Technical Report DIKU-TR-97/22
Department of Computer Science
University of Copenhagen
Universitetsparken 1
2100 Copenhagen East
Denmark

submitted for publication
August 1997

Why the Perona–Malik Filter Works

Joachim Weickert

Department of Computer Science, University of Copenhagen,
Universitetsparken 1, 2100 Copenhagen East, Denmark
E-mail: joachim@diku.dk

Brahim Benhamouda

Faculty of Mathematics, Technical University Chemnitz–Zwickau,
Reichenhainer Str. 41, 09107 Chemnitz, Germany
E-mail: brahim@mathematik.tu-chemnitz.de

Abstract

Although the widely-used Perona–Malik filter is regarded as ill-posed, straightforward implementations are often surprisingly stable. We give an explanation for this effect by applying a discrete nonlinear scale-space framework: a spatial discretization on a fixed pixel grid gives a well-posed scale-space with many image-simplifying properties, and an explicit time discretization leads to a scheme which does not introduce additional oscillations. This explains why staircasing is essentially the only practically appearing instability.

Key words: Nonlinear diffusion, well-posedness, discretization, scale-space.

1 Introduction

The nonlinear diffusion filter by Perona and Malik [1] was the starting point of a new area in image processing, where images are filtered by studying their evolutions under nonlinear partial differential equations (PDEs). Unfortunately, this starting point is usually regarded as ill-posed, because the filter is designed to behave like a backward diffusion across edges. Nevertheless, practitioners use the Perona–Malik (PM) equation very often, and rarely report instabilities, except for some staircasing effects at slowly varying edges.

The mathematical problems behind the *continuous* PM equation are highly complicated and shall not be treated here. Recent results, however, give evidence that the ill-posedness of the continuous PM process is less severe than it has been assumed in the past [2, 3, 4]. Sometimes it has also been conjectured that *discretizations* have a stabilizing effect [5, 6]. This paper shows that this is indeed the case, and it clarifies to which extend one can expect straightforward discretizations to be stable.

Applying a recently discovered nonlinear discrete scale-space framework, we show that already a spatial discretization on a fixed pixel grid is sufficient to turn the PM equation into a well-posed system of ordinary differential equations, which has many nice scale-space properties. Furthermore, we shall see that the simplest and most widely used explicit (Euler forward) time discretization is monotonicity preserving in 1-D: this means that applying such a scheme does not introduce oscillations into a sigmoid-like edge. Therefore, practically observed instabilities are basically restricted to staircasing effects.

The paper is organized as follows: Section 2 explains the continuous PM filter and presents an m-dimensional semidiscrete formulation. By semidiscrete we mean discrete in space and continuous in time. In Section 3 we discuss a semidiscrete well-posedness and scale-space theory for nonlinear diffusion filters, which we apply in Section 4 to the semidiscrete PM filter. Section 5 is devoted to fully discrete results, especially to establishing monotonicity preservation. The presented proofs are simple and instructive, and the techniques may be useful for similar image processing problems. We conclude with a summary in Section 6. A preliminary version of this paper has been presented at a workshop [7].

2 The Perona–Malik filter

2.1 Continuous formulation

We consider an m-dimensional rectangular image domain $\Omega = (0, a_1) \times \cdots \times (0, a_m)$ with boundary $\partial\Omega$, and a (grey-value) image which is given by a bounded mapping $f : \Omega \to \mathbb{R}$. In order to avoid the blurring and localization problems of linear diffusion filtering, Perona and Malik proposed a nonlinear diffusion method [1]. Their nonuniform process – which they name anisotropic[1] – reduces the diffusivity at those locations which have a larger likelihood to be edges, since they reveal larger gradients. Perona and Malik obtain a filtered image $u(x,t)$ as solution of a nonlinear diffusion equation with the original image as initial condition and reflecting boundary conditions (∂_n denotes the derivative normal to the image boundary $\partial\Omega$):

$$\partial_t u = \operatorname{div}\left(g(|\nabla u|^2)\, \nabla u\right) \qquad \text{on} \qquad \Omega \times (0,\infty), \tag{1}$$

$$u(x,0) = f(x) \qquad \text{on} \qquad \Omega, \tag{2}$$

$$\partial_n u = 0 \qquad \text{on} \qquad \partial\Omega \times (0,\infty). \tag{3}$$

