From Mumford-Shah to Perona-Malik in image processing

B. Kawohl*

October 4, 2003

submitted to Mathematical Methods in the Applied Sciences

Dedicated to Professor H.A. Levine on the occasion of his 60th birthday

Summary. In mathematical image processing we are often presented with amazing examples of image enhancement algorithms. Yet, when applied to different noisy images, they can produce unwanted effects. The analysis of such algorithms lags behind their intuitive development. Two essentially different models have found wide recognition: a variational approach due to Mumford and Shah, and an anisotropic diffusion approach leading to an evolution type equation by Perona and Malik. In this survey we shall explain a surprising connection between these two approaches.

Keywords. image processing, anisotropic diffusion, forward-backward diffusion, Gamma-convergence, variational method, edge enhancement

1 Introduction

Suppose u_0 is the grey-scale distribution of a noisy image. **Mumford and Shah** suggested studying the variational problem

$$J(v) = \alpha \int_\Omega |v-u_0|^2 \; dx + \beta \int_{\Omega \setminus S_v} |\nabla v|^2 \; dx + \; \mathcal{H}^{n-1}(S_v),$$

in which S_v denotes a set of discontinuities of v. A minimizer of this problem, if it exists, should have a small set of discontinuities, a relatively smooth appearance elsewhere and still resemble the original picture. One can replace the quadratic terms in the functional by strictly convex terms such as $|v|^p$ or $\sqrt{1+|\nabla v|^2}$ without

*Mathematisches Institut, Universität zu Köln, D 50923 Köln, Germany. short title: From Mumford-Shah to Perona-Malik, e-mail: kawohl@mi.uni-koeln.de

changing essential features of the problem. For reasons of exposition I stick to the simple functional above. It is nowadays minimized on $GSBV(\Omega)$, a space of special functions of bounded variation. One of the analytical problems is a lack of semicontinuity of $\mathcal{H}^{n-1}(S)$ with respect to the Hausdorff topology.

A set A is ε-close to C in the Hausdorff metric if C is contained in an ε-neighborhood of A and if A is contained in an ε-neighborhood of C.

If the Mumford-Shah functional has a minimizing sequence u_n, then the sets S_{u_n} might converge in the Hausdorff metric, but their Hausdorff measure is not necessarily lower semicontinuous, as one can see from the following example:

$S_1 := [0, \frac{1}{2}]$
$S_2 := [0, \frac{1}{4}] \cup [\frac{1}{2}, \frac{3}{4}]$
$S_3 := [0, \frac{1}{8}] \cup [\frac{1}{4}, \frac{3}{8}] \cup [\frac{1}{2}, \frac{5}{8}] \cup [\frac{3}{4}, \frac{7}{8}]$ etc.

The sequence S_i converges to the interval $S_\infty := [0, 1]$ in the Hausdorff metric, but $\mathcal{H}^1(S_\infty) = 1 > \liminf_{k\to\infty} \mathcal{H}^1(S_k) = 1/2$.

Nevertheless significant progress has been made on the study of this functional. It turns out that $GSBV(\Omega)$, a space of generalized functions of bounded variation, is the suitable space in which one can prove the existence of a minimizer, and also the regularity of the singular set of a minimizer is fairly well understood. Since we do not intend to elaborate on existence and regularity theory, we omit the definition of $GSBV(\Omega)$ and refer the interested reader to [3], [10] and [15].

Independently from this variational approach **Perona and Malik** suggested to take the noisy image u_0 as initial datum for a diffusion equation such as

$$u_t - \operatorname{div}\left(\frac{\nabla u}{(1+|\nabla u|^2)^2}\right) = 0, \tag{1.1}$$

$$\text{or more generally} \qquad u_t - \operatorname{div}\left(a(|\nabla u|^2)\nabla u\right) = 0, \tag{1.2}$$

with $a(s)$ positive and decreasing to zero as $s \to \infty$, and under no-flux boundary conditions. Small diffusion near discontinuities in u_0 was supposed to lead to edge preservation, while large diffusion elsewhere would sort of mollify the brightness function and take out noise. Let us explain why these equations are sometimes called “anisotropic” diffusion equations. If u is a classical solution, (1.2) becomes

$$u_t - a(|\nabla u|^2)\Delta u - 2a'(|\nabla u|^2)\nabla u D^2 u \nabla u = 0. \tag{1.3}$$

In noncritical points one can rewrite the Laplacian in intrinsic coordinates as

$$\Delta u = u_{\nu\nu} + (n-1)Hu_\nu = u_{\nu\nu} + \Delta_{n-1}u,$$

where $\nu = -\frac{\nabla u}{|\nabla u|}$ and where H is the mean curvature of a level surface of u, or where Δ_{n-1} is the Laplace Beltrami operator on the tangent plane to this level surface. Therefore (1.3) turns into

$$\begin{aligned} u_t - a(|\nabla u|^2)\Delta u - 2a'(|\nabla u|^2)|\nabla u|^2 u_{\nu\nu} &= 0 && (1.4)\\ u_t - b(|\nabla u|^2)u_{\nu\nu} - a(|\nabla u|^2)\Delta_{n-1} u &= 0, && (1.5)\end{aligned}$$

and the coefficient $b(s) = a(s) + 2sa'(s)$ in (1.5) can even become negative for large values of s. Then there is backward diffusion, which leads to a steepening of profiles, in direction of the gradient of u, but forward diffusion along level surfaces of u. Aside from a significant contribution of Kichenassamy [33], who proposed a notion of weak solution in the case of one space dimension, and considerations in [13] to approximate the problem by a truly parabolic one in which $|\nabla u|$ in the argument of a is mollified by convolution with a Gaussian, little is known about the theory of such diffusion equations. Nevertheless numerical results for this class of equations are extremely convincing.

