Advanced Studies in Pure Mathematics 67, 2015
Variational Methods for Evolving Objects
pp. 131–156

Anisotropic diffusions of image processing from Perona–Malik on

Patrick Guidotti

Abstract.

Many reasons can be cited for the desire to harness the power of nonlinear anisotropic diffusion in image processing. Perona and Malik proposed one of the pioneering models which, while numerically viable, proves mathematically ill-posed. This discrepancy between its analytical properties and those of its numerical implementations spurred a significant amount of research in the past twenty years or so. An overview of the latter is the topic of this article.

§1. Introduction

Mathematical image processing was profoundly influenced by the introduction of two prototypical models in the late 80ies, early 90ies: Mumford–Shah [60] and Perona–Malik [62]. The first is a variational model as it is formulated by means of the functional

$$E_{MS}(u,\Gamma)=\frac{1}{2}\int_{\Omega\setminus\Gamma}|\nabla u|^2\,dx+\alpha\int_{\Omega}|Ku-u_0|^2\,dx+\beta l(\Gamma)\,, \tag{1}$$

where a possibly non-smooth minimizer u is sought for which the length $l(\Gamma)$ of its discontinuity set Γ, in a typical two dimensional setting, is penalized. Away from the set Γ, the first terms enforces smoothness, while the second penalizes overall deviation from the observed data u_0. The smoothing operator K models any blurring present in the image. When $K=\mathrm{id}$, one deals with the "pure" denoising problem. The second

Received January 16, 2013.
2010 *Mathematics Subject Classification.* Primary 35-02; Secondary 35K55, 35K65.
Key words and phrases. Anisotropic diffusion, Perona–Malik equation, regularization, relaxation, semi-discretization, forward-backward diffusion, gradient systems.

model, while at least formally the gradient flow associated to

$$E_{PM}(u) = \frac{1}{2}\int_\Omega \log\bigl(1+|\nabla u|^2\bigr)\,dx\,, \tag{2}$$

is best viewed as an example of forward-backward diffusion in view of the convex-concave nature of

$$\varphi(s) = \frac{1}{2}\log(1+|s|^2)\,,\ s\in\mathbb{R}^2\,. \tag{3}$$

The corresponding PDE reads

$$u_t = \nabla\cdot\Bigl(\frac{1}{1+|\nabla u|^2}\nabla u\Bigr) = \frac{1}{1+|\nabla u|^2}\Bigl(\partial_{\tau\tau}u + \frac{1-|\nabla u|^2}{1+|\nabla u|^2}\partial_{\nu\nu}\Bigr)\,, \tag{4}$$

and clearly allows for backward diffusion in direction ν normal to the level sets of the function u for large enough gradients, while always maintaining a forward nature in tangential direction(s) τ. For this very reason, equation (4) is customarily become labeled as *a*, if not *the* example of, *anisotropic diffusion*. The two models above, or related ones derived from them, are typically used for so-called low level image processing tasks such as denoising, deblurring, or segmentation, to name only the most common. In such a context, u_0 represents a given image which needs to be enhanced (read denoised, deblurred, or segmented, respectively). It explicitly appears in (1) and enters into (4) as the initial condition for the Cauchy problem. The Mumford–Shah model (1) presents non-trivial numerical challenges and has therefore lead to a variety of approximations (often in the sense of Γ-convergence) including an early one by Ambrosio and Tortorelli [3] which replaces the perimeter term by an approximating bulk integral. It is, in a way, remarkable that, on the contrary, the Perona–Malik model (4) poses significant mathematical challenges while delivering better behaved numerical implementations than expected.

The focus of this overview paper is on anisotropic diffusion and no further mention will be made of the numerous subsequent developments in the variational arena revolving around or inspired by the Mumford and Shah model. It is only pointed out in passing that a, at least formal, connection between (1) and (4) was found by Chan and Vese [19] (see also the overview article by Kawohl [53]).

In the early 90ies Perona and Malik [62] proposed a nonlinear model for image processing in order to replace/improve on previous techniques

based on linear filtering followed by edge identification and reconstruction. Linear filtering essentially amounts to solving a linear heat equation

$$\begin{cases} u_t = \Delta u\,, \\ u(0,\cdot) = u_0\,, \end{cases}$$

whereas the second phase requires an edge detector typically based on the use of $|\nabla u|$ or $|\Delta u|$ such as is the case for the Canny or for the Marr–Hildreth edge detectors [16], [59]. Perona and Malik's idea consisted in combining the two steps into a single one by integrating edge detection into a nonlinear diffusion equation modulated by a diffusivity varying according to the local geometric features of the image. While they actually consider a discrete model in their paper, its "apparent" continuos counterpart has come to be known as the Perona–Malik equation. For the sake of completeness it should be mentioned that they also considered $e^{-|\nabla u|^2}$ as an alternative diffusivity function and that only the qualitative behavior of the flux function $\nabla\varphi$ really matters in applications. For this reason only (3) is considered in this paper. Perona and Malik observe in [62] that "solutions" should, contrary to the case of linear backward diffusion, satisfy a maximum principle, thus excluding the possibility of L^∞-instability. This is later proved to be true for weak C^1-solutions by Kawohl and Kutev [54] and independently by Weickert [72]. The existence of a backward regime, however, makes (4) ill-posed. A fact that was first formalized in a paper by Kichenassamy [55] in 1996.

Numerical experiments strongly suggest that the ill-posedness of (4) manifests itself through the staircasing effect, so-called for its characteristic appearance (see Fig. 1). In the intervening years after the original publication of [62] many attempts were made to either

- gain a satisfactory understanding of (4).
- find well-posed models preserving salient features of (4).
- produce a mathematically sound "rationale" for the behavior of typical solutions observed in numerical experiments.

The rest of the paper is devoted to providing a representative (albeit possibly incomplete) overview of these attempts.

§2. Results about the Perona–Malik equation

While the Perona–Malik equation is an example of a forward-backward diffusion and, as such, it is ill-posed, its instability does not manifest itself at the level of the function values of its solutions since

$$\|u(t,\cdot)\|_\infty \le \|u_0\|_\infty\,,\ t > 0\,, \tag{5}$$

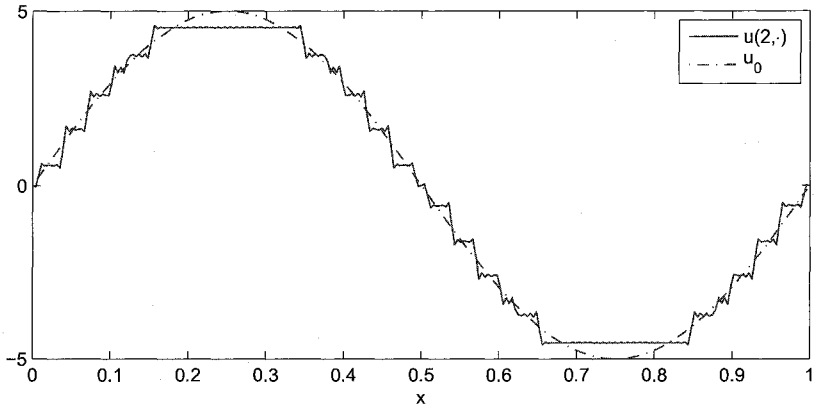

Fig. 1. A typical manifestation of staircasing. Depicted are an inital datum u_0 and the solution $u(t,\cdot)$ of (4) at $t = 2$.