The "time" t is a smoothing parameter: larger values correspond to simpler image representations $u(x,t)$. Using for instance diffusivities of type

$$g(|\nabla u|^2) = \frac{1}{1 + |\nabla u|^2/\lambda^2} \qquad (\lambda > 0), \tag{4}$$

they obtained visually impressive results [1]: 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.

[1] It would be more appropriate to regard the PM filter as an isotropic model, since it uses a scalar-valued diffusivity and not a diffusion tensor. For models with a diffusion tensor, see e.g. [8, 9].

2.2 Problems

Soon it became clear that the PM approach reveals a problem: It is not hard to see that (1) can be decomposed into a process across the edge and one perpendicular to it:

$$\begin{aligned}\partial_t u &= \Big(g(|\nabla u|^2) + 2|\nabla u|^2 g'(|\nabla u|^2)\Big)\,\partial_{\eta\eta} u \\ &+ g(|\nabla u|^2)\,\Big(\Delta u - \partial_{\eta\eta} u\Big),\end{aligned} \tag{5}$$

where $\eta := \nabla u/|\nabla u|$, and Δ denotes the m-dimensional Laplacian. This process is a usual (forward) diffusion equation as long as $|\nabla u| \le \lambda$, since the factors in front of $\partial_{\eta\eta} u$ and $(\Delta u - \partial_{\eta\eta} u)$ are nonnegative. For $|\nabla u| > \lambda$, however, the factor in front of $u_{\eta\eta}$ becomes negative. Thus, across high-contrast edges, the PM filter acts like a diffusion backward in time. This forward–backward diffusion behavior is explicitly intended in the PM method, since it gives the desirable result of blurring small fluctuations and sharpening edges. On the other hand, backward diffusion is well-known to be an ill-posed process where the solution – if it exists at all – is highly sensitive even to the slightest perturbations of the initial data.

Several numerical studies can be found where instabilities in the PM process have been reported [10, 11, 6, 12, 14]. However, the main observed instability is the so-called *staircasing effect*, where a smoothed step edge evolves into piecewise almost linear segments which are separated by jumps; see e.g. [6] for an illustration. The extend of this effect depends on the discretization. Typically, fine pixel grids [12, 14] or adaptive numerical schemes with grid refinement at large gradients [6] leads to a splitting into more stairs. Such instabilities which increase with finer discretizations can be regarded as a hint that the underlying continuous equation is ill-posed. Contributions to the explanation and avoidance of staircasing can be found in [13, 14, 15, 16, 3, 4]. However, the staircasing effect is mainly visible for fine spatial discretizations and for slowly varying ramp-like edges. Under practical situations, this is hardly observed, and it is an experimental fact that discretizations of the PM are not very unstable.

A rigorous mathematical analysis of forward–backward diffusion equations is very difficult, because many classical mathematical theories are not applicable. One reason why people became very pessimistic about the well-posedness of the PM equation was a result by Höllig [17]. He constructed a forward–backward diffusion process which can have infinitely many solutions. Although this process was different from the PM process, one was warned what can happen. In 1994 the general opinion, including the one of Perona and Malik, was that the PM filter might have weak solutions[2], but one should neither expect uniqueness nor stability [18]. In the meantime several theoretical results are available which provide a better insight into the actual degree of ill-posedness of the Perona–Malik filter:

Kawohl and Kutev [2] were able to establish that the PM process has a unique weak solution which is continuously differentiable and which exists for some *finite* time. On the other hand, they also showed that such a solution cannot exist for an infinite time interval.