It is the purpose of this note to reveal a connection between these two seemingly different approaches, the variational one and the evolutionary one.

2 Approximating the Mumford Shah functional

If one wants to minimize the Mumford Shah functional numerically, there are obvious problems. Where should one put grid-points to capture the set of discontinuities? How can one properly evaluate the Hausdorff measure of the discontinuity set (say in $n = 2$) if the discontinuities of the discrete solution can only lie on edges between gridpoints? A way out of this calamity is offered by the seminal observation of Modica and Mortola, that one can approximate $\mathcal{H}^{n-1}(S_u)$ by a domain integral. Modica & Mortola showed in 1977 that $H_0(u) = \mathcal{H}^{n-1}(u = 0)$ is the Γ-limit of

$$H_\varepsilon(u) = \int_\Omega \varepsilon|\nabla u|^2 + \frac{1}{4\varepsilon}(u^2-1)^2 \, dx.$$

Subsequently Ambrosio & Tortorelli showed in 1990 that the Mumford-Shah functional

$$J_0(v) := \int_\Omega \alpha|v-u_0|^2 \, dx + \beta \int_{\Omega\setminus S_v} |\nabla v|^2 \, dx + \mathcal{H}^{n-1}(S_v)$$

is the Γ-limit of

$$J_\varepsilon(v,\eta) := \int_\Omega \alpha|v-u_0|^2 + \beta\eta^2|\nabla v|^2 + \varepsilon|\nabla\eta|^2 + \frac{1}{4\varepsilon}(\eta-1)^2 \, dx, \qquad (2.1)$$

where the set $\{\eta = 0\}$ stands for S_v. Notice that by this trick one has approximated J_0 by a functional which contains only domain integrals over Ω but not over $\Omega \setminus S_v$.

The usefulness of this result is apparent when we recall the definition of and a principal result on Γ–convergence, see [21]. Let X be a metric space and F_ε :

$X \mapsto [0, \infty]$ a family of mappings. Then F **is the Γ-limit of** F_ε as $\varepsilon \to 0$, iff the following statements a) and b) hold.

a) For every $u \in X$ and every sequence $u_\varepsilon \to u$ in X

$$\liminf_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) \geq F(u) \ .$$

b) For every $u \in X$ there exists a sequence u_ε such that $u_\varepsilon \to u$ in X and

$$\lim_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) = F(u) \ .$$

Theorem 2.1 *If F is Γ–limit of F_ε and if u_ε is a minimizer of F_ε, then every cluster point u of $\{u_\varepsilon\}_{\varepsilon>0}$ minimizes F.*

We can therefore expect minimizers of the approximating functionals to be close to a minimizer of the limit functional, and the approximating functionals are better suited for numerical computations. Other ways to approximate the Mumford Shah functional are described in the recent book [6].

3 From Mumford-Shah to Perona-Malik

In view of the stability Theorem 2.1 let us consider minimizers of (2.1) and see how they behave as $\varepsilon \to 0$. If $(v_\varepsilon, \eta_\varepsilon)$ minimize J_ε, then $v_\varepsilon \to v$ and $\eta_\varepsilon \to 1$ in $L^2(\Omega)$, but $|\eta_\epsilon| << 1$ near S_v. The Euler equations are given by

$$\alpha(v_\varepsilon - u_0) - \beta \ \mathrm{div}(\eta_\varepsilon^2 \ \nabla v_\varepsilon) = 0 \tag{3.1}$$

for the variation with respect to the first argument of J_ε and by

$$\beta \ \eta_\varepsilon |\nabla v_\varepsilon|^2 - \varepsilon \Delta \eta_\varepsilon + \frac{1}{4\varepsilon}(\eta_\varepsilon - 1) = 0, \tag{3.2}$$

for the variation of J_ε with respect to η.

