Image Segmentation and Edge Enhancement with Stabilized Inverse Diffusion Equations

Ilya Pollak, Alan S. Willsky, *Fellow, IEEE*, and Hamid Krim, *Fellow, IEEE*

***Abstract*—We introduce a family of first-order multidimensional ordinary differential equations (ODE's) with discontinuous right-hand sides and demonstrate their applicability in image processing. An equation belonging to this family is an inverse diffusion everywhere except at local extrema, where some stabilization is introduced. For this reason, we call these equations "stabilized inverse diffusion equations" (SIDE's). Existence and uniqueness of solutions, as well as stability, are proven for SIDE's. A SIDE in one spatial dimension may be interpreted as a limiting case of a semi-discretized Perona–Malik equation [14], [15]. In an experimental section, SIDE's are shown to suppress noise while sharpening edges present in the input signal. Their application to image segmentation is also demonstrated.**

***Index Terms*—Diffusion, enhancement, scale-space, segmentation, sliding modes, synthetic aperture radar (SAR).**

I. Introduction

THE OBJECTIVE of this paper is to develop and analyze robust and fast image segmentation algorithms. They must be robust to pervasive, large-amplitude noise, which cannot be well characterized in terms of probabilistic distributions. This is because we are interested in applications such as synthetic aperture radar (SAR) segmentation in which speckle noise is a well-known problem that has defeated many algorithms. (A prototypical SAR log-magnitude image of two textural regions—forest and grass—is shown in Fig. 1.) Our methods must also be robust to blur, because many imaging techniques result in smoothed images. For example, SAR image formation has a natural blur associated with it, due to the finite aperture used in forming the image.

The algorithms we introduce are motivated by the great recent interest in using evolutions specified by partial differential equations (PDE's) as image processing procedures for tasks such as edge enhancement and segmentation, among others [1], [4], [12], [14]–[18], [21]. While the analysis of these techniques is most often performed in the continuous setting, where an image is identified with a function of two continuous spatial variables, the implementation of such equations generally involves their discrete approximation. As a consequence, as Weickert pointed out in [20], "a scale-space representation cannot perform better than its discrete realization." Following in the footsteps of his research on semi-discrete diffusions, we concentrate in this paper on semi-discrete scale spaces [i.e., continuous in scale (or time) and discrete in space]. More specifically, the main contribution of this paper is a new family of semi-discrete evolution equations which stably sharpen edges and suppress noise. The starting point for the development of these equations is a discrete interpretation of anisotropic diffusions such as that used by Perona and Malik [14], [15]. One motivation for the work in [14] and [15] is achieving both noise removal and edge enhancement through the use of an equation which, in essence, acts as an unstable inverse diffusion near edges and as a stable linear-heat-equation-like diffusion in homogeneous regions without edges. In the sense that we will make both precise and conceptually clear, the evolutions that we introduce may be viewed as a conceptually limiting case of such diffusions. These evolutions have discontinuous right-hand sides and act as inverse diffusions "almost everywhere" with stabilization resulting from the presence of the discontinuities in the vector field defined by the evolution. As we will see, the scale space of such an equation is a family of segmentations of the original image, with larger values of the scale parameter t corresponding to segmentations at coarser resolutions. Moreover, in contrast to continuous evolutions, the ones introduced here naturally define a sequence of logical "stopping times," i.e., points along the evolution endowed with useful information, and corresponding to times at which the evolution hits a discontinuity surface of its solution field.

In the next section we begin by describing a convenient mechanical analog for the visualization of many spatially-discrete evolution equations, including discretized linear or nonlinear

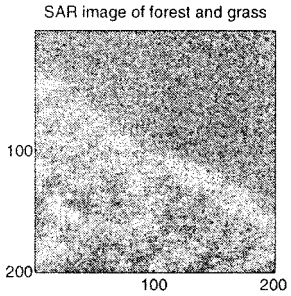


Fig. 1. SAR image of trees and grass.

Manuscript received September 4, 1997; revised October 30, 1998. This work was supported in part by AFOSR Grant F49620-98-1-0349, ONR Grant N00014-91-J-1004, and by Subcontract GC123919NGD from Boston University under the AFOSR Multidisciplinary Research Program on Reduced Signature Target Recognition. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Kannan Ramchandran.

I. Pollak is with the Division of Applied Mathematics, Brown University, Providence, RI 02912 USA (e-mail: ipollak@alum.mit.edu; WWW: http://www.cfm.brown.edu/people/pollak).