Kichenassamy [3, 4] proposed a notion of generalized solutions to the PM process, which are piecewise linear and contain jumps, and showed that an analysis of their moving and merging gives similar effects to those one can observe in practice.

Results of You et al. [19] give evidence that the PM process is unstable with respect to perturbations of the initial image. They showed that the energy functional leading to the PM process as steepest descent method has an infinite number of global minima which are dense in the image space. Each of these minima corresponds to a piecewise constant image, and slightly different initial images may end up in different minima for $t \to \infty$.

However, the practical behavior of finite difference approximations is much nicer than one would expect from the theory above: One can easily calculate a discrete solution for all times, and this solution converges to a flat image for $t \to \infty$. Thus, since it is well-known that discretizations may have a regularizing effect to ill-posed problems, it would be desirable to complement the preceding continuous results with a detailed analysis of the properties of practically used difference approximations. This will be done in the subsequent sections.

[2] A weak solution satisfies a generalized (integral) formulation of the PDE. In particular, a weak solution does not have to be twice differentiable in x.

2.3 Semidiscrete formulation

A discrete m-dimensional image can be regarded as a vector $\mathbf{f} \in \mathbb{R}^N$, whose components f_i, $i \in J := \{1, ..., N\}$ display the grey values at the N pixels. Pixel i represents the location x_i. Let h_l denote the grid size in l direction. By u_i and g_i we denote approximations to $u(x_i, t)$ and $g(|\nabla u(x_i, t)|^2)$, respectively. Then, a consistent spatial discretization of the PM equation with reflecting boundary conditions can be written as

$$\frac{du_i}{dt} = \sum_{l=1}^{m} \sum_{j \in \mathcal{N}_l(i)} \frac{g_j + g_i}{2h_l^2} (u_j - u_i), \tag{6}$$

where $\mathcal{N}_l(i)$ consist of the two neighbors of pixel i along the direction l (boundary pixels may have only one neighbor) and

$$g_i := g\left(\frac{1}{2}\sum_{l=1}^{m} \sum_{p,q \in \mathcal{N}_l(i)} \left(\frac{u_p - u_q}{2h_l}\right)^2\right) \tag{7}$$

uses a gradient approximation by central differences. In vector–matrix notation (6) becomes

$$\frac{d\mathbf{u}}{dt} = \sum_{l=1}^{m} \mathbf{A_l}(\mathbf{u})\, \mathbf{u} \tag{8}$$

where the matrix $\mathbf{A_l}(\mathbf{u}) = (a_{ijl}(\mathbf{u}))_{ij}$ is given by

$$a_{ijl} := \begin{cases} \frac{g_i + g_j}{2h_l^2} & (j \in \mathcal{N}_l(i)), \\ -\sum\limits_{l=1}^{m} \sum\limits_{k \in \mathcal{N}_l(i)} \frac{g_i + g_k}{2h_l^2} & (j = i), \\ 0 & (\text{else}). \end{cases} \tag{9}$$

We observe that a semidiscrete PM process leads to a nonlinear system of ordinary differential equations. Let us now study a general well-posedness and scale-space framework for problems of this type.

3 Semidiscrete nonlinear scale-space theory

While semidiscrete analogues to the linear diffusion scale-space are well-known [20], results for the nonlinear diffusion case have been found quite recently [21].

In the *continuous* setting a nonlinear well-posedness and scale-space framework has been established for *spatially regularized* variants of the PM process [5, 8, 9]. Such problems have a unique solution [5] which is stable under perturbations of the initial image [8, 9]. In addition to invariances such as the preservation of the average grey value, it has been shown that these equations create smoothing scale-spaces in spite of their contrast-enhancing potential: they obey a causality property in terms of a maximum–minimum principle, reveal many image simplifying properties, and converge to a constant steady state [8, 9].