"Now a miracle occurs ...". Let us think of (3.2) as a singular perturbation problem. As $\varepsilon \to 0$ we expect $\varepsilon \Delta \eta_\varepsilon$ to become negligible compared to the other terms in (3.2). This heuristically motivates (and is justified by the remarks below) the choice of

$$\eta = \frac{1}{1 + 4\beta\varepsilon|\nabla v|^2} \tag{3.3}$$

and the neglection of the term $\varepsilon \Delta \eta_\varepsilon$ in (3.2). Upon plugging (3.3) into (2.1) and neglecting $\varepsilon|\nabla \eta|^2$ as well, we arrive at yet another functional. In fact, this trick was used in [9], [49] to approximate J_ε by

$$\tilde{J}_\varepsilon(v) = \int_\Omega \alpha|v-u_0|^2 + \tilde{A}_\varepsilon(|\nabla v|^2)\,dx \quad \text{with} \quad \tilde{A}_\varepsilon(s^2) := \frac{s^2}{1+4\beta\varepsilon s^2}. \tag{3.4}$$

The differential operator that corresponds to the functional (3.4) is of the same type as the spatial operator in the Perona-Malik approach. The Euler equation for $\tilde{J}_\varepsilon$ is in fact

$$-\mathrm{div}\left(\frac{\nabla v}{(1+4\beta\varepsilon|\nabla v|^2)^2}\right) = -\alpha(v-u_0), \tag{3.5}$$

whereas a nonlinear diffusion equation according to Perona and Malik is given for instance by (1.1)

$$u_t - \mathrm{div}\left(\frac{\nabla u}{(1+|\nabla u|^2)^2}\right) = 0$$

Remark 3.1 *On the continuous analytical level it is a crime to neglect the term $\varepsilon\Delta\eta_\varepsilon$ in (3.2), but on a discrete level it appears to be justified by the following considerations. As long as $\Delta\eta_\varepsilon$ is bounded by some moderate constant, we may neglect $\varepsilon\Delta\eta_\varepsilon$. But in points where $\Delta\eta_\varepsilon$ becomes very large we have to take a closer look. In those points the gradient of v_ε becomes very large, but its discretization cannot exceed h^{-1}. If we discretize the Laplace operator by the usual five point stencil (in 2-d, and similarly in higher dimensions) we see that the discretized version of $\Delta\eta_\varepsilon$ is of magnitude not exceeding h^{-2}. Now set $\varepsilon = h^{2+\delta}$ with $0 < \delta << 1$. Then the term involving $\varepsilon\Delta\eta_\varepsilon$ in the discretized version of (3.2) is negligible compared to the other terms even where $\Delta\eta_\varepsilon$ becomes very large. Therefore the derivation of (3.3) can in fact be justified at the discretized level. By similar considerations one can justify the passage from (2.1) to (3.4).*

In view of (2.1), (3.4) and Remark 3.1 is natural to ask if the Mumford-Shah functional can be approximated by a discrete functional that has its effective domain of definition on a space of piecewise linear finite elements. To get a Γ-convergent approximation, however, even the approximating functionals should be defined (and maybe unbounded) for all functions in $GSBV(\Omega)$. This approach was successsfully pursued in [14] and [15] and in 2 space dimensions it provides a sound proof that under the right scaling of h versus ε the finite element minimizers of (3.4) have a subsequence which converges to a minimizer of the Mumford Shah functional.

What about the term involving α in (3.4)? We can think of α as a Lagrange parameter when minimizing $\hat{J}_\varepsilon(v) = \int_\Omega \tilde{A}_\varepsilon(|\nabla v|^2)\,dx$ on a δ-neigbourhood of u_0. If we solve

$$\min\ \hat{J}_\varepsilon(v) := \int_\Omega \tilde{A}_\varepsilon(|\nabla v|^2)\,dx \quad \text{on} \quad \int_\Omega |v-u_0|^2\,dx < \delta$$

with a descent method, and with initial datum u_0, then

$$u_t - \operatorname{div}\left(\frac{\nabla u}{(1+4\beta\varepsilon|\nabla u|^2)^2}\right) = 0 \tag{3.6}$$

is the corresponding flow equation, as long as the constraint is not active. For a short time, u will stay near u_0, but then the Perona Malik approach (for short times) is useful for lowering the energy of $\hat{J}_\varepsilon$. In fact, if u solves (3.6) under the no-flux condition on the boundary, then

$$\begin{aligned} \frac{d}{dt}\int_\Omega |u-u_0|^2\,dx &= \int_\Omega (u-u_0)u_t\,dx = -\int_\Omega \frac{(\nabla u - \nabla u_0)\nabla u}{(1+4\beta\varepsilon|\nabla u|^2)^2}\,dx \\ &< \int_\Omega \frac{\nabla u_0 \nabla u}{(1+4\beta\varepsilon|\nabla u|^2)^2}\,dx \le \int_\Omega |\nabla u_0|^2\,dx + c\ , \end{aligned}$$

so that u increaes its distance to u_0 with at most finite speed, at least provided u_0 is in $H^1(\Omega)$. Notice again, that on a discrete level, even discontinuous but bounded initial data are perceived as being in $H^1(\Omega)$, so that the right hand side in the last estimate may be large but is finite in numerical experiments.

To conclude this section let me remark that I have gone the way from Mumford Shah functionals to energies associated with Perona-Malik equations. There is a also a recent investigation [38] of Morini and Negri, in which the authors go from spatially discrete Perona-Malik energies to (anisotropic) Mumford Shah functionals. Moreover, in [23] Esedoglu has investigated a one-dimensional Perona-Malik equation with continuous time and discrete space variable and its asymptotic behaviour as the space discretization goes to zero. All these papers are directed at a better understanding of the unreasonable effectiveness of the Perona-Malik approach. Instead of discretizing the space variable, one can discretize time. In this case we can make the following observation:

Remark 3.2 *If we use an implicit Euler discretization for (3.6) in which k denotes the size of a time-step and $u_n(x) = u(nk,x)$, the evolution equation turns into*

$$\frac{u_{n-1}-u_n}{k} - \operatorname{div}\left(\frac{\nabla u_{n+1}}{(1+|\nabla u_{n+1}|^2)^2}\right) = 0 \tag{3.7}$$

or rather

$$-\operatorname{div}\left(\frac{\nabla u_{n+1}}{(1+|\nabla u_{n+1}|^2)^2}\right) = -\frac{1}{k}(u_{n-1}-u_n)\ . \tag{3.8}$$

If we set $v = u_{n+1}$ and compare (3.8) with (3.5), the stunning similarity of the success of the variational and the evolutionary approach comes as no surprise. Here a rescaling of time amounts to tuning the $1/k$ to α.

4 Image enhancement via diffusion

It has already been mentioned above that nonlinear diffusion approaches to image enhancement show better than expected numerical stability. In this section I want to describe qualitative results from [31] on solutions of the Perona Malik equation. Consider the following diffusion problem

$$
\begin{aligned}
u_t - \operatorname{div}(a(|\nabla u|^2)\nabla u) &= 0 \quad \text{in} \quad \Omega \times (0,T) && (4.1)\\
a(|\nabla u|^2)\nabla u &= 0 \quad \text{on} \quad \partial\Omega \times (0,T) && (4.2)\\
u(x,0) &= u_0(x) \quad \text{in} \quad \Omega && (4.3)
\end{aligned}
$$

in a smooth and bounded domain Ω and with $a \in C^1([0,\infty))$ positive. A typical candidate is $a(s) = 1/(1+ds)^2$ as in (1.1) or [20]. The ellipticity function $b(s). = a(s) + 2sa'(s)$ is positive for small s but changes sign exactly once at $s_0^2 > 0$. How can one classify this equation?

a) If $|\nabla u| < s_0$, it is a regular parabolic equation, and for $|\nabla u| \leq s_0$ it is degenerate parabolic.

b) The potential flow equation in gasdynamics of an ideal gas is of type (4.1), and s_0 is the speed of sound. So points in which $|\nabla u|$ is greater (smaller) than s_0 are in the supersonic (subsonic) regime of such flow.

c) In one space dimension, equation (4.1) becomes a forward-backward nonlinear diffusion equation. Backward equations transform smooth initial data like Gaussian distributions into very large and very steep functions such as δ-Distributions.

The last observation c) has led experts to believe that one cannot expext a maximum principle to hold. Therefore the following Theorem is somewhat surprising.

Theorem 4.1 *Suppose that u is a weak $C^{0,1}$-solution to the anisotropic problem* (4.1) – (4.3). *Then* $\max|u(x,t)| = \max|u_0(x)|$.

For a proof we choose $p \in (1,\infty)$ and calculate, using (4.1) and (4.2)

$$
\begin{aligned}
\frac{d}{dt}\int_\Omega |u(x,t)|^p \, dx &= p\int_\Omega |u|^{p-2}u \; u_t \, dx\\
&= p\int_\Omega |u|^{p-2}u \; \operatorname{div}(a(|\nabla u|^2)\nabla u) \, dx\\
&= -p(p-1)\int_\Omega |u|^{p-2}(a(|\nabla u|^2)|\nabla u|^2 \, dx\\
&< 0
\end{aligned}
$$

Therefore $L^p(\Omega)$-norms are decreasing in t. Now one can send p to infinity to see the claim. This short proof was found independently by Weickert [50] and Kutev and myself [31].

On the next pages you find in Figure 1 some numerical studies of M.Mester for the case that $a(s) = 1/(1+s)^2$ with $s_0^2 = \frac{1}{3}$ and with $u_0(x,y) = \sin(x\pi)\sin(y\pi)$ on a square of length 2, and in Figure 2 the same study for the slightly modified case that $a(s) = 1/(1+4s)^2$ with $s_0^2 = \frac{1}{12}$ and same intial data as before.

We observe that the size of s_0 determines what the equation recognizes as an edge that should be enhanced. If s_0 is large then diffusion is stronger.