A. S. Willsky is with the Department of Electrical Engineering and Computer Science and the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

H. Krim is with the Electrical and Computer Engineering Department, North Carolina State University, Raleigh, NC 27607 USA.

Publisher Item Identifier S 1057-7149(00)01259-8.

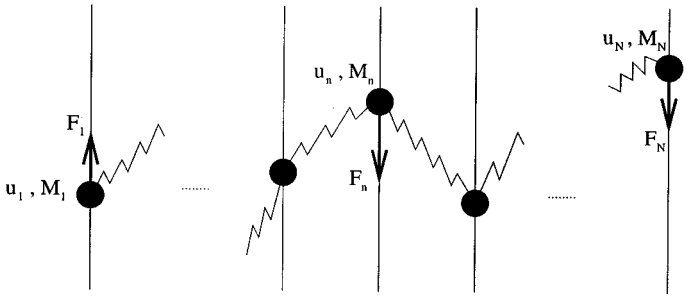


Fig. 2. Spring-mass model.

diffusions such as that of Perona and Malik, as well as the discontinuous equations that we introduce in Section III. The implementation of such a discontinuous equation naturally results in a recursive region merging algorithm. In Section IV, we point out its principal differences from Koepfler *et al.*'s [9] region-merging procedure for minimizing the Mumford–Shah functional [11]. The rest of that section is devoted to exploring the links with other important work in the field: the total variation approach [2], [16]; shock filters of Osher and Rudin [12]; the robust variational formulation of Geman and Reynolds [7]; and the stochastic modeling approach of Zhu and Mumford [22]. Because of the discontinuous right-hand side of our equations, some care must be taken in defining solutions, but as we show in Section V, once this is done, the resulting evolutions have a number of important properties. Moreover, as we have indicated, they lead to very effective algorithms for edge enhancement and segmentation, something that we demonstrate in Section VI. In particular, as we will see, they can produce sharp enhancement of edges in high noise as well as accurate segmentations of very noisy imagery such as SAR imagery subject to severe speckle.

II. A Spring-Mass Model for Certain Evolution Equations

As we indicated in the introduction, the focus of this paper is on discrete-space, continuous-time evolutions of the following general form:

$$\begin{aligned} \dot{\boldsymbol{u}}(t) &= \mathcal{F}(\boldsymbol{u})(t), \\ \boldsymbol{u}(0) &= \boldsymbol{u}_0, \end{aligned} \tag{1}$$

where $\boldsymbol{u}$ is either a discretized signal, i.e., an N-point discrete sequence ($\boldsymbol{u} = (u_1, \cdots, u_N)^T \in \mathbb{R}^N$), or an N-by-N image whose jth entry in the ith row is u_{ij} ($\boldsymbol{u} \in \mathbb{R}^{N^2}$). The initial condition $\boldsymbol{u}_0$ corresponds to the original signal or image to be processed, and $\boldsymbol{u}(t)$ then represents the evolution of this signal/image at time (scale) t, resulting in a scale-space family for $0 \le t < \infty$.

The nonlinear operators $\mathcal{F}$ of interest in this paper can be conveniently visualized through the following simple mechanical model. For the sake of simplicity in visualization, let us first suppose that $\boldsymbol{u} \in \mathbb{R}^N$ is a one-dimensional (1-D) sequence, and interpret $\boldsymbol{u}(t) = (u_1(t), \cdots, u_N(t))^T$ in (1) as the vector of vertical positions of the N particles of masses $M_1, \cdots, M_N$, depicted in Fig. 2. The particles are forced to move along N vertical lines. Each particle is connected by springs to its two neighbors (except the first and last particles, which are only connected to one neighbor). Every spring whose vertical extent is v has energy $E(v)$, i.e., the energy of the spring between the nth and $(n+1)$st particles is $E(u_{n+1} - u_n)$. We impose the usual requirements on this energy function:

$$\begin{aligned} E(v) &\ge 0, \\ E(0) &= 0, \\ E'(v) &\ge 0, \qquad \text{for } v > 0, \\ E(v) &= E(-v). \end{aligned} \tag{2}$$

Then the derivative of $E(v)$, which we refer to as "the force function" and denote by $F(v)$, satisfies

$$\begin{aligned} F(0) &= 0, \\ F(v) &\ge 0, \qquad \text{for } v > 0, \\ F(v) &= -F(-v). \end{aligned} \tag{3}$$