Of course, it is desirable to find discrete processes which reveal exactly the same qualities. To this end, criteria have been identified under which it is guaranteed that a semidiscrete scheme of type

$$\frac{d\mathbf{u}}{dt} = \mathbf{A}(\mathbf{u})\, \mathbf{u}, \tag{10}$$

$$\mathbf{u}(0) = \mathbf{f}, \tag{11}$$

possesses such properties [9, 21]. All one has to check are the following criteria for $\mathbf{A}(\mathbf{u}) = (a_{ij}(\mathbf{u}))$:

(S1) $\mathbf{A}$ is a continuously differentiable function in $\mathbf{u}$.

(S2) $\mathbf{A}$ is a symmetric matrix.

(S3) All row sums of $\mathbf{A}$ are zero.

(S4) All off-diagonal elements of $\mathbf{A}$ are nonnegative.

(S5) $\mathbf{A}$ is irreducible:
Loosely speaking, this means that we can connect any two pixels by a path with nonvanishing diffusivities.
Formally: For any $i, j \in J$ there exist $k_0, ..., k_r \in J$ with $k_0 = i$ and $k_r = j$ such that $a_{k_p k_{p+1}} \neq 0$ for $p = 0, ..., r-1$.

Under these prerequisites the filtering process satisfies the following properties [9, 21]:

(a) *Well-posedness:*
The considered problem has a unique solution for all $t > 0$, which depends continuously on the initial value and the right-hand side of the ODE system.
This is of significant practical importance, since it guarantees stability under parameter variations and under perturbations of the original image. Such a property is desirable when considering stereo images, image sequences or slices from medical CT or MR sequences, since we know that similar images remain similar after filtering.

(b) *Average grey level invariance:*
The average grey level $\mu := \frac{1}{N} \sum_{j \in J} f_j$ is not affected by the semidiscrete diffusion filter:

$$\frac{1}{N} \sum_{j \in J} u_j(t) = \mu \qquad \forall\, t > 0. \tag{12}$$

This invariance is useful for scale-space based segmentation algorithms [22] and for all applications where grey values are related to physical qualities, for instance in medical imaging.

(c) *Extremum principle:*

$$\min_{j \in J} f_j \leq u_i(t) \leq \max_{j \in J} f_j \qquad \forall\, i \in J, \quad \forall\, t > 0. \tag{13}$$

This property is much more than a stability result which forbids under- and overshoots. Since it also ensures that iso-intensity linking towards the original image is possible, it states an important causality property, cf. [23].

(d) *Smoothing Lyapunov functionals:*
In spite of some possible edge enhancement, the considered processes are simplifying, information-reducing transformations with respect to many aspects: The p-norms

$$\|\mathbf{u}(t)\|_p \ := \ \Big(\sum_{j \in J} |u_j(t)|^p \Big)^{1/p} \tag{14}$$

are decreasing in t for all $p \geq 2$, all even central moments

$$M_{2n}[\mathbf{u}(t)] \ := \ \frac{1}{N} \sum_{j \in J} (u_j(t) - \mu)^{2n} \tag{15}$$

are decreasing in t, and the entropy

$$S[\mathbf{u}(t)] \ := \ -\sum_{j \in J} u_j(t) \, \ln u_j(t), \tag{16}$$

a measure of uncertainty and missing information, is increasing in t (if f_j is positive for all j).

(e) *Convergence to a constant steady state:*
The scale-space evolution tends to the most global image representation that is possible: a constant image with the same average grey level as f:

$$\lim_{t \to \infty} u_i(t) = \mu \quad \forall\, i \in J. \tag{17}$$

These semidiscrete results are part of a more general framework which includes also a continuous and fully discrete theory [21]. Full details with proofs can be found in [9], where it is also shown that it is possible to derive semidiscrete scale-spaces from spatial discretizations of continuous scale-spaces. Due to its ill-posedness, however, the *continuous* PM equation cannot be treated within the *continuous* scale-space framework.