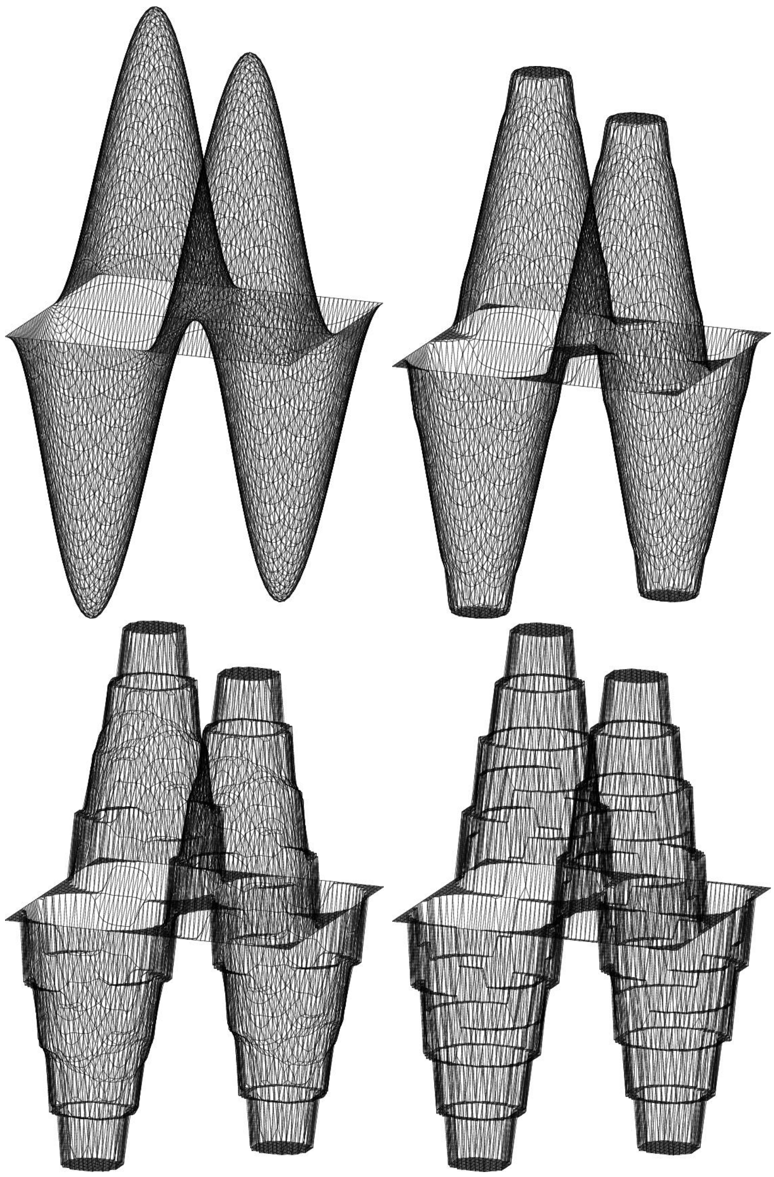

Figure 1: $a(s) = 1/(1+s)^2$ for $t = 0, t = 0.11111, t = 0.20444$ and $t = 0.44444$

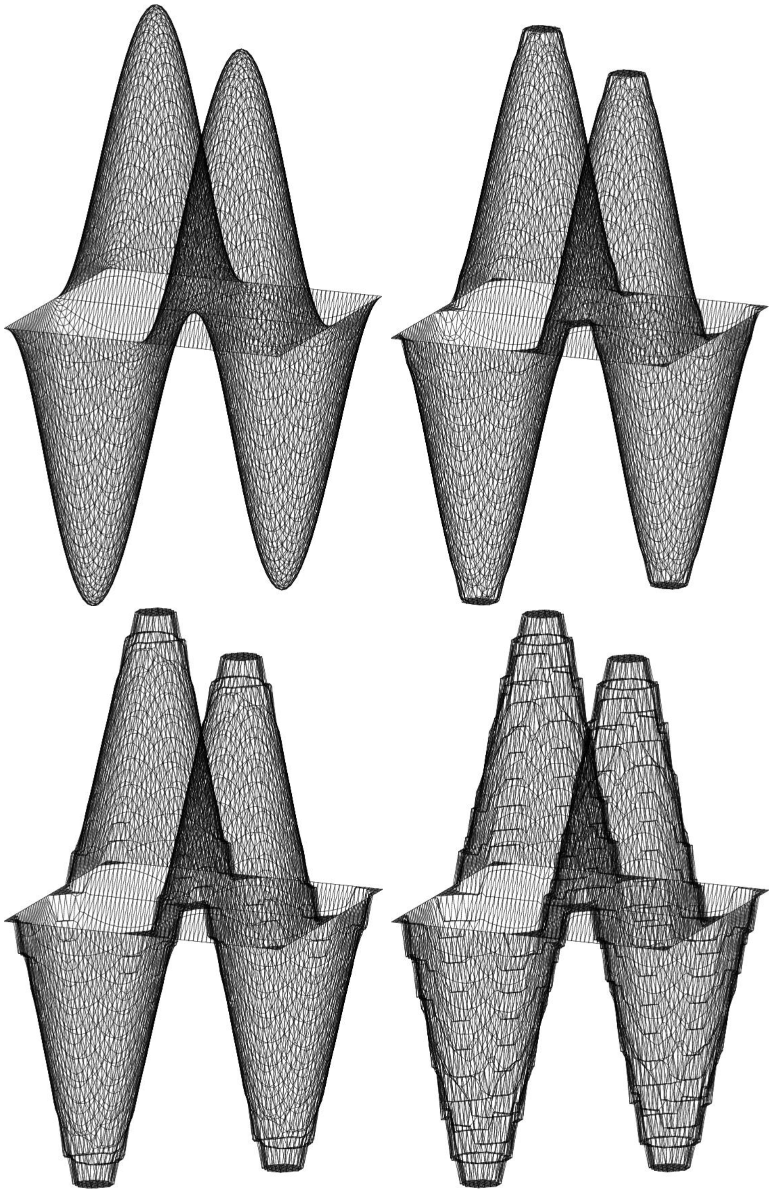

Figure 2: $a(s) = 1/(1+4s)^2$ for $t = 0, t = 0.11111, t = 0.20444$ and $t = 0.44444$