We also call $F(v)$ a "force function" and $E(v)$ an "energy" if $-E(v)$ satisfies (2) and $-F(v)$ satisfies (3). We make the movement of the particles non-conservative by stopping it after a small period of time Δt and restarting with zero velocity. (Note that this will make our equation non-hyperbolic.) We assume that during one such step, the total force $F_n = -F(u_n - u_{n+1}) - F(u_n - u_{n-1})$, acting on the nth particle, stays approximately constant. The displacement during one iteration is proportional to the product of acceleration and the square of the time interval:

$$u_n(t + \Delta t) - u_n(t) = \frac{(\Delta t)^2}{2} \frac{F_n}{M_n}.$$

Letting $\Delta t \to 0$, while fixing $2M_n/\Delta t = m_n$, where m_n is a positive constant, leads to

$$\begin{aligned} \dot{u}_n = \frac{1}{m_n} (F(u_{n+1} - u_n) - F(u_n - u_{n-1})), \\ n = 1, 2, \cdots, N, \end{aligned} \tag{4}$$

with the conventions $u_0 = u_1$ and $u_{N+1} = u_N$ imposed by the absence of springs to the left of the first particle and to the right of the last particle. We will refer to m_n as "the mass of the nth particle" in the remainder of the paper. Note that (4) is a (weighted) gradient descent equation for the following global energy:

$$\mathcal{E}(\boldsymbol{u}) = \sum_{i=1}^{N-1} E(u_{i+1} - u_i). \tag{5}$$

The examples below, where we set $m_n = 1$, clearly illustrate these notions.

Example 1: A linear force function $F(v) = v$ leads to the semi-discrete linear heat equation

$$\dot{u}_n = u_{n+1} - 2u_n + u_{n-1}.$$

This corresponds to a simple discretization of the 1-D linear heat equation and results in evolutions which produce increasingly low-pass filtered and smoothed versions of the original signal $\boldsymbol{u}_0$. ■

In general, we call $F(v)$ a "diffusion force" if, in addition to (3), it is monotonically increasing:

$$v_1 < v_2 \Rightarrow F(v_1) < F(v_2), \tag{6}$$

which is illustrated in Fig. 3(a). We shall call the corresponding energy a "diffusion energy" and the corresponding evolution (4) a "diffusion." The evolution in Example 1 is clearly a diffusion. We call $F(v)$ an "inverse diffusion force" if $-F(v)$ satisfies (3) and (6), as illustrated in Fig. 3(b). The corresponding evolution (4) is called an "inverse diffusion." Inverse diffusions have the characteristic of enhancing abrupt differences in $\boldsymbol{u}$ corresponding to "edges" in the 1-D sequence. Such pure inverse diffusions, however, lead to unstable evolutions (in the sense that they greatly amplify arbitrarily small noise). The following example, which is prototypical of the examples considered by Perona and Malik, defines a stable evolution that captures at least some of the edge enhancing characteristics of inverse diffusions.

Example 2: Taking $F(v) = v \exp(-(v/K)^2)$, as illustrated in Fig. 3(c), yields a 1-D semi-discrete (continuous in scale and discrete in space) version of the Perona–Malik equation [15, Eqs. (3.3), (3.4), and (3.12)]. In general, given a positive constant K, we shall call a force $F(v)$ a "Perona–Malik force of thickness K" if, in addition to (3), it satisfies the following conditions:

$$\begin{aligned} &F(v) \text{ has a unique maximum at } v = K, \\ &F(v_1) = F(v_2) \Rightarrow (|v_1| - K)(|v_2| - K) < 0. \end{aligned} \tag{7}$$

We shall call the corresponding energy a "Perona–Malik energy" and the corresponding evolution equation a "Perona–Malik equation of thickness K." As Perona and Malik demonstrate (and as can also be inferred from our results), evolutions with such a force function act like inverse diffusions in the regions of high gradient and like usual diffusions elsewhere. They are stable and capable of achieving some level of edge enhancement depending on the exact form of $F(v)$. ■