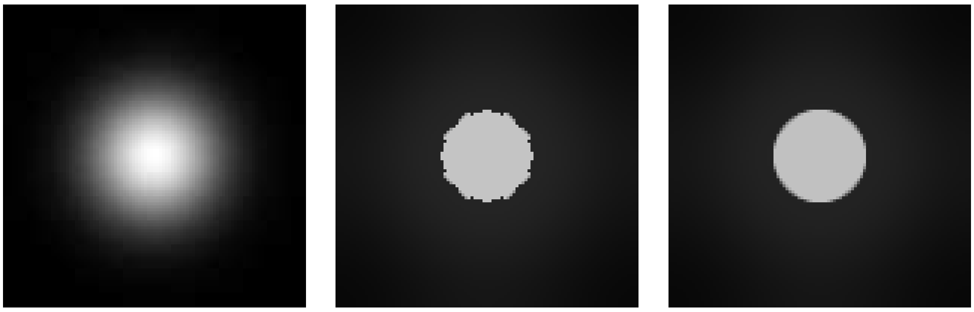

Figure 1: Nonlinear diffusion of a Gaussian-like image. (a) LEFT: Original image, $\Omega = (0, 101)^2$. (b) MIDDLE: Filtered without regularization, $\lambda = 9$, $\sigma = 0$, $t = 250$. (c) RIGHT: Filtered with regularization, $\lambda = 9$, $\sigma = 0.7$, $t = 250$.

4 Application to the semidiscrete Perona-Malik process

Interestingly, it is not hard to verify that the semidiscrete PM process satisfies the well-posedness and scale-space requirements (S1)–(S5).

(S1) Since g is infinitely times differentiable, it follows from (9) that the same also holds for $\mathbf{A} := \sum_l \mathbf{A_l}$.

(S2) The symmetry of $\mathbf{A}$ follows directly from (9) and the symmetry of the neighborhood relation:

$$i \in \mathcal{N}_l(j) \iff j \in \mathcal{N}_l(i).$$

(S3) By the construction of $\mathbf{A}$ it is also evident that all row sums vanish.

(S4) Since g is positive, it follows that $a_{ij} \geq 0$ for all $i \neq j$.

(S5) In order to show that $\mathbf{A}$ is irreducible, let us consider two arbitrary pixels i and j. Then we have to find $k_0,...,k_r \in J$ with $k_0 = s_1$ and $k_r = s_2$ such that $a_{k_q k_{q+1}} \neq 0$ for $q = 0,...,r-1$. If $i = j$, we know already from (9) that $a_{ii} < 0$. In this case we have the trivial path $i = k_0 = k_r = j$. For $i \neq j$, we may choose any arbitrary path $k_0,...,k_r$, such that k_q and k_{q+1} are neighbors for $q = 0,...,r-1$. Then,

$$a_{k_q k_{q+1}} = \frac{g_{k_q} + g_{k_{q+1}}}{2h_l^2} > 0$$

for some $l \in \{1, ..., m\}$. This proves (S5).

Despite the fact that the PM filter is regarded to be ill-posed in the continuous setting, we observe that its semidiscrete approximation on a fixed grid satisfies (S1)–(S5) and, thus, reveals all the before mentioned well-posedness and scale-space properties. Hence, the spatial discretization has a strong regularizing effect. It is caused by the fact that the extremum principle limits the modulus of discrete gradient approximations[3].

One of the remarkable consequences of the gradient restriction due to the grid regularization is the established continuous dependence of the filtered image on the initial one. This is quite unusual for a process which may act edge-enhancing. Another interesting property is the convergence to a constant image for $t \to \infty$. In some publications it has been stated that PM implementations create a segmentation for $t \to \infty$. Our result shows that this is not correct.

It should be noted that the grid regularization due to the semidiscretization is certainly not the best regularization strategy. Experiments show that grid refinement leads to more staircasing [6, 12, 14]. Moreover, grid regularization may cause artifacts which reflect the grid structure (Fig. 1(b)). One

[3]Our results are in accordance with a recent paper by Pollak et al. [24]. They study an image evolution under an ODE system with a discontinuous right hand side, which has some interesting relations to the limit case of a semidiscrete PM model. They also report stable behavior of their process.

can avoid these problems by introducing the regularization explicitly in the continuous PM model, for instance by a Gaussian smoothing of ∇u within the diffusivity [10, 5, 6, 13, 8, 9]. Figure 1(c) illustrates that – as soon as the standard deviation σ of the Gaussian approaches the pixel size – this regularization dominates over the pixel regularization and avoids their artifacts. Moreover, since the convolution with a Gaussian bounds the gradient already in the continuous formulation, it leads to continuous models which satisfy similar well-posedness and scale-space properties as in the semidiscrete case [8, 9]. Last but not least, it makes the filter more robust against noise [5] and reduces staircasing [6, 12, 14]. Thus, such regularizations are generally recommendable.

5 Fully discrete results

We have seen that a spatial discretization is sufficient to make the PM filter well-posed. Of course, in practice, one has to apply a temporal discretization as well. If we consider a time step size $\tau > 0$ and denote by u_i^k an approximation of $u(x_i, k\tau)$, then the simplest fully discrete PM equation is given by the explicit scheme

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \sum_{l=1}^{m} \sum_{j \in \mathcal{N}_l(i)} \frac{g_j^k + g_i^k}{2h_l^2} \, (u_j^k - u_i^k). \qquad (18)$$

In [9] it is shown that such a scheme inherits its well-posedness and scale-space properties from the semidiscrete one if the time step size satisfies the stability restriction[4] $\tau < 1/\sum_{l=1}^{m} \frac{2}{h_l^2}$. So we do not have to worry that we loose the nice semidiscrete results by the temporal discretization.

Finally, we prove an important property which is responsible for the fact that – apart from the grid artifacts depicted in Fig. 1 – staircasing is essentially the only practical instability that is observed when implementing the PM equation: the one-dimensional discrete PM process is *monotonicity preserving* [14]: if the grey values are increasing (decreasing) from left to right, then they remain increasing (decreasing) after filtering. This property explains why over- and undershoots and the creation of other oscillations are not possible.

Proposition 1 (Monotonicity preservation) *For $\tau < \frac{h_1^2}{2}$, the 1-D explicit PM scheme is (strictly) monotonicity preserving. For a consecutive pixel numbering this means that*

$$u_i^k < u_{i+1}^k \quad \forall i \qquad \Longrightarrow \qquad u_i^{k+1} < u_{i+1}^{k+1} \quad \forall i.$$

Proof. Consider some arbitrary inner pixel $i \in \{2, ..., N-2\}$. From

$$\begin{aligned} u_i^{k+1} &= \tau\, a_{i,i+1}\, u_{i+1}^k + (1 - \tau\, a_{i,i+1} - \tau\, a_{i,i-1})\, u_i^k \\ &+ \tau\, a_{i,i-1}\, u_{i-1}^k \\ u_{i+1}^{k+1} &= \tau\, a_{i+1,i+2}\, u_{i+2}^k + (1 - \tau\, a_{i+1,i+2} - \tau\, a_{i+1,i})\, u_{i+1}^k \\ &+ \tau\, a_{i+1,i}\, u_i^k \end{aligned}$$

and the symmetry of A we obtain

$$\begin{aligned} u_{i+1}^{k+1} - u_i^{k+1} &= \tau\, a_{i+1,i+2}\, (u_{i+2}^k - u_{i+1}^k) \\ &+ (1 - 2\tau a_{i,i+1})\, (u_{i+1}^k - u_i^k) \\ &+ \tau\, a_{i-1,i}\, (u_i^k - u_{i-1}^k). \end{aligned}$$

Let $\tau < \frac{h_1^2}{2}$ and $u_j^k < u_{j+1}^k$ for all j. Since $0 < a_{j,j+1} \le 1/h_1^2$ for all j, we know that all summands of the RHS are positive. Thus, it follows that $u_i^{k+1} < u_{i+1}^{k+1}$ for $i \in \{2, ..., N-2\}$. It is easy to see that the same reasoning is applicable to the modified equations arising at the boundaries when $i = 1$ or $N-1$. q.e.d.

[4] More efficient absolutely stable schemes can be found in [25].

6 Conclusions

In this paper we have presented a semidiscrete and discrete analysis of the Perona–Malik process. In order to explain the well-posedness of a spatially discretized PM equation, we have applied a recent framework for semidiscrete nonlinear scale-spaces. These results also carry over to the fully discrete setting, where a monotonicity preserving property has been proved. They explain that observed instabilities of PM discretizations are restricted to staircasing effects and grid artifacts affecting the rotational invariance. These effects may be tamed by introducing a grid independent spatial regularizations already in the continuous equation [10, 5, 6, 9].

Acknowledgments

This work has been funded by *Stiftung Innovation des Landes Rheinland–Pfalz*, *Deutscher Akademischer Austauschdienst*, the *Real World Computing Partnership*, and the *Danish National Science Research Council* under grant no. 9502164.

References

[1] P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, 629–639, 1990.

[2] B. Kawohl, N. Kutev, *Maximum and comparison principles for anisotropic diffusion*, Preprint, Mathematical Institute, University of Cologne, 50923 Cologne, Germany, 1997.

[3] 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.

[4] S. Kichenassamy, *The Perona–Malik paradox*, SIAM J. Appl. Math., to appear.

[5] 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.

[6] M. Nitzberg, T. Shiota, *Nonlinear image filtering with edge and corner enhancement*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, 826–833, 1992.

[7] 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, in press.

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

[9] J. Weickert, *Anisotropic diffusion in image processing*, Ph.D. thesis, Dept. of Mathematics, University of Kaiserslautern, Germany, 1996. Revised version to be published by Teubner Verlag, Stuttgart, 1997.

[10] E.S. Posmentier, *The generation of salinity finestructure by vertical diffusion*, J. Phys. Oceanogr., Vol. 7, 298–300, 1977.

[11] 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.

[12] 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.

[13] R.T. Whitaker, S.M. Pizer, *A multi-scale approach to nonuniform diffusion*, CVGIP: Image Understanding, Vol. 57, 99–110, 1993.

[14] 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.

[15] 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.

[16] 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.

[17] K. Höllig, *Existence of infinitely many solutions for a forward–backward heat equation*, Trans. Amer. Math. Soc., Vol. 278, 299–316, 1983.

[18] 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.

[19] Y.-L. You, W. Xu, A. Tannenbaum, M. Kaveh, *Behavioral analysis of anisotropic diffusion in image processing*, IEEE Trans. Image Proc., Vol. 5, 1539–1553, 1996.

[20] T. Lindeberg, *Scale-space for discrete signals*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, 234–254, 1990.

[21] 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.

[22] W.J. Niessen, K.L. Vincken, J.A. Weickert, M.A. Viergever, *Nonlinear multiscale representations for image segmentation*, Computer Vision and Image Understanding, Vol. 66, 233–245, 1997.

[23] 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.

[24] I. Pollak, A.S. Willsky, H. Krim, *Scale space analysis by stabilized inverse diffusion equations*, B. ter Haar Romeny, L. Florack, J. Koenderink, M. Viergever (Eds.), Scale-space theory in computer vision, Lecture Notes in Comp. Science, Vol. 1252, Springer, Berlin, 200–211, 1997.

[25] J. Weickert, B.M. ter Haar Romeny, M.A. Viergever, *Efficient and reliable schemes for nonlinear diffusion filtering*, IEEE Trans. Image Proc., 1998, to appear.