We also observe the staircasing effect which is described in many places. This is due to the fact, that the Perona-Malik energy is nonconvex. Therefore the flow tries to minimize the convex envelope of the Perona-Malik energy. On the continuous level, the convex envelope is identically zero and does not contain much information. But again on the discrete level, because $|\nabla u|$ cannot exceed $1/h$ with h denoting grid size, the convex envelope of the energy restricted to such arguments is nontrivial. This envelope coincides not everywhere with the original energy, and so instead of taking a value in the set where these energies are distinct, $|\nabla u|$ oscillates between large values of order $1/h$ and small values where the functionals coincide. This oscillation effect of minimizing sequences for nonconvex functionals is well understood. In the end the smooth initial datum is made "sharp", and that is a desired effect in the enhancement of satellite pictures or medical imaging.

Significantly more can be said about the qualitative behavior of solutions when the space domension is $n = 1$ and $\Omega = (-1, 1)$, say. In that case the differential equation and boundary condition become

$$\begin{aligned} u_t - b(|u_x|^2)u_{xx} &= 0 \quad \text{in } (-1,1)\times(0,T) && (4.4)\\ a(|u_x|^2)u_x &= 0 \quad \text{in } \{-1,1\}\times(0,T). && (4.5)\end{aligned}$$

Unfortunately, little appears to be known about such equations. The existence problem is wide open, and not even uniqueness can be expected if one allows for Lipschitz solutions as in [30]. Therefore Kutev and I considered C^1-solutions in [31]. Assuming that those exist we can identify if a point (x, t) in $\Omega \times (0, T)$ belongs to the forward or backward regime of equation (4.4), because $|u_x|^2$ can be pointwise evaluated. If it exceeds the threshhold s_0^2 for b in a point (x, t), then (x, t) is in the backward regime. In particular for a C^1 initial datum u_0 one can say which points of $\Omega \times \{0\}$ belong to the forward or backward regime.

Theorem 4.2 *If u is a weak C^1-solution of* (4.4), (4.5) *and if the initial datum is analytic, then the backward regime of u shrinks in time and does not migrate.*

The proof of this result is quite lengthy and can be found in [31]. One reason why the analyticity of u is assumed is the fact that we want only finitely many places where $|du_0/dx|$ equals s_0, so that at time $t = 0$ we have finitely many switches from one regime to the other. Theorem 4.2 states in other words, that the Perona-Malik flow enhances edges.

Theorem 4.2 can also be used to derive a comparison result for solutions $u(x, t)$ and $v(x, t)$ of (4.4), provided the initial data are ordered, i.e. $u_0(x) \le v_0(x)$ and have disjoint backward regimes. In that case, if $u(x, t) - v(x, t)$ should happen to have a positive maximum for some positive t_0 and some x_0, then $|u_x(x_0, t_0)| = |v_x(x_0, t_0)|$ and this value does not exceed s_0, because the backward regimes of u and v never intersect by Theorem 4.2. Hence one can apply comparison results for forward parabolic equations to reach a contradiction. For details I refer to [31].

The restrictive assumptions on such a comparison result are necessary. In [31] we have also exhibited a situation, where $u_0(x) \leq v_0(x)$ and where the two graphs touch each other with slope greater than s_0 in a point x_0. Then after arbitrarily short time they have crossed each other.

Finally we have shown in [31] that solutions of (4.4),(4.5) must cease to be of class C^1 after finite time, because some integral quantity must blow up in finite time. (I learned this trick from Howard Levine, see for instance the seminal paper [34].) In fact, numerical evidence suggests that they can (and should) even become discontinuous. In that case one can switch from the representation $u(x,t)$ to a representation $x = x(u,t)$ and still do some analysis. This line of research was recently (and successfully) pursued in [7]. In a different but related context, this transformation came also up in [32].

In this paper I have disregarded operators which lead to degenerate parabolic equations. The special function $a(s^2) = s^{-1/2}$, for instance, falls outside of our class of admissible a because it is singular at $s = 0$ and because $b \equiv 0$. There there has been significant progress on equations of this and of slightly similar type, see for instance papers of Evans and Spruck, Oliker and Uraltseva, Deckelnick and Dziuk, Mikula et al.[29] or Andreu, Caselles, Mazón and coauthors [2] and [1]. Alhough these equations have no "backward diffusion feature", they can justfiably be used for image enhancement as well, because their associated energies can be interpreted as convex relaxations of nonconvex Perona-Malik energies. Let me remark in passing that different nonconvex energy functionals can have virtually the same convex relaxation, if ∇u is restricted to the numerically relevant ball of radius $1/h$. To be specific, one can compare the relaxation of the functional $\int I_{[-1/h^2,1/h^2]}(|\nabla u|)\ \tilde{A}_\varepsilon(|\nabla u|^2)\ dx$ from (3.4) with that of the so called Blake-Zisserman functional (see [8])

$$\int_\Omega I_{[-1/h,1/h]}(|\nabla u|)\ \min\{\gamma|\nabla u|^2, k\}\ dx$$

for appropriate positive constants γ and k. Here $I_A(z)$ is defined as 1 if $z \in A$ and as $+\infty$ otherwise.

Recent books on the subject of this article and related problems are [6], [37], [46], [48] and [50]. There are also numerous generalizations of these nonlinear diffusion equations to equations involving fourth order derivatives with respect to x or functionals which depend on curvatures of u. These are too numerous to be listed here, and so I just refer to [4], [11], [12], as a starting point for the interested reader.

Acknowledgement. An essential part of this paper was written while the author enjoyed the hospitality of V. Oliker at Emory University in the spring of 2002. I thank M.Mester for producing and providing the calculations and Figures 1 and 2. Financial support came partly from Emory University and from DFG. I am grateful to the referee for helpful suggestions.

References

[1] Andreu, F., Ballester, C, Caselles, & Mazón, J.M. The Dirichlet problem for total variation flow. *J. Functional Anal.* 2001; **180**:347–403.

[2] Andreu, F., Caselles,V., Diaz, J.I. & Mazón, J.M. Some qualitative properties for the total variation flow. *J. Functional Anal.* 2002; **188**:516–547.

[3] Ambrosio, L. Existence theory for a new class of variational problems. *Arch. Ration. Mech. Anal.* 1990; **111**:291-322.

[4] Ambrosio, L., Faina, L. & March, R. Variational approximation of a second order free discontinuity problem in computer vision. *SIAM J. Math. Anal.* 2001; **32**:1171–1197.

[5] Ambrosio, L. & Tortorelli, V. M. Approximation of functionals depending on jumps by elliptic functionals via Γ- convergence. *Comm. Pure Appl. Math.* 1990; **43**:999–1036.

[6] Aubert, G. & Kornprobst, P. *Mathematical Problems in Image Processing* Springer, New York Berlin Heidelberg, 2002

[7] Barenblatt, G.I. & Vazquez, J.L. Nonlinear diffusion and image contour enhancement. Preprint Feb 2003

[8] Blake, A. & Zisserman, A. *Visual Reconstruction*, The MIT Press, Cambridge, Mass., 1987.

[9] Blomgren, P., Chan, T., Mulet, P., Vese, L. & Wan, W.L. Variational PDE models & methods for image processing, in *Numerical Analysis 1999*, D.F. Griffiths & G.A. Watson (Editors), Chapman & Hall/CRC Research Notes in Mathematics, Boca Raton 2000; **420**: 43–67.

[10] Bonnet, A. & Guy, D. Cracktip is a global Mumford-Shah minimizer. *Astérisque* 2001; **274** :1–259.

[11] Carriero, M., Leaci, A. & Tomarelli, F. Local minimizers for a free gradient discontinuity problem in image segmentation, in Variational methods for discontinuous structures, G. Dal Maso & F. Tomarelli (Editors) Birkhäuser Progress in Nonlinear Differential Equations and Their Applications, Basel 2002; **51**:67–80.

[12] Carriero, M., Leaci, A. & Tomarelli, F. Calculus of variations and image segmentation. Politecnico Preprint 506/P, Milano 2002;

[13] Catté, F., Lions, P.-L., Morel, J.-M. & Coll, T. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.* 1992 ; **29**:182–193.

[14] Chambolle, A. Image segmentation by variational methods: Mumford and Shah functionals and the discrete approximations. *SIAM J. Appl. Math.* 1995; **55**:827–863.

[15] Chambolle, A. & Dal Maso, G. Discrete approximation of the Mumford Shah functional in dimension two. *Mathem. Modelling and Num Anal.* 1999; **33**:651–672.

[16] Chambolle, A. & Lions P.L. Image recovery via total variation minimization and related problems. *Numer. Math.* 1997; **76**:167–188.

[17] Chan, T.F., Shen, J. & Vese, L. Variational PDE models in image processing. *Notices Amer. Math. Soc.* 2003; **50**:14–26.

[18] Chan, T. & Vese, L. Active contours without edges, *IEEE Trans. Image Processing* 2001; **10**:266–277.

[19] Chan, T. & Vese, L. Variational image restoration and segmentation models and approximations, *IEEE Trans. Image Processing*; to appear

[20] Chipot, M., March, R., Rosati, M. & Vergara Caffarelli, G. Analysis of a nonconvex problem related to signal selective smoothing. *Math. Models Meth. Appl. Sci* 1997; **7**:313–328.

[21] Dal Maso, G. *An Introduction to* Γ*–convergence*. Birkhäuser Verlag, Basel, 1993.

[22] Deckelnick, K. & Dziuk, G. Convergence of a finite element method for non-parametric mean curvature flow. *Numer. Math.* 1995; **72**:197–222.

[23] Esedoglu, S. An analysis of the Perona-Malik scheme. *Comm. Pure Appl. Math.* 2001; **54**:1442–1487.

[24] Esedoglu, S. Stability properties of Perona-Malik scheme. Preprint, Minneapolis 2001;

[25] Evans, L.C. & Spruck, J. Motion of level sets by mean curvature I. *J. Differ. Geom.* 1991; **33**:635–681.

[26] Evans, L.C. & Spruck, J. Motion of level sets by mean curvature II. *J. Trans. Amer. Math. Soc.* 1992; **330**:321–332.

[27] Evans, L.C. & Spruck, J. Motion of level sets by mean curvature III. *J. Geom. Anal.* 1992; **2**:121–150.

[28] Gajewski, H. & Gärtner, K. On a nonlocal model of image segmentation. WIAS-Preprint 763, Berlin 2002;

[29] Handlovičová, A. & Mikula, K. Analysis of a semidiscrete scheme for solving image smoothing equations of mean curvature flow type. *Acta Math. Univ. Comeniae* 2002; **71**:93–111.

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

[31] Kawohl, B. & Kutev, N. Maximum principle and comparison theorems for anisotropic diffusion. *Mathematische Annalen.* 1998; **311**:107–123.

[32] Kawohl, B. Remarks on quenching, *Documenta Mathematica (Electronic Journal of DMV)* 1996; **1**:199–208.

[33] Kichenassamy, S. The Perona-Malik Paradox. *SIAM J. Appl. Math.* 1997; **57**:1328–1342.

[34] Levine, H.A. Some nonexistence and instability theorems for solutions of formally parabolic equations of the form $Pu_t = -Au + \mathcal{F}(u)$. *Arch. Rational Mech. Anal.* 1973; **51**:371–386.

[35] March, R. Visual reconstruction with discontinuities using variational methods. *Image and Vision Computing* 1992; **10**:30–38.

[36] Mester, M. Mathematische Methoden zur Bildbearbeitung, doctoral thesis in preparation

[37] Morel, J.-M. & Solimini, S. *Variational methods in image segmentation* Birkhäuser Verlag, Boston, 1995

[38] Morini,M. & Negri, M. Mumford Shah functional as limit of discrete Perona Malik energies. *Math. Models and Meth. in Appl. Sci.* 2003; **13**:785–805.

[39] Mumford, D. & Shah, J. Optimal aproximations by piecewise smooth functions and associated variational problems, *Comm. Pure Appl. Math.* 1989; **42**:577–685.

[40] Negri, M. & Paolini, M. Numerical minimization of the Mumford-Shah functional. *Calcolo* 2001; **38**:67–84.

[41] Oliker, V.I. Evolution of nonparametric surfaces with speed depending on curvature I. The Gauss curvature case. *Indiana Univ. Math. J.* 1991; **40**:237–258.

[42] Oliker, V.I. & Uraltseva, N.N. Evolution of nonparametric surfaces with speed depending on curvature II. The mean curvature case. Comm. Pure Appl. Math. 1993; **43**:97–135.

[43] Oliker, V.I. & Uraltseva, N.N. Evolution of nonparametric surfaces with speed depending on curvature III. Some remarks on mean curvature and anisotropic flows, in *Degenerate diffusions (Minneapolis, MN, 1991)*, W.M. Ni, L.A. Peletier & J.L. Vazquez (Editors), Springer IMA Vol. Math. Appl., New York, Berlin, Heidelberg 1993; **47**:141–156.

[44] Oliker, V.I. & Uraltseva, Long time behavior of flows moving by mean curvature. in *Nonlinear evolution equations*, Amer. Math. Soc. Transl. Ser. 2 1995; **164**:163–170.

[45] Oliker, V.I. & Uraltseva, N.N. Long time behavior of flows moving by mean curvature. II. *Topol. Methods Nonlinear Anal.* 1997; **9**: 17–28.

[46] Osher, S. & Fedkiw, M: *Level set methods and dynamic implicit surfaces* Springer, New York Berlin Heidelberg, 2003

[47] Perona, P. & Malik, J. Scale space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.* 1990; **12**:629–639.

[48] Sapiro, G. *Geometric partial differential equations and image analysis* Cambridge University Press, Cambridge, 2001

[49] Vese, L. & Chan, T. Reduced non-convex functional approximations for image restoration and segmentation. UCLA CAM Report 97-56, 1997;

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

[51] Weickert, J. Applications of nonlinear diffusion in image processing and computer vision. *Acta Math. Univ. Comeniae* 2001; **70**:33–50.

Author's addresses:

Bernd Kawohl
Mathematisches Institut
Universität zu Köln
D-50923 Köln
Germany
kawohl@mi.uni-koeln.de