Finally, to extend our mechanical model of Fig. 2 to images, we simply replace the sequence of vertical lines along which the particles move with an N-by-N square grid of such lines. The particle at location (i, j) is connected by springs to its four neighbors: $(i-1, j)$, $(i, j+1)$, $(i+1, j)$, $(i, j-1)$, except for the particles in the four corners of the square (which only have two neighbors each), and the rest of the particles on the boundary of the square (which have three neighbors). This arrangement is reminiscent of (and, in fact, was suggested by) the resistive network of [14, Fig. 8]. The analog of (4) for images is then

$$\begin{aligned} \dot{u}_{ij} = \frac{1}{m_{ij}} &(F(u_{i+1,j} - u_{ij}) - F(u_{ij} - u_{i-1,j}) \\ &+ F(u_{i,j+1} - u_{ij}) - F(u_{ij} - u_{i,j-1})), \end{aligned} \tag{8}$$

with $i = 1, 2, \cdots, N, \quad j = 1, 2, \cdots, N,$ and the conventions $u_{0,j} = u_{1,j}$, $u_{N+1,j} = u_{N,j}$, $u_{i,0} = u_{i,1}$ and $u_{i,N+1} = u_{i,N}$ imposed by the absence of springs outside of $1 \le i \le N, 1 \le j \le N$.

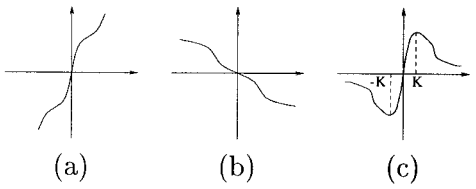


Fig. 3. Force functions: (a) diffusion, (b) inverse diffusion, and (c) Perona–Malik.

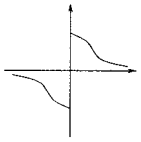


Fig. 4. Force function for a stabilized inverse diffusion equation.

III. Stabilized Inverse Diffusion Equations (SIDE's): The Definition

In this section, we introduce a discontinuous force function, resulting in a system (4) that has discontinuous right-hand side (RHS). Such equations received much attention in control theory because of the wide usage of relay switches in automatic control systems. More recently, deliberate introduction of discontinuities has been used in control applications to drive the state vector onto lower-dimensional surfaces in the state space [19]. As we will see, this objective of driving a trajectory onto a lower-dimensional surface also has value in image analysis and in particular in image segmentation. Segmenting a signal or image, represented as a high-dimensional vector $\boldsymbol{u}$, consists of evolving it so that it is driven onto a comparatively low-dimensional subspace which corresponds to a segmentation of the signal or image domain into a small number of regions.

The type of force function of interest to us here is illustrated in Fig. 4. More precisely, we wish to consider force functions $F(v)$ which, in addition to (3), satisfy the following properties:

$$\begin{aligned} &F'(v) \le 0 \qquad \text{for } v \ne 0, \\ &F(0^+) > 0 \\ &F(v_1) = F(v_2) \Leftrightarrow v_1 = v_2. \end{aligned} \tag{9}$$

Contrasting this form of a force function to the Perona–Malik function in Fig. 3, we see that in a sense one can view the discontinuous force function as a limiting form of the continuous force function in Fig. 3(c). However, because of the discontinuity at the origin of the force function in Fig. 4, there is a question of how one defines solutions of (4) for such a force function. Indeed, if (4) evolves toward a point of discontinuity of its RHS, the value of the RHS of (4) apparently depends on the direction from which this point is approached [because $F(0^+) \ne F(0^-)$], making further evolution nonunique. We therefore need a special definition of how the trajectory of our evolution proceeds at these discontinuity points.[1] For this definition to be useful, the resulting evolution must satisfy well-posedness properties: the existence and uniqueness of solutions, as well as stability of solutions with respect to the ini-

[1]Having such a definition is crucial because, as we will show in Section V, (4) will reach a discontinuity point of its RHS in finite time, starting with any initial condition.

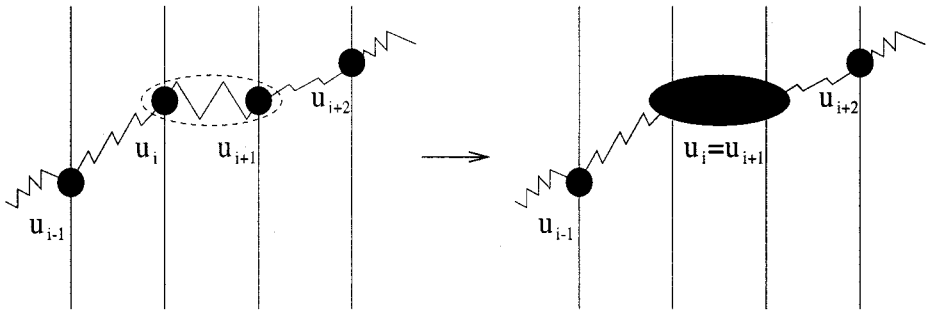


Fig. 5. A horizontal spring is replaced by a rigid link.

tial data. In the rest of this section we describe how we define solutions to (4) for force functions (9). Assuming the resulting evolutions to be well-posed, we demonstrate that they have the qualitative properties we desire, namely that they both are stable and also act as inverse diffusions and hence enhance edges. We address the issue of well-posedness and other properties in Section V.

Consider the evolution (4) with $F(v)$ as in Fig. 4 and (9) and with all of the masses m_n equal to 1. Notice that the RHS of (4) has a discontinuity at a point $\boldsymbol{u}$ if and only if $u_i = u_{i+1}$ for some i between 1 and $N-1$. It is when a trajectory reaches such a point $\boldsymbol{u}$ that we need the following definition. In terms of our spring-mass model of Fig. 2, once the vertical positions u_i and u_{i+1} of two neighboring particles become equal, the spring connecting them is replaced by a rigid link. In other words, the two particles are simply merged into a single particle which is twice as heavy (Fig. 5), yielding the following modification of (4) for $n = i$ and $n = i+1$:

$$\dot{u}_i = \dot{u}_{i+1} = \tfrac{1}{2}\left(F(u_{i+2} - u_{i+1}) - F(u_i - u_{i-1})\right).$$

(The differential equations for $n \neq i$, $i+1$ do not change.) Similarly, if m consecutive particles reach equal vertical position, they are merged into one particle of mass m ($1 \leq m \leq N$):

$$\begin{aligned}\dot{u}_n &= \cdots = \dot{u}_{n+m-1} \\ &= \frac{1}{m}\left(F(u_{n+m} - u_{n+m-1}) - F(u_n - u_{n-1})\right) \\ &\qquad \text{if } u_{n-1} \neq u_n = u_{n+1} = \cdots \\ &\qquad\quad = u_{n+m-2} = u_{n+m-1} \neq u_{n+m}. \end{aligned} \tag{10}$$

Notice that this system is the same as (4), but with possibly unequal masses. It is convenient to rewrite this equation so as to explicitly indicate the reduction in the number of state variables:

$$\begin{aligned}&\dot{u}_{n_i} = \frac{1}{m_{n_i}}\left(F(u_{n_{i+1}} - u_{n_i}) - F(u_{n_i} - u_{n_{i-1}})\right), \\ &u_{n_i} = u_{n_i+1} = \cdots = u_{n_i+m_{n_i}-1}, \\ &\qquad \text{where } i = 1, \cdots p, \\ &\qquad 1 = n_1 < n_2 < \cdots < n_{p-1} < n_p \leq N, \\ &\qquad n_{i+1} = n_i + m_{n_i}. \end{aligned} \tag{11}$$

The compound particle described by the vertical position u_{n_i} and mass m_{n_i} consists of m_{n_i} unit-mass particles u_{n_i}, $u_{n_i+1}, \cdots, u_{n_i+m_{n_i}-1}$ that have been merged, as shown in Fig. 5. The evolution can then naturally be thought of as a sequence of stages: during each stage, the right-hand side of (11) is continuous. Once the solution hits a discontinuity surface of the right-hand side, the state reduction and reassignment of m_{n_i}'s, described above, takes place. The solution then proceeds according to the modified equation until it hits the next discontinuity surface, etc. This definition of a solution is reminiscent of the pixel-grouping method used by Bouman and Sauer in [2] for solving a nondifferentiable optimization problem.

Notice that such an evolution automatically produces a multiscale segmentation of the original signal if we view each compound particle as a region of the signal. Viewed as a segmentation algorithm, our evolution can be summarized as follows.

1) Start with the trivial initial segmentation: each sample is a distinct region.
2) Evolve (11) until the values in two or more neighboring regions become equal.
3) Merge the neighboring regions whose values are equal.
4) Go to step 2.

The same algorithm can be used for 2-D images, which is immediate upon rewriting (11):

$$\dot{u}_{n_i} = \frac{1}{m_{n_i}} \sum_{n_j \in A_{n_i}} F(u_{n_j} - u_{n_i}) p_{ij}, \tag{12}$$

where

- m_{n_i} mass of the compound particle n_i (= number of pixels in the region n_i);
- A_{n_i} set of the indexes of all the neighbors of n_i, i.e., of all the compound particles that are connected to n_i by springs;
- p_{ij} number of springs between regions n_i and n_j (always 1 in 1-D, but can be larger in 2-D).

Just as in 1-D, two neighboring regions n_1 and n_2 are merged by replacing them with one region n of mass $m_n = m_{n_1} + m_{n_2}$ and the set of neighbors $A_n = A_{n_1} \cup A_{n_2} \backslash \{n_1, n_2\}$.

We close this section by describing one of the basic and most important properties of these evolutions, namely that the evolution is stable but nevertheless behaves like an inverse diffusion. Notice that a force function $F(v)$ satisfying (9) can be represented as the sum of an inverse diffusion force $F_{id}(v)$ and a positive multiple of sign(v): $F(v) = F_{id}(v) + C\,\text{sign}(v)$, where $C = F(0^+)$ and $-F_{id}(v)$ satisfies (3) and (6). Therefore, if $u_{n_{i+1}} - u_{n_i}$ and $u_{n_i} - u_{n_{i-1}}$ are of the same sign

[which means that u_{n_i} is not a local extremum of the sequence $(u_{n_1}, \cdots, u_{n_p})$], then (11) can be written as

$$\dot{u}_{n_i} = \frac{1}{m_{n_i}} \left(F_{id}(u_{n_{i+1}} - u_{n_i}) - F_{id}(u_{n_i} - u_{n_{i-1}})\right). \quad (13)$$

If $u_{n_i} > u_{n_{i+1}}$ and $u_{n_i} > u_{n_{i-1}}$ (i.e., u_{n_i} is a local maximum), then (11) is

$$\dot{u}_{n_i} = \frac{1}{m_{n_i}} \left(F_{id}(u_{n_{i+1}} - u_{n_i}) - F_{id}(u_{n_i} - u_{n_{i-1}}) - 2C\right). \quad (14)$$

If $u_{n_i} < u_{n_{i+1}}$ and $u_{n_i} < u_{n_{i-1}}$ (i.e., u_{n_i} is a local minimum), then (11) is

$$\dot{u}_{n_i} = \frac{1}{m_{n_i}} \left(F_{id}(u_{n_{i+1}} - u_{n_i}) - F_{id}(u_{n_i} - u_{n_{i-1}}) + 2C\right). \quad (15)$$

Equation (13) says that the evolution is a pure inverse diffusion at the points which are not local extrema. It is not, however, a *global* inverse diffusion, since pure inverse diffusions drive local maxima to $+\infty$ and local minima to $-\infty$ and thus are unstable. In contrast, (14) and (15) show that at local extrema, our evolution is an inverse diffusion plus a stabilizing term which guarantees that the local maxima do not increase and the local minima do not decrease. For this reason, we call the new evolution (11), (12) a "stabilized inverse diffusion equation" ("SIDE"), a force function satisfying (9) a "SIDE force," and the corresponding energy a "SIDE energy."

IV. Related Approaches

A. Mumford–Shah, Geman–Reynolds, and Zhu–Mumford

The SIDE force function $F(v)$ defined in the previous section and illustrated in Fig. 4 results in the SIDE spring energy $E(v)$ which is concave everywhere except at zero and nondifferentiable at zero, and which looks like Υ [for example, $\arctan(v)$ or $1 - (1 + |v|)^{-1}$]. In this case, the global energy (5) is similar to the first term of the image restoration model of D. Geman and Reynolds [7], as well as to Zhu and Mumford's potential function [22]. As we will show in the experimental section, this not only leads to sharp segmentations, but also allows our method to be more robust to heavy-tailed noise than algorithms which use quadratic energy terms. Since our equations are implemented via recursive region merging, it is instructive to compare them with other recursive region merging algorithms, such as Koepfler, Lopez, and Morel's [9] implementation of Mumford–Shah [11] segmentation, in which merging of neighboring regions occurs if it reduces the energy $\mathcal{E}_{\mathrm{MS}}(\boldsymbol{u}) = (\boldsymbol{u} - \boldsymbol{u}_0)^T(\boldsymbol{u} - \boldsymbol{u}_0) + \lambda l$. (Here l is the total length of the boundaries, and λ is a scale parameter: a larger λ imposes greater penalty on boundaries, which results in a coarser segmentation $\boldsymbol{u}$.) The first term of this functional makes it nonrobust to noise outliers. This term quadratically penalizes the difference between the initial image and its approximation $\boldsymbol{u}$, thereby causing very large outliers present in the original image $\boldsymbol{u}_0$ to reappear in $\boldsymbol{u}$, even for large values of λ. The absence of such a term from the SIDE energy allows the evolution to diffuse strong bursts of noise, making it robust.

Since the exhaustive survey of variational models in image processing is beyond the scope of this paper, we defer to the much more complete bibliography in [10]. In particular, [10, Ch. 3] contains a very nice discussion of region merging segmentation algorithms, starting with Brice and Fennema's [3] and Pavlidis' [13], which may be considered as ancestors to both [9], the snakes [8], and our SIDE's.

B. Shock Filters and Total Variation

Replacing the discrete vector $\boldsymbol{u}(t)$ with a function $u(t, x)$ of a continuous spatial variable x, and replacing first differences with derivatives in (4), we see that, for $m_n = 1$, (4) is a discretization of $u_t = \partial/\partial x[F(u_x)]$, where letter indexes denote corresponding partial derivatives. Expanding the SIDE force function again as $F(v) = F_{id}(v) + C\,\mathrm{sign}(v)$, we obtain:

$$u_t = C\,\frac{\partial}{\partial x}\left[\mathrm{sign}(u_x)\right] + F'_{id}(u_x)u_{xx}. \quad (16)$$

The first of the RHS terms is the 1-D version of the gradient descent on total variation, proposed independently by Bouman and Sauer for edge-preserving tomographic reconstruction [2] and by Rudin, Osher, and Fatemi for image restoration [16]. It has very good noise removal properties, but, if used alone, it will ultimately blur the signal. If $F_{id}(v) = -(1/2)\,v|v|$, then the second term is equal to the RHS of one of the shock filters considered by Osher and Rudin in [12]. It is excellent for edge enhancement, but, as mentioned in [12], it cannot remove noise (and, in fact, if our simple discretization scheme is used, it will be unstable and noise-enhancing). Thus, our equation combines the noise-suppressive properties of the total variation approach with the edge-sharpening features of shock filters.

Using the system of ODE's (4) as a starting point for our investigations, rather than PDE (16), is due to the variety of interpretations (16) may have. Indeed, the discontinuity of the RHS leads one to interpret the equation in the weak sense, while no theory of weak solutions of such an equation is known to us. Moreover, the numerical solutions depend on the discretization scheme employed—which is also true of the Perona–Malik equation. In contrast, much more can be said about (4)—in particular, there are no difficulties in defining what one means by a "solution." An interesting theoretical question, which remains on our agenda, is on the relations of the solutions of SIDE's and those of (16).

V. Properties of SIDE's

The SIDE's described in the two preceding sections enjoy a number of interesting properties which validate and explain their adaptability to segmentation problems. We first examine the SIDE's in one spatial dimension for which we can make the strongest statements.

We define the n_ith discontinuity hyperplane of a SIDE (11) by $S_{n_i} = \{\boldsymbol{u} \in \mathbb{R}^p\colon u_{n_i} = u_{n_{i+1}}\}$, $i = 1, \cdots, p-1$. Sometimes it is more convenient to work with the vector $\boldsymbol{v} = (v_{n_1}, \cdots, v_{n_{p-1}})^T \in \mathbb{R}^{p-1}$ of the first differences of $\boldsymbol{u}$: $v_{n_i} = u_{n_{i+1}} - u_{n_i}$, for $i = 1, \cdots, p-1$. We abuse notation by also denoting $S_{n_i} = \{\boldsymbol{v} \in \mathbb{R}^{p-1}\colon v_{n_i} = 0\}$.

On such hyperplanes, we defined the solution of a SIDE as the solution to a modified, lower-dimensional, equation whose RHS is continuous on S_{n_i}. In what follows, we will assume that

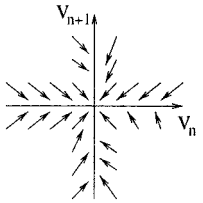


Fig. 6. Solution field near discontinuity surfaces.

the SIDE force function $F(v)$ is sufficiently regular away from zero, so that the ODE (11), restricted to the domain of continuity of its RHS, is well-posed. As a result, existence and uniqueness of solutions of SIDE's immediately follow from the existence and uniqueness of solutions of ODE's with continuous RHS. