion ($\lambda = 3$, $\sigma = 4$), $t = 0$, 40, 400, 1500. (C) RIGHT COLUMN: Edge-enhancing anisotropic diffusion ($\lambda = 3$, $\sigma = 1$), $t = 0$, 250, 875, 3000.

4.2 The Perona–Malik Model

Basic Idea. The historically first nonlinear diffusion filter has been proposed by Perona and Malik in 1987 [54]. They use more rapidly decreasing diffusivities than (10), for instance

$$g(|\nabla u|^2) = \frac{1}{1+|\nabla u|^2/\lambda^2} \qquad (\lambda > 0). \tag{21}$$

The results were visually impressive: edges remained stable over a very long time. It was demonstrated that edge detection based on this process clearly outperforms the linear Canny edge detector, even without applying non-maximum suppression and hysteresis thresholding [56]. However, the Perona–Malik approach reveals some problems which we shall analyse next.

Edge Enhancement and Ill-Posedness. To study the theoretical behaviour of the Perona–Malik filter, let us for simplicity of notation restrict ourselves to the one-dimensional case.
For the diffusivity (21) it follows from Table 1 that the potential function $\Psi(s)$ is only convex for $|s| \le \lambda$. Consequently, the flux function $\Phi(s)$ satisfies $\Phi'(s) \ge 0$ for $|s| \le \lambda$, and $\Phi'(s) < 0$ for $|s| > \lambda$; see Figure 3.

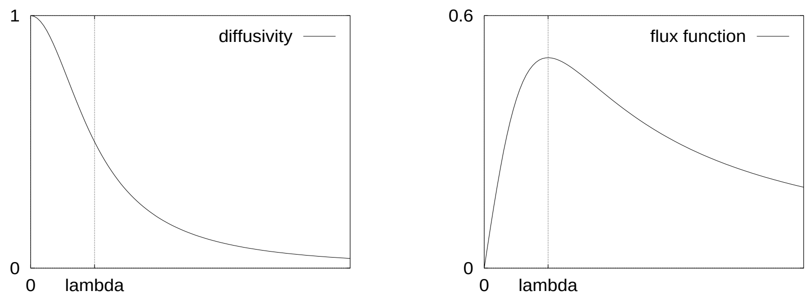

Fig. 3. (a) LEFT: Diffusivity $g(s^2) = \frac{1}{1+s^2/\lambda^2}$. (b) RIGHT: Flux function $\Phi(s) = \frac{s}{1+s^2/\lambda^2}$.

Since (12) can be rewritten as

$$\partial_t u = \partial_x(\Phi(\partial_x u)) = \Phi'(\partial_x u)\, \partial_{xx} u, \tag{22}$$

we observe that – in spite of its nonnegative diffusivity – the Perona–Malik model is of *forward parabolic type* for $|\partial_x u| \le \lambda$, and of *backward parabolic type* for $|\partial_x u| > \lambda$. In the backward region the Perona–Malik equation resembles the backward diffusion equation $\partial_t u = -\partial_{xx} u$, a classical example for an ill-posed equation. In the same way as forward diffusion smoothes contrasts, backward diffusion enhances them. Thus, the Perona–Malik model may sharpen edges, if their gradient is larger than the contrast parameter λ; see [56] for more details.

Unfortunately, for general smooth nonmonotone flux function Φ there is no mathematical theory available which guarantees well-posedness, and a counterexample by Höllig [34] shows that certain diffusion processes with nonmonotone fluxes can have an infinite number of solutions. Recently Kichenassamy [39,40] introduced a notion of generalized solutions to the Perona–Malik process which are piecewise linear and contain jumps. However, one should not expect that a solution of this type is unique or stable with respect to the initial image [12,57,40].

Forward–backward diffusion equations are not as unnatural as they look at first glance: they have been proposed as models for the evolution of stepwise constant temperature or salinity profiles in the ocean, and related equations appear in population dynamics and viscoelasticity; see [8] and the references therein.

In the context of oceanography, numerical experiments were carried out by Posmentier [58] in 1977. Starting from a smoothly increasing initial distribution he reported the creation of perturbations which led to a stepwise constant salinity profile after some time. He also observed instabilities, a first experimental hint to the ill-posedness of this equation. Instabilities were also reported later on by Dzhu Magazieva [21].

In image processing, numerical simulations on the ill-posedness of the one-dimensional Perona–Malik filter were performed by Nitzberg and Shiota [51], Fröhlich and Weickert [28], and Benhamouda [9]. All results point in the same direction: the solution depends strongly on the regularizing effect of the discretization. Finer discretizations are less regularizing and reveal a larger danger of *staircasing effects*, where a smoothed step edge evolves into piecewise linear segments which are separated by jumps. Contributions to the explanation and avoidance of staircasing can be found in [1,9,14,40,81,87].

Interestingly, practical implementations of the Perona–Malik process work often better than one would expect from theory: Staircasing is essentially the only observable instability. A discrete explanation for this so-called *Perona–Malik paradox* [40] has been given by Weickert and Benhamouda [81]. They proved that a standard spatial finite difference discretization is sufficient to turn the Perona–Malik process into a well-posed system of nonlinear ordinary differential equations (ODEs). If an explicit time discretization is applied, then the resulting scheme is *monotonicity preserving* in the 1-D case [9]: a monotone function remains monotone after filtering. Thus, oscillations cannot appear and artefacts are restricted to staircasing.

4.3 Regularized Models with Nonmonotone Flux Functions

Although numerical schemes may provide implicit regularizations which stabilize the process, it seems to be more natural to introduce the regularization directly into the continuous Perona–Malik equation in order to become more independent of the numerical implementation [12,51].
One can apply spatial or temporal regularization (and of course, a combination of both). The following three examples illustrate the variety of possibilities and their tailoring towards a specific task.

(a) The first spatial regularization attempt is probably due to Posmentier who observed numerically the stabilizing effect of averaging the gradient within the diffusivity [58].
A mathematically sound formulation of this idea is given by Catté, Lions, Morel and Coll [12]. By replacing the diffusivity $g(|\nabla u|^2)$ of the Perona–Malik model by $g(|\nabla u_\sigma|^2)$ with $u_\sigma := K_\sigma * u$ they end up with

$$\partial_t u = \operatorname{div}\,(g(|\nabla u_\sigma|^2)\;\nabla u) \tag{23}$$

and establish existence, uniqueness and regularity of a solution to a corresponding initial boundary value problem.
Whitaker and Pizer [87] and Li and Chen [43] have suggested to make the parameters σ and λ time-dependent, and a systematic study of the parameter influence in a 1-D version of (23) has been carried out by Benhamouda [9]. Other spatial regularizations of equations of the Perona–Malik type will be described in Section 5.1.
Spatial regularizations lead to processes where the solution converges to a constant steady-state [75]. This is a desirable property for scale-spaces, since a constant image can be regarded as the simplest and coarsest image representation.
From a practical point of view, spatial regularizations offer the advantage that they make the filter insensitive to noise at scales smaller than σ. Therefore, when regarding (23) as an image restoration equation, it exhibits besides the contrast parameter λ an additional *scale parameter* σ. This avoids a shortcoming of the genuine Perona–Malik process which misinterprets strong oscillations due to noise as edges which should be preserved or even enhanced.

(b) P.-L. Lions proved in a private communication to Mumford that the one-dimensional process

$$\partial_t u = \partial_x\,(g(v)\;\partial_x u), \tag{24}$$

$$\partial_t v = \tfrac{1}{\tau}\,(|\partial_x u|^2 - v) \tag{25}$$

leads to a well-posed filter (cf. [57]). We observe that v is intended as a time-delay regularization of $|\partial_x u|^2$ where the parameter $\tau > 0$ determines the delay. These equations arise as a special case of the spatio-temporal regularizations of Nitzberg and Shiota [51] when neglecting any spatial regularization. Mumford conjectures that this model gives piecewise constant steady-states. In this case, the steady-state solution would solve a segmentation problem.

(c) In the context of shear flows, Barenblatt et al. [8] regularized the one-dimensional forward–backward heat equation by considering the third-order equation

$$\partial_t u = \partial_x(\Phi(u_x)) + \tau \partial_{xt}(\psi(u_x)) \tag{26}$$

where ψ is strictly increasing and uniformly bounded in $\mathbb{R}$, and $|\Phi'(s)| = O(\psi'(s))$ as $s \to \pm\infty$. This regularization was physically motivated by introducing a relaxation time τ into the diffusivity.

For the corresponding initial boundary value problem with homogeneous Neumann boundary conditions they proved the existence of a unique generalized solution. They also showed that smooth solutions may become discontinuous within finite (!) time, before they finally converge to a piecewise constant steady-state.

These examples demonstrate that regularization is much more than stabilizing an ill-posed process: *Regularization is modelling. Appropriately chosen regularizations create the desired filter features.* We observe that spatial regularizations are closer to scale-space ideas while temporal regularization are more related to image restoration and segmentation, since they may lead to nontrivial steady-states.

Figure 2 (a) depicts the temporal evolution under the spatially regularized filter (23). At the chin we observe that this equation is indeed capable of enhancing edges. All structures are extremely well-localized and the results are segmentation-like. On the other hand, also small structures exist over long range of scales if they differ from their vicinity by a sufficiently large contrast. One can try to make this filter faster and more insensitive to small-size structures by increasing the regularizing Gaussian kernel size σ (cf. Fig. 2 (b)), but this also leads to stronger blurring of large structures, and it is no longer possible to enhance the contour of the whole head.

5 Nonlinear Anisotropic Models

All nonlinear diffusion filters that we have investigated so far utilize a scalar-valued diffusivity g which is adapted to the underlying image structure. Therefore, they are isotropic and the flux $j = -g\nabla u$ is always parallel to ∇u. Nevertheless, in certain applications it would be desirable to rotate the flux towards the orientation of interesting features. These requirements cannot be satisfied by a scalar diffusivity anymore, a diffusion tensor leading to anisotropic diffusion filters has to be introduced.

5.1 Spatial Regularizations

Anisotropic diffusion filters often apply spatial regularization strategies. We shall study two examples: The first one offers advantages at noisy edges, and the second one is adapted to the processing of one-dimensional features such as parallel lines and flow-like structures.

Edge-Enhancing Anisotropic Diffusion. In the interior of a segment the nonlinear isotropic diffusion equation (23) behaves almost like the linear diffusion filter (4), but at edges diffusion is inhibited. Therefore, noise at edges cannot be eliminated successfully by this process.
To overcome this problem, a desirable method should prefer diffusion along edges to diffusion perpendicular to them. To this end, we construct the orthonormal system of eigenvectors v_1, v_2 of the diffusion tensor D such that $v_1 \parallel \nabla u_\sigma$ and $v_2 \perp \nabla u_\sigma$. In order to prefer smoothing along the edge to smoothing across it, one can choose the corresponding eigenvalues λ_1 and λ_2 as [76]

$$\lambda_1 := g(|\nabla u_\sigma|^2), \tag{27}$$

$$\lambda_2 := 1. \tag{28}$$

In general, ∇u is not parallel to one of the eigenvectors of D as long as $\sigma > 0$. Hence, this model behaves really anisotropic. If we let the regularization parameter σ tend to 0, we end up with the isotropic Perona–Malik process.
Figure 2 (c) depicts an evolution under edge enhancing anisotropic diffusion. We observe that it creates fairly realistic segments and does not suffer from the problem that small structures are too robust. The larger diffusion along edges, however, creates also slightly stronger rounding of high-curved objects such as the nose.
Another anisotropic model which can be regarded as a regularization of an isotropic nonlinear diffusion filter has been described in [73].

Coherence-Enhancing Anisotropic Diffusion. A second motivation for introducing anisotropy into diffusion processes arises from the wish to process one-dimensional features such as line-like structures. We shall now investigate a modification of a model by Cottet and Germain [18] which is specifically designed for the enhancement of coherent flow-like structures [74].
For this purpose one needs more sophisticated structure descriptors than ∇u_σ. For instance we may use the *structure tensor (second-moment matrix, scatter matrix, interest operator)* [26,60]

$$J_\rho(\nabla u_\sigma) := K_\rho * (\nabla u_\sigma \nabla u_\sigma^T). \tag{29}$$

The componentwise convolution with the Gaussian K_ρ averages orientation information over an integration scale ρ. Since J_ρ is a symmetric positive semidefinite matrix, there exists an orthonormal basis of eigenvectors v_1 and v_2 with corresponding eigenvalues $\mu_1 \geq \mu_2 \geq 0$. The eigenvalues measure the average contrast (grey value variation) in the eigendirections within a scale ρ. Therefore, v_1 is the orientation with the highest grey value fluctuations, and v_2 gives the preferred local orientation, the *coherence direction*. The expression $(\mu_1 - \mu_2)^2$ is a measure of the local coherence. If one wants to enhance coherent structures, one should smooth mainly along the coherence direction v_2 with a diffusivity λ_2 which increases with respect to the coherence $(\mu_1 - \mu_2)^2$. This can be accomplished by designing D such that it possesses the same eigenvectors v_1, v_2 as J_ρ and choosing its corresponding eigenvalues as

$$\lambda_1 := \alpha, \tag{30}$$

$$\lambda_2 := \begin{cases} \alpha & \text{if } \mu_1 = \mu_2, \\ \alpha + (1-\alpha)\exp\left(\frac{-C}{(\mu_1-\mu_2)^{2m}}\right) & \text{else,} \end{cases} \tag{31}$$

where $C > 0$, $m \in \mathbb{N}$, and the small positive parameter $\alpha \in (0,1)$ keeps the diffusion tensor uniformly positive definite.

Fig. 4. Coherence-enhancing anisotropic diffusion of a fingerprint image. (a) LEFT: Original image, $\Omega = (0,256)^2$. (b) RIGHT: Filtered, $\sigma = 0.5$, $\rho = 4$, $t = 20$. From [74].

Figure 4 shows the restoration properties of coherence-enhancing anisotropic diffusion when being applied to a fingerprint image. The diffusion filter encourages smoothing along the coherence orientation v_2 and is therefore well-suited for closing interrupted lines. Due to its reduced diffusivity at noncoherent structures, the locations of the semantically important singularities in the fingerprint remain the same. It should be noted that this filter cannot be regarded as a regularization of a Perona-Malik process. Moreover, a pure local analysis cannot detect interrupted lines. This requires *semilocal* information from the structure tensor which averages orientation information over an integration scale ρ.

Well-Posedness and Scale-Space Properties. Linear diffusion filters, the regularized isotropic filter (23), and the preceding nonlinear anisotropic models can be cast in a general form where the diffusion tensor D is a smooth function of $J_\rho(\nabla u_\sigma)$, which remains uniformly positive definite. Generalizing the reasonings from [12] it follows that problems of this type have a unique solution which is infinitely often differentiable for $t > 0$. Moreover, this solution depends continuously on the original image with respect to the $\mathrm{L}^2(\Omega)$ norm [76,80]. This is of importance when considering stereo images, image sequences or slices from medical CT or MR sequences, since we know that similar images remain similar after filtering.
Interestingly, this filter class – which allows edge enhancement – can be regarded as a smoothing scale-space evolution: It can be proved that the solutions satisfy causality properties in terms of an extremum principle and nonenhancement of local extrema [76,80]. Other interpretations as smoothing transformations are based on Lyapunov functional and will be discussed next.

Lyapunov Functionals. Nonlinear diffusion filters with a spatial regularization by means of Gaussian derivatives cannot be derived as a descent method of a suitable energy functional. However, there are many alternatives: it can be shown [76,80] that for all convex, twice differentiable functions r the expression

$$E(u(t)) := \int_\Omega r(u(x,t))\, dx \tag{32}$$

is a *Lyapunov functional*, if $u(x,t)$ is a solution of the diffusion filter: $E(u(t))$ is decreasing in t, and bounded from below by $\mathrm{vol}(\Omega)\, r(\mu)$, where μ is the average grey value of f.
These Lyapunov functionals allow to prove that the filtered image converges to a constant image as $t \to \infty$. Moreover, they show that – in spite of their image enhancing qualities – the considered filters create smoothing transformations: The special choices $r(s) := |s|^p$, $r(s) := (s-\mu)^{2n}$ and $r(s) := s \ln s$, respectively, imply that all L^p norms with $2 \le p \le \infty$ are decreasing (e.g. the energy $\|u(t)\|^2_{\mathrm{L}^2(\Omega)}$), all even central moments are decreasing (e.g. the variance), 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 [76]. Using Parseval's equality we know that a decreasing energy is equivalent to a decreasing sum of the squared Fourier coefficients. Thus, in spite of the fact that the filters may act image enhancing, their global smoothing properties in terms of Lyapunov functionals can be interpreted in a deterministic, stochastic, information-theory based and Fourier based manner.

Since we know the value of a Lyapunov functional for $t = 0$ and $t = \infty$, we may regard an actual value $E(u(t))$ as a descriptor of the distance to the initial and the steady-state. Prescribing a certain value gives an a-posteriori criterion for the stopping time of the nonlinear diffusion process [84]. The temporal evolution of Lyapunov-like expressions such as the entropy has also been used for selecting the most important scales [37,67], and there are interesting relations between Lyapunov functionals, generalized entropies and fractal theory [69].

5.2 Temporal Regularizations

A temporal regularization for anisotropic diffusion filtering has been proposed by Cottet and El-Ayyadi [17]; see also [16]. Motivated from neural network dynamics they consider a coupled PDE–ODE system of type

$$\frac{\partial u}{\partial t} = \operatorname{div}\,(D\,\nabla u), \tag{33}$$

$$\frac{dD}{dt} = \tfrac{1}{\tau}\,(F(\nabla u) - D) \tag{34}$$

where $\tau > 0$ is some delay parameter, and $F(\nabla u)$ is basically a projection orthogonal to ∇u, if $|\nabla u|$ is larger than some contrast parameter λ. The unit matrix is chosen as the initial value for the diffusion tensor D.
This restoration method combines intraregional isotropic smoothing with anisotropic smoothing along edges. It has been used for edge-preserving denoising, and experiments indicate that it converges to a piecewise constant steady-state. The currently available well-posedness results consist of an existence proof.

6 Generalizations

6.1 Additional Reaction Terms

Similarity Terms. Diffusion filters with a constant steady-state require to specify a stopping time T, if one wants to get nontrivial results. Sometimes it is attempted to circumvent this task by adding an additional reaction term which keeps the steady-state solution close to the original image:

$$\partial_t u = \operatorname{div}\,(g(|\nabla u|^2)\;\nabla u) + \beta(f - u) \qquad (\beta > 0). \tag{35}$$

This equation can be regarded a the descent method of the energy functional

$$E_f(u) := \int_\Omega \left(\tfrac{\beta}{2} \cdot (u - f)^2 \;+\; \Psi(|\nabla u|) \right) dx. \tag{36}$$

In practice, such a modification shifts the problem of specifying a stopping time T to the problem of determining β; so it appears to be a matter of taste

which formulation is preferred. People interested in image restoration usually prefer the reaction term, while for scale-space people it is more natural to have a constant steady-state as the simplest image representation.
From a theoretical viewpoint, it is advantageous to choose a convex potential function Ψ, since this guarantees well-posedness and stable algorithms [10,13,20,62,70]. For nonconvex potentials such as in [52,29] many theoretical questions are open.

Attractors to Certain Grey Values. Cottet and Germain [18] have suggested a reaction term $h(u)$ which attracts the image u to a finite number of specified grey levels. They are chosen as zeros of h. This idea can be used to ease image segmentation and quantization; see also [4]. This is an example where it is explicitly intended to create an energy functional with multiple local minima.

6.2 Higher Dimensions

It is easily seen that many of the previous results can be generalized to higher dimensions. This may be useful when considering e.g. CT or MR image sequences arising from medical applications or when applying diffusion filters to the postprocessing of fluctuating higher-dimensional numerical data. Spatially regularized 3-D nonlinear diffusion filters have been investigated by Gerig et al. [31] in the isotropic case, and by Rambaux and Garçon [59] in the edge-enhancing anisotropic case. A generalization of coherence-enhancing anisotropic diffusion to higher dimensions is proposed in [82].

6.3 Vector-Valued Models

Vector-valued images can arise either from devices measuring multiple physical properties or from a feature analysis of one single image. Examples for the first category are colour images, multi-spectral Landsat exposures and multi-spin echo MR images, whereas representatives of the second class are given by statistical moments or the jet space induced by the image itself and its partial derivatives up to a given order. Feature vectors play an important role for tasks like texture segmentation.
The simplest idea how to apply diffusion filtering to multichannel images would be to diffuse all channels separately and independently from each other. This leads to the undesirable effect that edges may be formed at different locations for each channel. In order to avoid this, one should use a common diffusivity which combines information from all channels. Such isotropic vector-valued diffusion models were studied by Gerig et al. [31] and Whitaker [85] in the context of medical imagery. Extensions to anisotropic vector-valued models with a common tensor-valued structure descriptor for all channels have been studied by Weickert [75,78].

7 Discrete and Numerical Aspects

For nonlinear diffusion filtering numerous numerical methods have been applied:
In [28] three schemes for a one-dimensional regularized nonlinear diffusion filter are compared: a wavelet method of Petrov–Galerkin type, a spectral method and a finite-difference (FD) scheme. It turned out that – especially for large σ – all results were fairly similar. Since the computational effort is of a comparable order of magnitude, it seems to be a matter of taste which scheme is preferred.
Other numerical methods have been applied as well, e.g. finite elements [6,38,62,70]. Bänsch and Mikula reported a significant speed-up by supplementing them with an adaptive mesh coarsening [6]. Neural network approximations to nonlinear diffusion filters are investigated by Cottet [15,16] and Fischl and Schwartz [23]. Perona and Malik [55] propose hardware realizations by means of analogue VLSI networks with nonlinear resistors. A very detailed VLSI proposal has been developed by Gijbels et al. [32].
In most applications of nonlinear diffusion filters, finite differences are preferred, since they are easy to handle and the pixel structure of a real digital image already provides a natural discretization on a fixed rectangular grid. Explicit schemes are very simple to implement and, therefore, they are used almost exclusively. Due to their local behaviour, they are well-suited for parallel architectures. Nevertheless, they suffer from the fact that fairly small time step sizes are needed in order to ensure stability. Semi-implicit schemes (which approximate the diffusivity or the diffusion tensor in an explicit way and the rest implicitly) are considered in [12,80]. They possess much better stability properties. A fast multigrid technique using a pyramid algorithm for the Perona–Malik filter has been studied by Acton et al. [2,1].
While the preceding techniques are focusing on *approximating* a continuous equation, it is often desirable to have a genuinely discrete theory which guarantees that an algorithm *exactly* reveals the same qualitative properties as its continuous counterpart. Such a framework is presented in [77], both for the semidiscrete (discrete in space, continuous in time) and for the fully discrete case. Table 2 gives an overview of the requirements which are needed in order to prove well-posedness, average grey value invariance, causality in terms of an extremum principle and Lyapunov functionals, and convergence to a constant steady-state.
We observe that the requirements belong to five categories: smoothness, symmetry, conservation, nonnegativity and connectivity requirements. These criteria are easy to check for many discretizations. In particular, it turns out that suitable explicit and semi-implicit finite difference discretizations of essentially all before mentioned models except for temporal regularizations create discrete scale-spaces; see Table 3. Moreover, the discrete nonlinear scale-space concept has also led to the development of fast novel schemes, which are based on an additive operator splitting [79]. Under typical accuracy requirements,

Table 2. Requirements for continuous, semidiscrete and fully discrete nonlinear diffusion scale-space. See [77] for more details.

requirement	continuous $\partial_t u = \text{div}\,(D\nabla u)$ $u(t=0) = f$ $\langle D\nabla u, n\rangle = 0$	semidiscrete $\frac{du}{dt} = A(u)u$ $u(0) = f$	discrete $u^0 = f$ $u^{k+1} = Q(u^k)u^k$
smoothness	$D \in \mathrm{C}^\infty$	A Lipschitz-cont.	Q continuous
symmetry	D symmetric	A symmetric	Q symmetric
conservation	div form; refl. b.c.	column sums 0	column sums 1
nonnegativity	pos. semidefinite	nonneg. off-diags.	nonneg. elements
connectivity	uniform pos. def.	irreducible	irred.; pos. diagonal

these *AOS schemes* are about 10 times more efficient than the widely used explicit scheme. A speed-up by another order of magnitude has been achieved by a parallel implementation [84]. This makes nonlinear diffusion filtering attractive for novel application areas: Figure 5 shows a rendering of a large 3-D ultrasound image of a phoetus. It was denoised in less than one minute (8 iterations on four R10000 processors on an SGI Challenge XL).

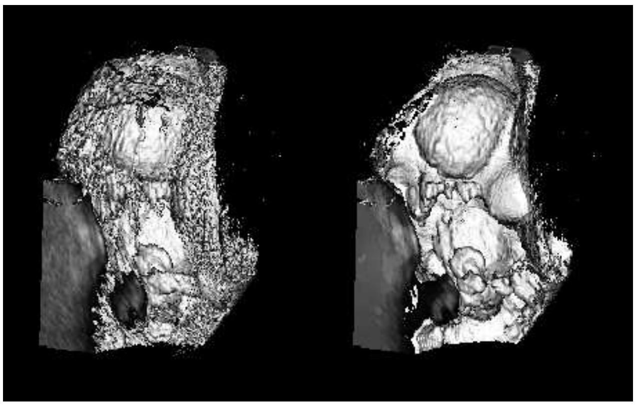

Fig. 5. Rendering of a 3-D ultrasound image of a 10-week old phoetus. LEFT: Original, size $138 \times 208 \times 138$. RIGHT: Filtered with the regularized nonlinear diffusion process (23). From [84].

Table 3. Finite difference schemes which create a discrete nonlinear diffusion scale-space. Upper indices denote the time level, τ is the time step size, h the grid size, m the dimension, and A_l is a discretization of $\partial_{x_l}(g\,\partial_{x_l}u)$.

scheme	formula	stability	costs/iter.	efficiency
explicit	$u^{k+1} = \left(I+\tau \sum_{l=1}^{m} A_l(u^k)\right) u^k$	$\tau < \frac{h^2}{2m}$	very low	low
semi-implicit	$u^{k+1} = \left(I-\tau \sum_{l=1}^{m} A_l(u^k)\right)^{-1} u^k$	$\tau < \infty$	high	fair
AOS	$u^{k+1} = \frac{1}{m} \sum_{l=1}^{m} \left(I-m\tau A_l(u^k)\right)^{-1} u^k$	$\tau < \infty$	low	high

8 Applications

Nonlinear diffusion filters have been applied for postprocessing fluctuating data [76] and for visualizing quality-relevant features in computer aided quality control [73,76,74]. They are useful for improving subsampling [25], for blind image restoration [89], for segmentation of textures [85,86] and remotely sensed data [3,1,2], and for target tracking in infrared images [1]. Most applications, however, are concerned with the filtering of medical images, see e.g. [7,31,44,47,48,65,71].

Besides such specific problem solutions, nonlinear diffusion filters can be found in commercial software packages such as the medical visualization tool Analyze.[4]

Diffusion–reaction approaches have been applied to edge detection [62,70], to the restoration of inverse scattering images [46], to SPECT [13] and 3-D vascular reconstruction in medical imaging [53], and to optic flow [63,19] and stereo problems [61]. They can also be extended to vector-valued images [64] and to corner-preserving smoothing of curves [22].

9 Summary and Conclusions

We have studied possibilities to include a-priori knowledge into the concept of diffusion filtering by a stepwise model refinement: starting with homogeneous linear blurring, we introduced an adaptation leading to inhomogeneous blurring. A feedback in the diffusivity gives nonlinear filters which are well-posed for monotonously increasing fluxes, but they do not allow to enhance edges: Feature enhancement in a well-posed way requires models with spatial or temporal regularizations. While spatial regularizations are closer to scale-space ideas, temporal regularizations are better suited for image restoration. Anisotropic models with a diffusion tensor offer the most degrees of freedom

[4] Analyze is a registered trademark of Mayo Medical Ventures, 200 First Street SW, Rochester, MN 55905, U.S.A.

in the modelling: even the enhancement of coherent structures becomes possibles within a scale-space evolution. Regularizations in nonlinear diffusion filters are far more than a stabilization strategy: they become part of the model, and they are a natural way for integrating semilocal or even global information into the evolution.
It is remarkable that for most models many theoretical results regarding well-posedness and smoothing scale-space properties have been established, and that feature enhancement and and scale-space smoothing is not really contradictive.
Efficient well-founded algorithms build a bridge between the flexible and mathematical sound theoretical basis of nonlinear diffusion filtering and a lot of challenging real-world problems in image processing and computer vision.

References

1. S.T. Acton, *Edge enhancement of infrared imagery by way of the anisotropic diffusion pyramid*, Proc. IEEE Int. Conf. Image Processing (ICIP–96, Lausanne, Sept. 16–19, 1996), Vol. 1, 865–868, 1996.
2. S.T. Acton, A.C. Bovik, M.M. Crawford, *Anisotropic diffusion pyramids for image segmentation*, Proc. IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), Vol. 3, IEEE Computer Society Press, Los Alamitos, 478–482, 1994.
3. S.T. Acton, M.M. Crawford, *A mean field solution to anisotropic edge detection of remotely sensed data*, Proc. 12th Int. Geoscience and Remote Sensing Symposium (IGARSS '92, Houston, May 26–29, 1992), Vol. 2, 845–847, 1992.
4. L. Alvarez, J. Esclarin, *Image quantization by nonlinear smoothing*, L.I. Rudin, S.K. Bramble (Eds.), Investigative and trial image processing, SPIE Vol. 2567, 182–192, 1995.
5. L. Alvarez, F. Guichard, P.-L. Lions, J.-M. Morel, *Axioms and fundamental equations in image processing*, Arch. Rational Mech. Anal., Vol. 123, 199–257, 1993.
6. E. Bänsch, K. Mikula, *A coarsening finite element strategy in image selective smoothing*, Preprint No. 18/1996, Faculty of Mathematics, University of Freiburg, Hebelstr. 27, 79104 Freiburg, Germany, 1996; submitted to Computation and Visualization in Sciences.
7. I. Bajla, M. Marušiak, M. Šrámek, *Anisotropic filtering of MRI data based upon image gradient histogram*, D. Chetverikov, W.G. Kropatsch (Eds.), Computer analysis of images and patterns, Lecture Notes in Comp. Science, Vol. 719, Springer, Berlin, 90–97, 1993.
8. G.I. Barenblatt, M. Bertsch, R. Dal Passo, M. Ughi, *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., Vol. 24, 1414–1439, 1993.
9. B. Benhamouda, *Parameter adaptation for nonlinear diffusion in image processing*, master thesis, Dept. of Mathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1994.

10. L. Blanc–Féraud, P. Charbonnier, G. Aubert, M. Barlaud, *Nonlinear image processing: modelling and fast algorithm for regularization with edge detection*, Proc. IEEE Int. Conf. Image Processing (ICIP–95, Washington, Oct. 22–25, 1995), Vol. 1, 474–477, 1995.
11. H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North Holland, Amsterdam, 1973.
12. F. Catté, P.-L. Lions, J.-M. Morel, T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal., Vol. 29, 182–193, 1992.
13. P. Charbonnier, L. Blanc–Féraud, G. Aubert, M. Barlaud, *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.
14. G. Cong, S.D. Ma, *Nonlinear diffusion for early vision*, Proc. 13th Int. Conf. Pattern Recognition (ICPR 13, Vienna, Aug. 25–30, 1996), Vol. A, 403–406, 1996.
15. G.-H. Cottet, *Diffusion approximation on neural networks and applications for image processing*, F. Hodnett (Ed.), Proc. Sixth European Conf. on Mathematics in Industry, Teubner, Stuttgart, 3–9, 1992.
16. G.-H. Cottet, *Neural networks: Continuous approach and applications to image processing*, J. Biological Systems, Vol. 3, 1131–1139, 1995.
17. G.-H. Cottet, M. El Ayyadi, *Nonlinear PDE operators with memory terms for image processing*, Proc. IEEE Int. Conf. Image Processing (ICIP–96, Lausanne, Sept. 16–19, 1996), Vol. 1, 481–483, 1996.
18. G.-H. Cottet, L. Germain, *Image processing through reaction combined with nonlinear diffusion*, Math. Comp., Vol. 61, 659–673, 1993.
19. R. Deriche, P. Kornprobst, G. Aubert, *Optical-flow estimation while preserving its discontinuities: A variational approach*, Proc. Second Asian Conf. Computer Vision (ACCV '95, Singapore, December 5–8, 1995), Vol. 2, 290–295, 1995.
20. R. Deriche, O. Faugeras, *Les EDP en traitement des images et vision par ordinateur*, Traitement du Signal, in press.
21. S.K. Dzhu Magazieva, *Numerical study of a partial differential equation*, U.S.S.R. Comput. Maths. Math. Phys., Vol. 23, No. 4, 45–49, 1983.
22. P. Fiddelaers, E.J. Pauwels, L.J. Van Gool, *Geometry-driven curve evolution*, J.-O. Eklundh (Ed.), Computer vision – ECCV '94, Lecture Notes in Comp. Science, Vol. 800, Springer, Berlin, 427–432, 1994.
23. B. Fischl, E.L. Schwartz, *Adaptive nonlinear filtering for nonlinear diffusion approximation in image processing*, Proc. 13th Int. Conf. Pattern Recognition (ICPR 13, Vienna, Aug. 25–30, 1996), Vol. D, 276–280, 1996.
24. L.M.J. Florack, *The structure of scalar images*, Kluwer, Dordrecht, 1997, to appear.
25. G.E. Ford, R.R. Estes, H. Chen, *Scale-space analysis for image sampling and interpolation*, Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP–92, San Francisco, March 23–26, 1992), Vol. 3, 165–168, 1992.
26. M.A. Förstner, E. Gülch, *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, 1987.
27. D.S. Fritsch, *A medial description of greyscale image structure by gradient-limited diffusion*, R.A. Robb (Ed.), Visualization in biomedical computing '92, SPIE Vol. 1808, 105–117, 1992.

28. J. Fröhlich, J. Weickert, *Image processing using a wavelet algorithm for nonlinear diffusion*, Report No. 104, Laboratory of Technomathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1994.
29. D. Geiger, A. Yuille, *A common framework for image segmentation*, Int. J. Comput. Vision, Vol. 6, 227–243, 1991.
30. D. Geman, G. Reynolds, *Constrained restoration and the recovery of discontinuities*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, 367–383, 1992.
31. G. Gerig, O. Kübler, R. Kikinis, F.A. Jolesz, *Nonlinear anisotropic filtering of MRI data*, IEEE Trans. Medical Imaging, Vol. 11, 221–232, 1992.
32. T. Gijbels, P. Six, L. Van Gool, F. Catthoor, H. De Man, A. Oosterlinck, *A VLSI-architecture for parallel non-linear diffusion with applications in vision*, Proc. IEEE Workshop on VLSI Signal Processing, Vol. 7, 398–707, 1994.
33. B.M. ter Haar Romeny (Ed.), *Geometry-driven diffusion in computer vision*, Kluwer, Dordrecht, 1994.
34. K. Höllig, *Existence of infinitely many solutions for a forward–backward heat equation*, Trans. Amer. Math. Soc., Vol. 278, 299–316, 1983.
35. R.A. Hummel, *Representations based on zero-crossings in scale space*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '86, Miami Beach, June 22–26, 1986), IEEE Computer Society Press, Washington, 204–209, 1986.
36. T. Iijima, *Basic theory of pattern normalization (for the case of a typical one-dimensional pattern)*, Bulletin of the Electrotechnical Laboratory, Vol. 26, 368–388, 1962 (in Japanese).
37. M. Jägersand, *Saliency maps and attention selection in scale and spatial coordinates: an information theoretic approach*, Proc. Fifth Int. Conf. on Computer Vision (ICCV '95, Cambridge, June 20–23, 1995), IEEE Computer Society Press, Los Alamitos, 195–202, 1995.
38. J. Kačur, K. Mikula, *Solution of nonlinear diffusion appearing in image smoothing and edge detection*, Appl. Num. Math., Vol. 17, 47–59, 1995.
39. S. Kichenassamy, *Nonlinear diffusions and hyperbolic smoothing for edge enhancement*, M.-O. Berger, R. Deriche, I. Herlin, J. Jaffré, J.-M. Morel (Eds.), ICAOS '96: Images, wavelets and PDEs, Lecture Notes in Control and Information Sciences, Vol. 219, Springer, London, 119–124, 1996.
40. S. Kichenassamy, *The Perona–Malik paradox*, SIAM J. Appl. Math., to appear.
41. J.J. Koenderink, *The structure of images*, Biol. Cybern., Vol. 50, 363–370, 1984.
42. S.Z. Li, *On discontinuity-adaptive smoothness priors in computer vision*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 17, 576–586, 1995.
43. X. Li, T. Chen, *Nonlinear diffusion with multiple edginess thresholds*, Pattern Recognition, Vol. 27, 1029–1037, 1994.
44. C. Lamberti, M. Sitta, F. Sgallari, *Improvements to the anisotropic diffusion model for 2-D echo image processing*, Proc. Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Vol. 14, 1872–1873, 1992.
45. T. Lindeberg, *Scale-space theory in computer vision*, Kluwer, Boston, 1994.
46. P. Lobel, L. Blanc-Féraud, C. Pichot, M. Barlaud, *A new regularization scheme for inverse scattering*, Research Note 95-72, I3S Laboratory, University of Nice–Sopia Antipolis, Bât. 4, 250 rue Albert Einstein, Sophia Antipolis, 06560 Valbonne, France, 1996; submitted to Inverse Problems.
47. M.H. Loew, J. Rosenman, J. Chen, *A clinical tool for enhancement of portal images*, M.H. Loew (Ed.), Image processing, SPIE Vol. 2167, 543–550, 1994.

48. D.-S. Luo, M.A. King, S. Glick, *Local geometry variable conductance diffusion for post-reconstruction filtering*, IEEE Trans. Nuclear Sci., Vol. 41, 2800–2806, 1994.
49. J.L. Marroquin, *Deterministic interactive particle models for image processing and computer graphics*, CVGIP: Graphical Models and Image Processing, Vol. 55, 408–417, 1993.
50. W.J. Niessen, K.L. Vincken, J.A. Weickert, M.A. Viergever, *Nonlinear multiscale representations for image segmentation*, Computer Vision and Image Understanding, 1997, in press.
51. M. Nitzberg, T. Shiota, *Nonlinear image filtering with edge and corner enhancement*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, 826–833, 1992.
52. N. Nordström, *Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection*, Image Vision Comput., Vol. 8, 318–327, 1990.
53. E. Payot, R. Guillemaud, Y. Trousset, F. Preteux, *An adaptive and constrained model for 3D X-ray vascular reconstruction*, P. Grangeat, J.-L. Amans (Eds.), Three-dimensional image reconstruction in radiation and nuclear medicine, Kluwer, Dordrecht, 47–57, 1996.
54. P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, Proc. IEEE Comp. Soc. Workshop on Computer Vision (Miami Beach, Nov. 30 – Dec. 2, 1987), IEEE Computer Society Press, Washington, 16–22, 1987.
55. P. Perona, J. Malik, *A network for multiscale image segmentation*, Proc. IEEE Int. Symp. Circuits and Systems (ISCAS–88, Espoo, June 7–9, 1988), 2565–2568, 1988.
56. P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, 629–639, 1990.
57. P. Perona, T. Shiota, J. Malik, *Anisotropic diffusion*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 72–92, 1994.
58. E.S. Posmentier, *The generation of salinity finestructure by vertical diffusion*, J. Phys. Oceanogr., Vol. 7, 298–300, 1977.
59. I. Rambaux, P. Garçon, *Nonlinear anisotropic diffusion filtering of 3D images*, project work, Département Génie Mathématique, INSA de Rouen and Laboratory of Technomathematics, University of Kaiserslautern, 1994.
60. A.R. Rao, B.G. Schunck, *Computing oriented texture fields*, CVGIP: Graphical Models and Image Processing, Vol. 53, 157–185, 1991.
61. L. Robert, R. Deriche, *Dense depth map reconstruction: A minimization and regularization approach which preserves discontinuities*, B. Baxton, R. Cipolla (Eds.), Computer vision – ECCV '96, Volume I, Lecture Notes in Comp. Science, Vol. 1064, Springer, Berlin, 439–451, 1996.
62. C. Schnörr, *Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals*, J. Math. Imag. Vision, Vol. 4, 189–198, 1994.
63. C. Schnörr, *Segmentation of visual motion by minimizing convex non-quadratic functionals*, Proc. 12th Int. Conf. Pattern Recognition (ICPR 12, Jerusalem, Oct. 9–13, 1994), Vol. A, IEEE Computer Society Press, Los Alamitos, 661–663, 1994.
64. C. Schnörr, *Convex variational segmentation of multi-channel images*, M.-O. Berger, R. Deriche, I. Herlin, J. Jaffré, J.-M. Morel (Eds.), ICAOS '96: Images, wavelets and PDEs, Lecture Notes in Control and Information Sciences, Vol. 219, Springer, London, 201–207, 1996.

65. J. Sijbers, P. Scheunders, M. Verhoye, A. Van der Linden, D. Van Dyck, E. Raman, *Watershed-based segmentation of 3D MR data for volume quantization*, Magnetic Resonance Imaging, Vol. 15, No. 4, 1997; to appear.
66. W. Snyder, Y.-S. Han, G. Bilbro, R. Whitaker, S. Pizer, *Image relaxation: Restoration and feature extraction*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 17, 620–624, 1995.
67. J. Sporring, *The entropy of scale-space*, Proc. 13th Int. Conf. Pattern Recognition (ICPR 13, Vienna, Aug. 25–30, 1996), Vol. A, 900–904, 1996.
68. J. Sporring, M. Nielsen, L. Florack, P. Johansen (Eds.), *Gaussian scale-space theory*, Kluwer, Dordrecht, 1997.
69. J. Sporring, J. Weickert, *On generalized entropies and scale-space*, this volume.
70. R. Sprengel, *Entwurf und Analyse nichtlinearer Diffusionsverfahren für die Bildverarbeitung*, Ph.D. thesis, Dept. of Computer Science, University of Hamburg, Germany, 1995.
71. E. Steen, B. Olstad, *Scale-space and boundary detection in ultrasonic imaging using nonlinear signal-adaptive anisotropic diffusion*, M.H. Loew (Ed.), Image processing, SPIE Vol. 2167, 116–127, 1994.
72. W. Walter, *Differential and integral inequalities*, Springer, Berlin, 1970.
73. J. Weickert, *Anisotropic diffusion filters for image processing based quality control*, A. Fasano, M. Primicerio (Eds.), Proc. Seventh European Conf. on Mathematics in Industry, Teubner, Stuttgart, 355–362, 1994.
74. J. Weickert, *Multiscale texture enhancement*, V. Hlaváč, R. Šára (Eds.), Computer analysis of images and patterns, Lecture Notes in Comp. Science, Vol. 970, Springer, Berlin, 230–237, 1995.
75. J. Weickert, *Scale-space properties of nonlinear diffusion filtering with a diffusion tensor*, Report No. 110, Laboratory of Technomathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1994.
76. J. Weickert, *Theoretical foundations of anisotropic diffusion in image processing*, W. Kropatsch, R. Klette, F. Solina (Eds.), Theoretical foundations of computer vision, Computing Suppl. 11, Springer, Wien, 221–236, 1996.
77. J. Weickert, *Nonlinear diffusion scale-spaces: From the continuous to the discrete setting*, M.-O. Berger, R. Deriche, I. Herlin, J. Jaffré, J.-M. Morel (Eds.), ICAOS '96: Images, wavelets and PDEs, Lecture Notes in Control and Information Sciences, Vol. 219, Springer, London, 111–118, 1996.
78. J. Weickert, *Coherence-enhancing diffusion of colour images*, Proc. Seventh National Symposium on Pattern Recognition and Image Analysis (VII NSPRIA, Barcelona, April 21–25, 1997), in press.
79. J. Weickert, *Recursive separable schemes for nonlinear diffusion filters*, this volume.
80. J. Weickert, *Anisotropic diffusion in image processing*, Teubner Verlag, Stuttgart, 1997, to appear.
81. J. Weickert, B. Benhamouda, *A semidiscrete nonlinear scale-space theory and its relation to the Perona–Malik paradox*, F. Solina, W.G. Kropatsch, R. Klette, R. Bajcsy (Eds.), Theoretical foundations of computer vision (TFCV '96, Dagstuhl, March 18–22, 1996), Springer, Wien, 1997, in press.
82. J. Weickert, B.M. ter Haar Romeny, A. Lopez, W.J. van Enk, *Orientation analysis by coherence-enhancing diffusion*, Proc. Symp. Real World Computing (RWC '97, Tokyo, Jan. 29–31, 1997), 96–103, 1997.

83. J. Weickert, S. Ishikawa, A. Imiya, *On the history of Gaussian scale-space axiomatics*, J. Sporring, M. Nielsen, L. Florack, P. Johansen (Eds.), Gaussian scale-space theory, Kluwer, Dordrecht, 1997, in press.
84. J. Weickert, K.J. Zuiderveld, B.M. ter Haar Romeny, W.J. Niessen, *Parallel implementations of AOS schemes: A fast way of nonlinear diffusion filtering*, submitted.
85. R.T. Whitaker, *Geometry limited diffusion in the characterization of geometric patches in images*, CVGIP: Image Understanding, Vol. 57, 111–120, 1993.
86. R. Whitaker, G. Gerig, *Vector-valued diffusion*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 93–134, 1994.
87. R.T. Whitaker, S.M. Pizer, *A multi-scale approach to nonuniform diffusion*, CVGIP: Image Understanding, Vol. 57, 99–110, 1993.
88. A.P. Witkin, *Scale-space filtering*, Proc. Eighth Int. Joint Conf. on Artificial Intelligence (IJCAI '83, Karlsruhe, Aug. 8–12, 1983), Vol. 2, 1019–1022, 1983.
89. Y.-L. You, M. Kaveh, *Anisotropic blind image restoration*, Proc. IEEE Int. Conf. Image Processing (ICIP–96, Lausanne, Sept. 16–19, 1996), Vol. 2, 461–464, 1996.