as was proved in [54] and was already mentioned above. Numerical experiments strongly suggest that ill-posedness affects the behavior of the gradient of the solution leading to the staircasing phenomenon. The equation possesses, however, a classical regime in which solutions are smooth and converge to a trivial steady-state. It indeed follows from the maximum principle [54, Theorem 6.1] that a solution, which is initially subcritical, satisfies

$$\|\nabla u(t,\cdot)\|_\infty \le \|\nabla u_0\|_\infty\,,\ t > 0\,,$$

and thus remains subcritical for all times. A function is called *subcritical* if its gradient is (everywhere) in the forward regime [$|s| < 1$] of the equation. If it exhibits non-trivial regions both in the backward and forward regimes of the equation it is called *transcritical*. A proof of the ill-posedness of was formalized by Kichenassamy in [55], where he showed that a solution can only exist (in general) for a transcritical initial datum if it is very special (read very smooth). Most of the early results obtained for the Perona–Malik equation pertain the one dimensional case. They will be described first. Kichenassamy proposed a concept of generalized solution for the one-dimensional version (4) which would include piecewise smooth functions and would allow piecewise constant functions, in particular, to be considered stationary for the evolution. This would provide some intuitive, if not rigorous and quantitative, explanation for the generic staircasing behavior of solutions. The

definition of generalized solution used relies, however, on taking a limit of solutions in which the equation goes lost. More precisely, a function $u : U \to \mathbb{R}$ is a generalized solution according to [55] of

$$P(u) := u_t - \partial_x\Big(\frac{1}{1+|\partial_x u|^2}\partial_x u\Big) =: u_t - \partial_x R(\partial_x u) = 0$$

if an approximating sequence of Lipschitz continuous functions $(u_n)_{n\in\mathbb{N}}$ exists such that

$$\begin{cases} u_n \to u & \text{in } \mathrm{L}^1_{loc}(U) \\ R(\partial_x u_n) \rightharpoonup R(\partial_x u) & \text{in } \mathrm{L}^2(U) \\ P(u_n) \rightharpoonup 0 & \text{in } \mathcal{D}'(U) \end{cases} , \text{ as } n \to \infty . \tag{6}$$

According to this definition, piecewise constant functions are generalized steady-states. In particular, this would indicate that the location of discontinuity points do not migrate with time. This is in contrast with the behavior of generalized solutions of (4) which are piecewise smooth (but not piecewise constant) solutions of which exhibit moving discontinuity locations according to

$$[u]dx + [R(\partial_x u)]dt = 0$$

where $[f]$ denotes the jump of a function f across a space-time discontinuity curve. More can be found in the original paper [55]. Observe that this concept of generalized solution is such that no equation is found which is satisfied by the limiting function but rather is such that the residual $P(u_n)$ converges to zero in the very weak sense of distributions. It is worthwhile pointing out that, while piecewise constant functions seem to steal the show in numerical experiments, they cannot be allowed as stationary solutions of (4) in any standard sense since the term $\frac{1}{1+|\partial_x u|^2}$ cannot be made sense of since it is a non-convex nonlinearity of a measure for such piecewise constant functions. Notice also that numerical experiments indicate that the location of singularities do not move, if present and/or once formed, but that their intensity (jump height) is observed to decrease as a function of time.

The emphasis of [54] is, by contrast, mainly on weak C^1-solutions of (4). In particular, the authors show the non-existence of weak global C^1-solutions to transcritical initial data and, that, beyond (5), no comparison principle can hold for solutions in general. If two comparable initial data are separated by a fully subcritical initial datum or an alternative very specific structure condition is satisfied, a comparison principle still holds. Paper [54] also contain other results concerning the transcritical region, uniqueness, and other qualitative properties.

In a series of papers Gobbino and collaborators investigate questions pertaining classical solutions to (4). In [40] the authors analyze the behavior of the subcritical region extending some related results about the one dimensional case contained in [54]. They show that this region grows as a function of time in the one dimensional setting as well as in the radially symmetric case, while it can be non-expanding in higher dimensions, in general. Since they are able to exactly characterize the growth rate in the radially symmetric case, they also devise a class of initial conditions for which finite time blow up has to occur. In essence initial conditions with large enough gradients in the supercritical regime cannot decrease their gradient fast enough for it to have disappeared by the estimated time the subcritical region would have to have conquered the whole domain. This contradiction means that such a solution has to become singular. While this results and the one contained in [54] yield non-existence of smooth solutions in general, Ghisi and Gobbino showed in [38] that the set of initial conditions for which a classical solution can be constructed in some small interval of time (depending on the initial datum) is actually dense in C^1. The proof is constructive and is based on alternatingly solving the equation forward and backward in the subcritical and supercritical regimes, respectively. Such solutions on adjacent time cylinders can then be successfully glued together to generate a rich enough class of initial data to yield density. For the sake of completeness let [41] and [37] be mentioned. In the first the authors show that time-independent affine functions are the only C^1 solutions of the one dimensional Perona–Malik equation on the whole real line. In the second paper various a priori estimates are obtained including a total variation estimate for C^2-solutions.

While global C^1-solutions of (4) cannot exist to transcritical initial data in a one-dimensional setting, this turns out not to be the case in higher dimensions. In [39] the authors indeed describe a family of smooth solutions with transcritical initial data which become subcritical in finite time, and, therefore, exist globally.

Smooth and classical solutions to the Perona–Malik equation certainly yield some insight into its properties. Such solutions are, however, unlikely to be observable in numerical experiments as they are highly unstable. If one is interested in explaining numerical results, a more appropriate type of solutions need to be found/considered. Zhang and co-authors [65], [75], [20] make this very point in a series of papers where they investigate existence, uniqueness, and stability of certain weak solutions. The common starting point of these articles consists in recasting the original equation in one and two dimensions as a (non-evolutionary) differential inclusion problem. This opens the door to the use of a set

of (relaxation) technologies developed to deal with micro-structures in material science [26], [5], [6], [27], [74]. The main idea is to exploit the non-convexity of (2) to prove non-uniqueness (and instability) of weak ($W^{1,\infty}$- and Young measure) solutions. Consider the function

$$\sigma(s) = \nabla\varphi(s) = \frac{s}{1+|s|^2}\,,\ s \in \mathbb{R}^n\,.$$

which, in one space dimension, has the form depicted in Fig. 2. The construction performed in [75], [20] in order to show the existence of infinitely many Young measure solutions in one and two dimensions begins by fixing a center of mass solution u^* (for the Young measure solution) obtained by solving the uniformly parabolic equation associated to $\sigma^*(s)$. The flux function σ^* is chosen to yield uniform parabolicity, to coincide with σ in a small neighborhood of the origin, and to become affine at infinity with positive slope (see Fig. 2). The initial data u_0 are required to be smooth, $u_0 \in C^{3+\alpha}$, and to satisfy homogenous Neumann boundary conditions. Notice that this enforces subcriticality close to the boundary. In a carefully chosen region, the gradient of u^* is then replaced by the appropriate convex combination of the two gradient values satisfying

$$\sigma^*(\nabla u^*) = \sigma(s) \text{ and } \nabla u^* \parallel s$$

via a laminate type construction yielding the effective gradient ∇u^*. The solutions constructed in this fashion are classical for small gradients, satisfy homogeneous Neumann boundary condition (in a classical sense), and have bounded gradients which can be estimated (indirectly) in terms of the gradient of the initial condition. The non-uniqueness of such solutions essentially stems from the apparent arbitrariness of the choice of the "center of mass flux", i.e. of σ^*. It is also clear that the construction does not deliver all standard classical solutions that exist purely within the subcritical regime since the flux function is also modified in the region in which it is monotone increasing. For this very same reason, this construction cannot deliver staircasing solutions even as it highlights important consequences of the lack of convexity of (2) as it excludes the onset of very large gradients by confining the support of the Young measure component of the solution to a region which is bounded away from zero and from infinity. The convexification φ^{**} of a non-convex energy functional φ would typically deliver the natural (and unique) center of mass for the construction of Young measure solutions. The relaxation of (2) is identically zero, which means $\sigma^* = 0$ in this case. Thus the natural center of mass evolution would be any time independent function $u(t,\cdot) \equiv u_0$. Indeed if a laminate construction

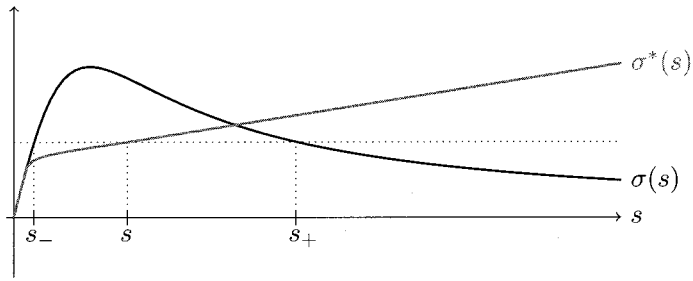

Fig. 2. The flux functions σ and σ^*.

were possible using a combination of gradients with zero and infinite size, staircasing solutions would be described by them. Even in such an event, however, the relative volume fraction of the different gradients could not be uniquely determined. Intuitively staircasing solutions would be obtained by letting the coincidence region $[\sigma^* = \sigma]$ shrink to the origin. This point of view is taken in [45] where a forward-backward regularization of (4) is proposed.

By using similar ideas but a more sophisticated construction, Zhang [75] also shows the existence of infinitely many $\mathrm{W}^{1,\infty}$-solutions for (4) in the one dimensional case. The construction yields a class of solutions without uniqueness and which doesn't include all classical solutions for reasons similar to the ones given above.

To better illustrate the basic idea of Zhang's approach, the two dimensional result of [20] is described in some detail. First observe that (4) can be rewritten as

$$\mathrm{div}_{(t,x)} \begin{pmatrix} u \\ -\frac{\nabla u}{1+|\nabla u|^2} \end{pmatrix} = 0$$

if $\mathrm{div}_{(t,x)}$ denotes the space-time divergence operator. It is known that the above implies the existence of a vector field ψ such that

$$\mathrm{curl}\,\psi = \begin{pmatrix} u \\ -\frac{\nabla u}{1+|\nabla u|^2} \end{pmatrix}.$$

Then, if $(\psi, u) \in \mathrm{W}^{1,\infty}(\Omega, \mathbb{R}^3)$ solves

$$(7) \qquad \begin{pmatrix} \mathrm{curl}\,\psi \\ Du \end{pmatrix} \in \left\{ \begin{pmatrix} u & -\frac{x_1}{1+|x|^2} & -\frac{x_2}{1+|x|^2} \\ y & x_1 & x_2 \end{pmatrix} \;\middle|\; x_1, x_2, y \in \mathbb{R} \right\} =: K(u)$$

it would follow that

$$u_t - \nabla \cdot \Big(\frac{1}{1+|\nabla u|^2}\nabla u\Big) = 0 \text{ in } \mathcal{D}'(\Omega)\,,$$

since div(curl) = 0 in the sense of distributions. The problem is then reduced to constructing a bounded sequence $(u_j)_{j\in\mathbb{N}}$ in $\mathrm{W}^{1,\infty}$ such that $u_j \overset{*}{\rightharpoonup} u^*$ in $\mathrm{W}_0^{1,\infty}(Q_T)$ where $u^* \in \mathrm{C}^{2,1}(\overline{Q}_T)$ is a center of mass solution as described above, obtained from a σ^* as in Fig. 2, and satisfying

$$u_j = u^* \text{ in } [\mathrm{dist}(\cdot, \partial\Omega) < \delta_j]\,,$$

for $\delta_j > 0$ and

$$\mathrm{div}_{(t,x)} \begin{pmatrix} u_j \\ -\frac{1+|\nabla u_j|^2}{\nabla u_j} \end{pmatrix} \longrightarrow 0 \text{ in } \mathrm{W}^{-1,p}(Q_T)\,,\ p < \infty\,. \tag{8}$$

The construction relies on the solvability of a relaxed problem

$$\begin{pmatrix} \mathrm{curl}\,\psi \\ Du \end{pmatrix} \in L(u)$$

and an explicit device by which the solution of the relaxed problem can be successively modified to get closer and closer to satisfying (7) in a sense leading to (8). The gradient Young measure $\nu_{(t,x)}$ generated by $(\nabla_{(t,x)} u_j)_{j\in\mathbb{N}}$ yields a Young measure solution of the original Perona–Malik equation, i.e. satisfying

$$0 = \int_{Q_T} \Big\{ \int_{\mathbb{R}^3} \tau\, d\nu_{(t,x)}(\tau,\lambda)\Phi(t,x) + \int_{\mathbb{R}^3} \frac{1}{1+|\lambda|^2}\lambda\, d\nu_{(t,x)}(\tau,\lambda)\cdot\nabla\Phi(t,x)\Big\}\, dxdt\,, \tag{9}$$

$$\forall\, \Phi \in \mathrm{W}_0^{1,p}(Q_T)\,,$$

where

$$\int_{\mathbb{R}^3} \tau\, d\nu_{(t,x)}(\tau,\lambda) = \partial_t u^*\,,\ \int_{\mathbb{R}^3} \lambda\, d\nu_{(t,x)}(\tau,\lambda) = \nabla u^*\,,$$

$$\text{and } \int_{\mathbb{R}^3} \frac{\lambda}{1+|\lambda|^2}\, d\nu_{(t,x)}(\tau,\lambda) = \sigma^*(\nabla u^*)\,.$$

While the details of the procedure are not given here, it is pointed out that the construction of the center of mass solution requires an initial datum $u_0 \in \mathrm{C}^{3,1}(\overline{\Omega})$ and a modified flux function σ^* derived from the

original taking into consideration the value of $\|\nabla u_0\|_\infty$. Again non-uniqueness stems in essence from the arbitrariness of the effective flux function σ^*.

§3. Regularizations and relaxations

3.1. Spatial regularizations

In [68] Weickert contends that "regularization is modeling". While undoubtedly interesting from a mathematical point of view, the question of capturing some or all solutions of the Perona–Malik equation through an approximation procedure, is somewhat in conflict with the pragmatic need for viable alternate (image) models. Regularizations and (time-) relaxations of (4) are an important source of such alternatives and are therefore of independent interest as they essentially rely on their own distinct "edge detectors". This is an important point as it makes engineering and studying the mathematical properties of regularizations a worthwhile investigative goal.

Regularization is typically taken to mean the replacement of the unknown in an equation's nonlinear term, in the situation at hand $\frac{1}{1+|\nabla u|^2}$, by a more regular version of it. In [17], the authors propose the use of a Gaussian kernel for the purpose, yielding

$$u_t = \nabla \cdot \Big(\frac{1}{1+|\nabla G_\sigma * u|^2}\nabla u\Big)\,. \tag{10}$$

As the a priori-bound in the L^∞-norm (5) remains valid for the regularized equation, one readily obtains that

$$\|\nabla G_\sigma * u\|_{\mathrm{L}^\infty} \leq c(\sigma)\|u\|_{\mathrm{L}^\infty} \leq c(\sigma)\|u_0\|_{\mathrm{L}^\infty}\,.$$

The equation therefore never degenerates and solutions can exist globally as shown in [17]. This approach has the disadvantage of re-introducing blur into solutions (at least for small scales, initially) and thus blunting the sharpening capabilities of Perona–Malik. This is one of the reasons why alternative, milder regularizations were sought. An important feature of the Perona–Malik equation is that it admits, at least on a purely formal level (as already mentioned when discussing [55]), piecewise constant functions (built from characteristic functions of sets with smooth boundaries) as steady-states. This lends some heuristic justification to the strong sharpening effect of the equation. Any standard regularization such as that of [17] does not preserve this salient and, arguably, essential property of (4). Some authors have attempted to overcome this short-coming. Cottet and Germain [25], for instance, looked for a model

which would not exhibit trivial dynamical behavior. They implement the idea by considering the regularized reaction-diffusion model

$$u_t = \sigma\varepsilon^2\nabla\cdot\big(A_\varepsilon(u)\nabla u\big) + f(u) \tag{11}$$

and then deriving a linearized/discrete version of it with the desired properties. The matrix A_ε is chosen as (essentially) the orthogonal projection to the direction τ_ε perpendicular to ∇u_ε where u_ε is a mollified version of u, i.e. such that

$$A_\varepsilon(u)v = (\tau_\varepsilon\cdot v)\tau_\varepsilon \text{ where } \tau_\varepsilon = \frac{1}{\sqrt{|\nabla u_\varepsilon|^2+\varepsilon^2}}\begin{pmatrix}-\partial_{x_2}u_\varepsilon\\ \partial_{x_1}u_\varepsilon\end{pmatrix}$$

whereas the reaction term satisfies

$$f(\pm1)=0\,,\ rf(r)>0\,,\ r\in(-1,1)\setminus\{0\}\,.$$

Notice that, unlike [17], this regularization does not formally approach (4) in the limit of vanishing ε. It actually converges to a diffusionless limiting equation. The reaction term is indeed essential in their analysis which is limited to special piecewise constant functions of the form

$$u=\begin{cases}-1\,, & x\notin D\,,\\ 1\,, & x\in D\,,\end{cases}$$

for D satisfying some assumptions. This reaction-diffusion is taylor-made to have the above type of piecewise constant functions behave like almost steady-states and thus generates a non-trivial behavior of solutions. Their effort to devise a model with this kind of properties underscores the desire to obtain well-posed models with non-trivial dynamics of Perona–Malik type. In a similar vein of truly anistropic diffusions, Weickert in [71] (see also [68], [72], [69], [70]) proposed two models, one which he termed edge-enhancing and one labeled as coherence enhancing. The first more closely mimics the Perona–Malik paradigm in that the diffusion tensor $D(u)$ in

$$u_t = \nabla\cdot\big(D(u)\nabla u\big)$$

is chosen as to have eigenvectors v_1 and v_2 satisfying

$$v_1\parallel\nabla u_\sigma \text{ and } v_2\perp\nabla u_\sigma\,,$$

and eigenvalues chosen as

$$\lambda_1 = g(|\nabla u_\sigma|^2) \text{ and } \lambda_2 = 1\,.$$

While the diffusion so obtained never inverts its direction it still can degenerate in direction perpendicular to level sets of u. In the second model a measure of local coherence is introduced via the spectral gap $\mu_1 - \mu_2$ which is obtained from the eigenvalues $0 \leq \mu_2 \leq \mu_1$ of $J_\rho(\nabla u_\sigma) = K_\rho * \left(\nabla u_\sigma \nabla u_\sigma^T\right)$, where K_ρ serves a regularizing purpose. The choice of tensor is similar to the above. One takes a diffusion tensor with the same eigenvectors v_1 and v_2 as J_ρ and sets the corresponding eigenvalues to

$$\lambda_1 = \alpha \text{ and } \lambda_2 = \begin{cases} \alpha\,, & \mu_1 = \mu_2\,, \\ \alpha + (1-\alpha)e^{-c/(\mu_1-\mu_2)^{2m}}\,, & \text{otherwise} \end{cases}$$

where $\alpha \in (0,1)$.

A well-posed model proposed in [47] shows that a strong edge-enhancing character is not incompatible with spatial regularization. The idea is to weaken the nonlinear diffusion coefficient in Perona–Malik only very slightly by the use of fractional derivatives. The corresponding model reads

$$u_t - \nabla \cdot \Big(\frac{1}{1+|\nabla^{1-\varepsilon} u|^2}\nabla u\Big) = 0\,. \tag{12}$$

The fractional gradient can be defined component-wise in various ways. In [47] the authors choose to work in a periodic setting and to define it through its Fourier symbol in a standard way. Well-posedness (the problem is no longer of forward-backward type) is not the only benefit of this model. It actually turns out that the use of slightly less than one derivative, obtained choosing $\varepsilon \in (0, 1/2)$, allows one to prove that characteristic functions of smooth sets (or combinations thereof) are stationary solutions of the equation and greatly influence the dynamic behavior of the equation. Numerical experiments show that typical solutions do relatively quickly evolve towards a piecewise constant function in their vicinity and subsequently converge to a trivial steady-state over the long run. Latter convergence is slow and occurs through gradual decrease in the size of the jumps. Implementations of Perona–Malik show that this indeed happens on a much longer time scale than the "simplification" to piecewise constancy. As (12) is locally well-posed, it suppresses the initial phase of almost instantaneous singularity formation (onset of staircasing) observed in implementations of Perona–Malik and related to lack of convexity. It does so, however, without preventing singularity formation all together. Analytical results concerning (12) and a related model are obtained in [44], [46] and [43]. These results are mostly local in time, however, and thus fall short of clarifying the interesting issue of global existence of smooth or weak solutions for the model

relating to potential finite time singularity formation, which would help better understand the numerical findings. Known PDE techniques do not seem to apply in this case. While it is possible to regularize the equation by adding viscosity and obtain limits for approximating sequences of solutions, it is not clear how to carry the equation to the limit. This is a price paid by the loss of variational structure incurred when introducing the fractional derivative. It is also observed that there is a nice transition from Perona–Malik to classical linear diffusive behavior as ϵ wanders from 0 to 1 as shown in [46] in the one dimensional case.

An even milder regularization was proposed in [45] where the Perona–Malik equation is regularized by a viscous term ($\delta > 0$)

$$u_t - \nabla \cdot \left(\left[\frac{1}{1 + |\nabla u|^2} + \delta \right] \nabla u \right) = 0. \tag{13}$$

Even after regularization, the equation remains of forward-backward type (at least for $\delta < 1/8$). It is the formal gradient flow of the modified energy functional

$$PM_\delta(u) = \frac{1}{2} \int_\Omega \left[\log(1 + |\nabla u|^2) + \delta |\nabla u|^2 \right] dx\,, \tag{14}$$

which, unlike (2), is eventually convex. This regularization preserves the time scale of singularity formation but reduces staircasing ("microstructured" jumping) to a milder micro-structured ramping with alternate gradients of finite size (which depends on δ). Functional (14) has a non-trivial convexification which can be taken as to uniquely (and naturally) determine a center of mass solution in a Young-measure type construction of weak solutions. While [45] follows a procedure first devised in [31] (see also [21], [57] for precursors in the elliptic case) in order to construct weak Young measure valued solutions of (13), an approach based on minimizing movements (see [2] or the classical [30], [28]) can also be taken to directly construct a center of mass solution first and, only, subsequently generate a Young measure solution explicitly by exploiting the specifics of the flux function. While (as in [65], [75], [20]) general Young measure solutions are not unique, those constructed in this fashion are (see [45], [31]). Regularizations of this type were subsequently considered in [48] by using the p-Laplacian in place of the Laplacian, i.e., using

$$E_{\delta,p}(u) = \int_\Omega \left[\frac{1}{2} \log(1 + |\nabla u|^2) + \frac{\delta}{p} |\nabla u|^p \right] dx\,,$$

for $p \in [1, \infty)$. Gamma convergence techniques allow to identify a time scale such that the vanishing viscosity limit ($\delta = 0$) solutions of the

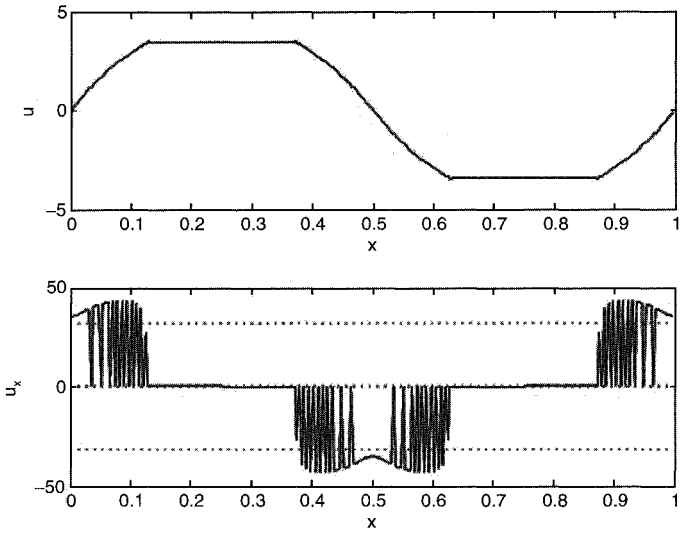

Fig. 3. Microstructured gradient for a typical solution of (13).

associated gradient flows satisfy the Total Variation minimization flow

$$u_t - \nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) = 0\,.$$

The proof is carried out for $p = 2$ in [23]. Notice that the relaxation of the limiting functional for $p = 1$ does coincide with the TV functional. From a dynamical point of view, regularization (13) replaces the staircasing effect of Perona–Malik with what could be called a micro-ramping phenomenon by which the effective (center of mass) solution is actually Hölder (even Lipschitz) continuous while its gradient exhibits a micro-structure (captured by a Young-measure) comprised of alternating gradients of small and large size. See [45] for more detail and Fig. 3 for a one dimensional illustration. Micro-ramping would formally be replaced by stair-casing in the limit of vanishing viscosity on the fast time scale of its occurrence. The slow time scale is indeed captured by the Gamma limit mentioned above.

Fourth order regularizations have also been considered. In [9] this approach is taken in the one dimensional context. The driving energy

functional is modified to

$$E^{\varepsilon}_{BF}(u) = \frac{1}{2}\int_0^1 \left[\varepsilon u_{xx}^2 + \log(1+u_x^2)\right] dx\,. \tag{15}$$

The authors of [9] are able to capture the long time behavior of the equation (evolution of piecewise constant data) by using Γ-convergence techniques in combination with the appropriate scaling. Latter reads $\varepsilon^2 = \nu^4 \log(1+1/\nu^2)$ and is motivated by an estimate of the energy of a single jump discontinuity which also leads to the rescaled energy

$$E^{\nu}_{BF}(u) = \frac{E^{\varepsilon}_{BF}}{\nu \log(1+1/\nu^2)}\,.$$

This family of energies L^1-Γ-converges to

$$E_{BF}(u) = \gamma \sum_{x\in J(u)} |[u]_x|^{1/2}\,,$$

an energy defined on $\mathrm{L}^1\big([0,1]\big)$ which is finite on a subspace of $\mathrm{SBV}\big([0,1]\big)$-functions with vanishing absolutely continuous part of the derivative. The constant γ has to be chosen appropriately and $[u]_x$ is the jump of u at $x \in J(u)$, where $J(u)$ is the not necessarily finite jump set of u. They also studied the minimizing movements evolution engendered by this energy originating in initial data with finitely many and with infinitely many jumps. In the case of finitely many jumps, the evolution is mostly determined by a system of ODEs for the jump height with the exception of finitely many "singularization" times at which one or more jump discontinuities disappear (their size vanishes). The location of discontinuities (as long as they are present) is an invariant of the evolution.

3.2. Temporal regularizations/relaxations

The Perona–Malik equation can also be regularized by convolution or relaxation in time. The latter approach is proposed in an early image processing paper by [61]. In this paper the authors suggest to delay or smooth out the effect of the nonlinearity by replacing (4) by the system

$$\begin{cases} u_t = \nabla\cdot\left(\frac{1}{1+v^2}\nabla u\right), \\ v_t = \omega\big(\rho * |\nabla u|^2 - v\big)\,, \end{cases} \tag{16}$$

or the system

$$\begin{cases} u_t = \nabla\cdot\left(\frac{1}{1+v^2}\nabla u\right), \\ v_t = \omega\big(|\nabla u|^2 - v + \frac{\sigma^2}{2}\Delta u\big)\,. \end{cases} \tag{17}$$

In both case some delay ($\omega^{-1} > 0$) along with some regularization (ρ and $\sigma > 0$) is introduced in an effort to partially tame the strength of the original nonlinearity. Perona–Malik is recovered in the limit as $\omega \to \infty$, $\sigma \to 0$, and $\rho \to \delta$, where δ now denotes the Dirac distribution. Weickert in [68] cites [63] as a reference that P. L. Lions, in a private communication to Mumford, claimed that (16) is a well-posed regularization of (4). A number of researcher have taken up the analysis of variations of the above models. Among them one finds [24], [7], [8], [1]. The last paper considers a regularization of (4) by time convolution with a positive kernel θ

$$u_t = \nabla \cdot \Big(\frac{1}{1 + \theta *_t |\nabla u|^2} \nabla u\Big)\,,$$

where $\theta *_t |\nabla u|^2(t) = \int_0^t \theta(t-\tau)|\nabla u|^2(\tau)\,d\tau$. This model turns out to be locally well-posed and globally well-posed when the support of θ is bounded away from zero, so that the nonlinearity does not include information up to the present time. It is referred to [1] for a detailed account. However, the rationale put forth by the author for considering this time-delayed model is its formal connection to a semi-explicit Euler discretization of the original (4)

$$\frac{u^{n+1} - u^n}{h} = \nabla \cdot \Big(\frac{1}{1 + |\nabla u^n|^2} \nabla u^{n+1}\Big)$$

where the nonlinearity would be evaluated at the previous time step. The authors of [24] considered the following relaxation

$$\begin{cases} u_t - \nabla \cdot \big(L \nabla u\big) = 0\,, \\ L_t - L = F(\nabla G_\sigma * u)\,, \end{cases}$$

where L is a matrix-valued function representing the diffusion tensor connected to the gradient of u by the second equation. If L is assumed to be initially positive definite, F to be bounded with bounded continuous derivative with nonnegative definite range, and G_σ(for $\sigma > 0$) to be a bounded C^1-kernel with bounded derivative, it is proved in [24] that the system possesses a unique solution on any time interval in appropriately chosen function spaces. Finally in [7], [8] the limiting case where ρ becomes the Dirac delta function is considered with the modified relaxation equation

$$v_t = F(|\nabla u|^2) - v\,,$$

where F is introduced as a technical device (read regularizing truncation) to ensure the validity of a priori estimates. More recently [64] analyzes

the well-posedness of a model of type (17) with relaxation equation given by

$$v_t = \lambda \Delta v + (1-\lambda)\big(|\nabla u| - v\big)\,,$$

for a parameter $\lambda \in (0,1)$ which can be chosen dynamically in numerical implementations.

§4. Semi-discrete models

Regularizations of the type described above have the distinct advantage of working in any space dimension. This is particularly important when it comes to applications to image processing where the dimension is typically at least two. In dimension one, however, more detailed analyses can often be performed. In such a context, an appealing alternative to regularizations of the type described in the previous section is given by spatial discretization. This way the PDE reduces to an ODE system and the mathematical problems stemming from concavity of the continuous limiting equation can be avoided altogether. Such regularizations do, furthermore, mimic numerical schemes used in practical implementations of the model and their analysis can shed light on the observed behavior of such algorithms. Particularly well understood is the long time behavior of (4) which can be captured by letting the mesh size of the discretization converge to zero by concurrently speeding up the evolution in order to capture a non-trivial limit in the sense of Γ convergence. This approach is taken in a number of papers concerning the one dimensional case including [11], [12], [10], [9], [22]. Paper [12] contains a result on the asymptotic expansion (in the sense of Γ-convergence) of the discrete Perona–Malik functional and [10] derives a system of ODEs connecting jump sizes and locations describing the evolution of piecewise constant initial data in a long time limit of the semi-discrete Perona–Malik scheme with Dirichlet boundary conditions which reads

$$u_t = D_n^-\Big(\frac{D_n^+ u}{1+|D_n^+ u|^2}\Big)\,. \tag{18}$$

Here $D_n^{\pm}$ stand for the for forward and backward difference quotients of the piecewise constant function u on $[0,1]$ with a finite number of jump discontinuities at the grid points of a uniform grid of mesh size $1/n$ for $n \in \mathbb{N}$. For convenience the collection of all such functions is denoted by PC_n. For technical reasons some of the results contained in these two papers do, however, not apply to the precise nonlinearity in (18) nor can they handle the evolution beyond the first singularization time. Latter is the time when plateaus of piecewise constancy of the solution

merge removing one or more discontinuities. Relying on these results and making an elegant use the theory of maximal slope curves [30], [29], [2], the authors of [22] obtain a complete description of the asymptotic slow time evolution of limiting solutions to (18) with more natural (from the application point of view) Neumann boundary conditions. They are able to obtain a global in time evolution by renormalizing the driving energy after each singularization in the following manner. The discrete Perona–Malik energy for which (18) can be viewed as a gradient flow is given by

$$E^n_{PM}(u) = \frac{1}{2}\int \log\bigl(1+|D^+_n u|^2\bigr)\,dx\,,\ u\in PC_n\,. \tag{19}$$

A reinterpretation of an asymptotic expansion of [12] yields

$$\Gamma-\lim_{n\to\infty} G^k_n(u) = \begin{cases} -\infty\,, & u\in PC_D\,,\ |D|<k\,,\\ G^k_\infty(u)\,, & u\in PC_D\,,\ |D|=k\,,\\ \infty\,, & \text{otherwise,}\end{cases}$$

where PC_D is the space of piecewise constant functions with exactly k jumps $d\in D=\{d_1,\dots,d_k\}$ of size J_d in the open interval $(0,1)$ and

$$G^k_n(u) = nE^n_{PM}(u) - k\log n\,,\ G^k_\infty(u) = \sum_{j=1}^{k}\log\bigl(|J_{d_j}|\bigr)$$

are defined on the space PC_n and on the subspace of piecewise Lipschitz and subcritical functions PS_D, i.e. satisfying the subcriticality inequality $|D^+u(x)|<1$ on their regularity subintervals and having jumps located at exactly k points collected in the set D. The functionals G^k_n lead, up to the renormalization term, to an accelerated version of the gradient flow associated to (19) with acceleration factor given by the number of grid points. The renormalization term $k\log n$ is important but inconsequential as far as the associated gradient system on PC_n

$$u_t = -n\nabla PM_n(u) = -\nabla G^n_k(u)\,,\ \text{for } t>0\,,$$

is concerned. All function spaces used can be thought of as subspaces of $\mathrm{L}^2(0,1)$ and the all functionals are thought of as being defined on the latter, extending the original by the value ∞ if necessary. Then, in essence, it is possible to describe the limiting evolution of initial data in

PC_D via the limiting system of ODEs given by

$$(20)\qquad \begin{cases} a_0'(t) = \frac{1}{d_1 - d_0}\frac{1}{a_1(t) - a_0(t)}\,, & a_0(0) = a_{00} \\ a_i'(t) = \frac{1}{d_{i+1} - d_i}\left[\frac{1}{a_{i+1} - a_i} - \frac{1}{a_i - a_{i-1}}\right], & a_i(0) = a_{0i} \\ a_k'(t) = -\frac{1}{d_{k+1} - d_k}\frac{1}{a_k - a_{k-1}}\,, & a_k(0) = a_{0k}\,, \end{cases}$$

where $i = 1, \ldots, k-1$, the initial datum u_0 takes values a_{0i} on the interval (d_i, d_{i+1}) for $i = 0, \ldots, k$, and $d_0 := 0$ as well as $d_{k+1} := 1$. It is recalled that $d_i \in (0,1)$ for $i = 1, \ldots, k$. System (20) has to be interpreted as to evolve initial data with no jump trivially (constant in time) and piecewise constant ones with exactly k jumps by the given equations up to the first singularization time where the number of jumps has necessarily to decrease (otherwise the solution could be continued a little longer). It then evolves the newly created piecewise constant functions according to (20) with a different k reflecting the new number of jumps. This evolution is well-defined as it can be shown that solutions are Hölder continuous up to the singularization time. See [22] for a detailed account.

An alternative approach to taking the vanishing mesh size limit on a mesh dependent time scale was taken in [35]. Instead of speeding up the evolution by rescaling time, the author chooses the threshold parameter κ of

$$u_t = \Big(\frac{1}{1 + u_x^2/\kappa} u_x\Big)_x$$

as a function of the mesh size h (of a uniform grid on $[0,1]$). The relation leading to a distinguished non-trivial limit for the evolution of piecewise smooth initial data with discontinuities at a finite number of points is given by $\kappa(h) = h^{-1}$, in which case nonlinear boundary conditions are obtained in the limiting problem across the discontinuity points and linear diffusion away from these. If u_i denotes the solution on the interval $[d_j, d_{j+1})$ for $j = 0, \ldots, n$ where $d_1, \ldots, d_{n-1}$ are the jump locations, d_0 and d_n are the end points of the unit interval, and

$$J_i = u_{i+1}(d_i, t) - u_i(d_i, t)$$

is the jump at d_i for $i = 1, \ldots, n-1$, then the limiting problem can be formulated as

$$(21)\qquad \begin{cases} \partial_t u_i = \Delta u_i\,, & \text{in } (d_{i-1}, d_i)\,, \ \text{for } t > 0\,, \\ \partial_x u_i(d_i) = \partial_x u_{i+1}(d_i) = \frac{1}{J_i}\,, & \text{for } i = 1, \ldots, n-1\,, \\ \partial_x u_i(d_i) = 0\,, & \text{for } i = 0, n\,. \end{cases}$$

The idea of choosing a mesh size dependent parameter to obtain an interesting limit was already present in [18] where it was used in a stationary context for the Mumford–Shah functional (1). As rescaling κ is at least formally akin to rescaling time, the limiting system also qualitatively captures the medium and long time behavior of solutions to numerical experiments with diffusive behavior away from the discontinuities and jump height reduction at the discontinuities. As in [22], a Hölder continuity argument allows the author of [35] to continue solutions beyond singularization times when one or more jumps vanish. For a detailed analysis it is referred to the original paper. Finally in [36] a (partial) two-dimensional comparison principle (other than the one of [54]) is obtained. Latter is subsequently applied in order to derive an interesting stability property of the (one dimensional) numerical scheme for (4) which shows that small enough "random" perturbations to initial data are quickly reabsorbed during the evolution, in agreement with numerical observations.

§5. A hybrid model

Since the ill-posedness and instability of Perona–Malik in its continuous formulation manifests itself in the onset of an uncontrolled number of jumps (captured qualitatively by Young-measure valued solutions discussed earlier), another approach can be taken which avoids this issue completely by preventing the unstable region to loose thickness in the limit. This is done in [56] where the author proposes a new image model by which the jump set is given and fixed but jump size is allowed to evolve. More precisely let $u : \Omega \to \mathbb{R}$ be an image which is slowly varying away from an edge set Γ (in fact a finite union of Lipschitz curves). Then u is evolved from its inital value I by the gradient flow of the functional

$$E_K(u) = \int_{\Omega\setminus\Gamma} \big[G(|\nabla u|) + \frac{1}{2}\alpha(u-I)^2\,dx + h\int_\Gamma F(\big|[u]_\Gamma\big|/h)\,d\sigma_\Gamma \tag{22}$$

so that it is smoothed out away from Γ and the size of its jumps is reduced over time. The nonlinear terms are chosen so that G behaves like F in the subcritical region and becomes quadratic at infinity, whereas F is of Perona–Malik type. The parameter h measures the edge "thickness". It is referred to the original article for a comprehensive description and for the rationale behind this choice. Assuming that $G - \lambda|\cdot|^2$ and $F + \mu|\cdot|^2$ are convex for some choice of $\lambda, \mu \in [0, \infty)$, nonlinear semigroup theory ([14]) can be applied to obtain an exponentially growing semigroup and, thus, a unique solution to the evolutionary inclusion

problem

$$\partial_t u + \partial E_K(u) \ni 0 \text{ in } t > 0\,,$$

on $\mathrm{L}^2(\Omega)$. As the assumptions suggest, the result is obtained by shifting the generator to $\partial E_K(u) + 2\beta u$ and treating $\partial E_K(u)$ as a Lipschitz perturbation of the shifted operator. The resulting semigroup is therefore exponentially growing in general. Thus Perona–Malik's instability has been effectively tamed but only mildly, as exponential divergence of nearby initial data over the long run shows. This model, like the one proposed in [47], has the benefit of delivering a well-posed model which predicts and agrees with numerical observations. Unlike the latter, however, it requires that the discontinuity set be known and fixed in advance.

§6. Fourth order models

Finally it should be mentioned that, while the mathematical community focussed on obtaining satisfactory insight into the Perona–Malik model and its many modifications, the more applied or application oriented community followed, among other avenues the trail of fourth order models. The motivation was almost always to try and prevent the cartoonish look generated by second order methods due to their strong tendency to produce piecewise constant output. Without any claim of completeness, here is a list of references to modeling papers [66], [67], [73], [58], [33], [32], [4], [51], [52], [49] where higher order models are considered and another, [13], [42], [34], [50], of papers dealing with related mathematical questions. Mathematical results are much more scarce for higher order methods than for their second order counterpart and many important basic questions remain open.

§7. A final remark

PDE based methods in image processing are only one of a variety of techniques employed in practice. While they often enjoy the advantage of allowing for a geometric interpretation and a clean theoretical analysis, they typically lead to slower numerical implementations. Therefore statistical and/or wavelet-based methods are often preferred (see [15] and references therein). These, however, fall far outside the scope of this paper, which aimed at presenting an overview of the research activities of the past twenty years engendered by the Perona–Malik model only.

References

[1] H. Amann, Time-delayed Perona–Malik type problems, Acta Math. Univ. Comenian. (N.S.), **76** (2007), 15–38.

[2] L. Ambrosio, N. Gigli and G. Savaré, Gradient Flows: In Metric Spaces and in the Space of Probability Measures, Lectures Math. ETH Zurich, Birkhäuser Verlag, 2008.

[3] L. Ambrosio and V.M. Tortorelli, Approximation of functionals depending on jumps by elliptic functionals via Γ- convergence, Comm. Pure Appl. Math., **43** (1990), 999–1036.

[4] J. Bai and X.-C. Feng, Fractional Order Anisotropic Diffusion for Image Denoising. IEEE Trans. Image Process., **16** (2007), 2492–2502.

[5] J. D. Ball and R. D. James, Fine phase mixtures as minimizers of energy, Arch. Rational Mech. Anal., **100** (1987), 13–52.

[6] J. D. Ball and R. D. James, Proposed experimental tests of a theory if fine microstructures and the two-well problem, Philos. Trans. Roy. Soc. London Ser. A, **338** (1992), 389–450.

[7] A. Belahmidi, Equations aux dérivées partielles appliquées à la restoration et à l'agrandissement des images, Ph. D. thesis, Univ. Paris-Dauphine, Paris, 2003.

[8] A. Belahmidi and A. Chambolle, Time-delay regularization of anisotropic diffusion and image processing, M2AN Math. Model. Numer. Anal., **39** (2005), 231–251.

[9] A. Bellettini and G. Fusco, The Γ-limit and the related gradient flow for singular perturbation functionals of Perona–Malik type, Trans. Amer. Math. Soc., **360** (2008), 4929–4987.

[10] G. Bellettini, M. Novaga and M. Paolini, Convergence for long-times of a semidiscrete Perona–Malik Equation in one dimension, Math. Models Methods Appl. Sci., **21** (2011), 1–25.

[11] G. Bellettini, M. Novaga, M. Paolini and C. Tornese, Convergence of discrete schemes for the Perona–Malik equation, J. Differential Equations, **245** (2008), 892–924.

[12] G. Bellettini, M. Novaga, M. Paolini and C. Tornese, Classification of the Equilibria for the semi-discrete Perona–Malik Equation, Calcolo, **46** (2009), 221–243.

[13] A. Bertozzi and J. Greer, Low-curvature image simplifiers: Global regularity of smooth solutions and Laplacian limiting schemes, Comm. Pure Appl. Math., **57** (2004), 764–790.

[14] H. Brézis, Opérateurs Maximaux Monotones et Semi-Groups de Contractions dans les Espaces de Hilbert, Nort-Holland, Amsterdam, 1973.

[15] A. Buades, B. Coll and J. M. Morel, A review of image denoising algorithms, with a new one, Multiscale Model. Simul., **4** (2005), 490–530.

[16] J. Canny, Finding Edges and Lines in Images, Tech Report, **720**, Artificial Intelligence Laboratory, MIT, 1983.

[17] F. Catté, P.-L. Lions, J.-M. Morel and T. Coll, Image selective smoothing and edge-detection by non-linear diffusion, SIAM J. Numer. Anal., **29** (1992), 182–193.

[18] A. Chambolle, Image segmentation by variational methods: Mumford and shah functional and the discrete approximations, SIAM J. Appl. Math., **55** (1995), 827–863.

[19] T. F. Chan and L. Vese, Active contours without edges, IEEE Tans. Image Processing, **10** (2001), 266–277.

[20] Y. Chen and K. Zhang, Young measures solutions of the two-dimensional Perona–Malik equation of image processing, Comm. Pure and Appl. Math., **5** (2006), 617–637.

[21] M. Chipot and D. Kinderlehrer, Equilibrium configurations of crystals, Arch. Rational Mech. Anal., **103** (1988), 237–277.

[22] M. Colombo and M. Gobbino, Slow time behavior of the semidiscrete Perona–Malik scheme in one dimension, SIAM J. Math. Anal., **43** (2011), 2564–2600.

[23] M. Colombo and M. Gobbino, Passing to the limit in maximal slope curves: from a regularized Perona–Malik equation to the total variation flow, Math. Models Methods Appl. Sci., **22** (2012), 1250017.

[24] G.-H. Cottet and M. El Ayyadi, A volterra type model for image processing, IEEE Trans. Image Processing, **7** (1998), 292–303.

[25] G.-H. Cottet and L. Germain, Image processing through reaction combined with nonlinear diffusion, Math. Comp., **61** (1993), 659–673.

[26] B. Dacorogna, Weak continuity and weak lower semicontinuity of nonlinear functionals, Lecture Notes in Math., **922**, Springer-Verlag, 1982.

[27] B. Dacorogna and I. Fonseca, A-B-quasiconvexity and implicit partial differential equations, Calc. Var. Partial Differential Equations, **14** (2002), 115–149.

[28] E. de Giorgi, New problems on minimizing movements, In: Boundary Value Problems for PDE and Applications, RMA Res. Notes Appl. Math., **29**, Masson, Paris, 1993, pp. 81–98.

[29] E. de Giorgi, Evolution problems in metric spaces and steepest descent curves, In: Ennio de Giorgi. Selected Papers, (eds. L. Ambrosio, G. Dal Maso, M. Forti and S. Spagnolo), Springer-Verlag, 2006, pp. 527–533.

[30] E. de Giorgi, A. Marino and M. Tosques, Problemi di evoluzione in spazi metrici e curve di massima pendenza, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur., **68** (1980), 180–187.

[31] S. Demoulini, Young measure solutions for a nonlinear parabolic equation of forward-backward type, SIAM J. Math. Anal., **27** (1996), 376–403.

[32] S. Didas, B. Burgeth, A. Imiya and J. Weickert, Regularity and Scale-Space Properties of Fractional High Order Linear Filtering, In: Scale Space and PDE Methods in Computer Vision, Lecture Notes in Comput. Sci., **3459**, Springer-Verlag, 2005.

[33] S. Didas, J. Weickert and B. Burgeth, Stability and local feature enhancement of higher order nonlinear diffusion filtering, Pattern Recognition, **3663** (2005), 451–458.

[34] S. Didas, J. Weickert and B. Burgeth, Properties of higher order nonlinear diffusion filtering, J. Math. Imaging Vision, **35** (2009), 208–226.

[35] S. Esedoglu, An analysis of the perona-malik scheme, Comm. Pure Appl. Math., **54** (2001), 1442–1487.

[36] S. Esedoglu, Stability properties of the Perona–Malik scheme, SIAM J. Numer. Anal., **44** (2006), 1297–1313.

[37] M. Ghisi and M. Gobbino, Gradient estimates for the Perona–Malik equation, Math. Ann., **337** (2007), 557–590.

[38] M. Ghisi and M. Gobbino, A class of local classical solutions of Perona–Malik, Trans. Amer. Math. Soc., **361** (2009), 6429–6446.

[39] M. Ghisi and M. Gobbino, An example of global classical solution for the Perona–Malik equation, Comm. Partial Differential Equations, **36** (2011), 1318–1352.

[40] M. Ghisi and M. Gobbino, On the evolution of subcritical regions for the Perona–Malik equation, Interfaces Free Bound., **13** (2011), 105–125.

[41] M. Gobbino, Entire solutions of the one-dimensional Perona–Malik equation, Comm. Partial Differential Equations, **32** (2007), 719–743.

[42] J. B. Greer and A. L. Bertozzi, Traveling wave solutions of fourth order PDEs for image processing, SIAM J. Math. Anal., **36** (2004), 38–68 (electronic).

[43] P. Guidotti, A new nonlocal nonlinear diffusion of image processing, J. Differential Equations, **246** (2009), 4731–4742.

[44] P. Guidotti, A new well-posed nonlinear nonlocal diffusion, Nonlinear Anal., **72** (2010), 4625–4637.

[45] P. Guidotti, A backward-forward regularization of the Perona–Malik equation, J. Differential Equations, **252** (2012), 3226–3244.

[46] P. Guidotti, A family of nonlinear diffusions connecting Perona–Malik to standard diffusion, Discrete Contin. Dyn. Syst. Ser. S, **5** (2012), 581–590.

[47] P. Guidotti and J. Lambers, Two new nonlinear nonlocal diffusions for noise reduction, J. Math. Imaging Vision, **33** (2009), 25–37.

[48] P. Guidotti, J. lambers and Y. Kim, Image restoration with a new class of forward-backward-forward diffusion equations of Perona–Malik type with applications to satellite image enhancement, under review.

[49] P. Guidotti and K. Longo, Two enhanced fourth order diffusion models for image denoising, J. Math. Imaging Vision, **40** (2011), 188–198.

[50] P. Guidotti and K. Longo, Well-posedness for a class of fourth order diffusions for image processing, NoDEA Nonlinear Differential Equations Appl., **18** (2011), 407–425.

[51] M. R. Hajiaboli, A self-governing hybrid model for noise removal, In: Advances in Image and Video Technology, Lecture Notes in Comput. Sci., **5414**, Springer-Verlag, 2008.

[52] M. R. Hajiaboli, An anisotropic fourth-order partial differential equation for noise removal, In: Scale Space and Variational Methods in Computer Vision, Lecture Notes in Comput. Sci., **5567**, Springer-Verlag, 2009.

[53] B. Kawohl, Variational versus PDE-based approaches in mathematical image processing, In: Singularities in PDE and the Calculus of Variations, CRM Proc. Lecture Notes, **44**, Amer. Math. Soc., Providence, RI, 2008, pp. 113–126.

[54] B. Kawohl and N. Kutev, Maximum and comparison principle for one-dimensional anisotropic diffusion, Math. Ann., **311** (1998), 107–123.

[55] S. Kichenassamy, The Perona–Malik paradox, SIAM J. Appl. Math., **57** (1997), 1328–1342.

[56] S. Kichenassamy, The Perona–Malik method as an edge pruning algorithm, J. Math. Imaging Vision, **30** (2008), 209–219.

[57] D. Kinderlehrer and P. Pedregal, Weak convergence of integrands and the young measure representation, SIAM J. Math. Anal., **23** (1992), 1–19.

[58] M. Lysaker, A. Lundervold and X.C. Tai, Noise removal using fourth order differential equations with applications to medical magnetic resonance images in space-time, IEEE Trans. Image Process., **12** (2003), 1579–1590.

[59] D. Marr and E. Hildreth, Theory of edge detection, Proc. Roy. Soc. London, **207** (1980), 187–217.

[60] D. Mumford and J. Shah, Optimal Approximation by piecewise smooth functions and associated variational problems, Comm. Pure Appl. Math., **42** (1989), 577–685.

[61] M. Nitzberg and T. Shiota, Nonlinear image filtering with edge and corner enhancement, IEEE Trans. Pattern Anal. Machine Intelligence, **14** (1992), 826–833.

[62] P. Perona and J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. Pattern Anal. Machine Intelligence, **12** (1990), 161–192.

[63] P. Perona, T. Shiota and J. Malik, Anisotropic Diffusion, In: Geometry-Driven Diffusion in Computer Vision, (ed. B. M. ter Haar Romeny), Kluwer, Dordrecht, 1994, pp. 72–92.

[64] V. B. Surya Prasath and D. Vorotnikov, On a system of adaptive coupled PDEs for image restoration, to appear in J. Math. Imaging Vision.

[65] S. Taheri, Q. Tang and K. Zhang, Young measure solutions and instability of the one-dimensional Perona–Malik equation, J. Math. Anal. Appl., **308** (2005), 467–490.

[66] J. Tumblin and G. Turk, LCIS: A boundary hierarchy for detail-preserving contrast reduction, In: Proceedings of the SIGGRAPH 1999 Annual Conference on Computer Graphics, August 8-13, 1999, Los Angeles, CA, USA, Siggraph Annual Conference Series, ACM Siggraph, Addison-Wesley, Longman, 1999, pp. 83–90.

[67] G. Wei, Generalized Perona–Malik equation for image restoration, IEEE Signal Processing Letters, **6** (1999), 165–167.

[68] J. Weickert, A review of nonlinear diffusion filtering, In: Scale-Space Theory in Computer Vision, Lecture Notes in Comput. Sci., **1252**, Springer-Verlag, 1997, pp. 3–28.
[69] J. Weickert, Coherence-enhancing diffusion filtering, Int. J. Comput. Vis., **31** (1999), 111–127.
[70] J. Weickert, Coherence-enhancing diffusion of colour images, Image and Vision Computing, **17** (1999), 201–212.
[71] J. Weickert, Anisotropic diffusion in image processing, Ph. D. thesis, Univ. Kaiserslautern, Kaiserlautern, 1996.
[72] J. Weickert, Anisotropic Diffusion in Image Processing, ECMI Series. Teubner Verlag, Stuttgart, 1998.
[73] Y. L. You and M. Kaveh, Fourth order partial differential equations for noise removal, IEEE Trans. Image Process., **9** (2000), 1723–1730.
[74] K. Zhang, On the structure of quasi-convex hulls, Ann. Inst. H. Poincaré Anal. Non Linéaire, **15** (1998), 663–686.
[75] K. Zhang, Existence of infinitely many solutions for the one-dimensional Perona–Malik model, Calc. Var. Partial Differential Equations, **26** (2006), 171–199.

Department of Mathematics
340 Rowland Hall
Irvine, CA 92697, USA
E-mail address: gpatrick@math.uci.edu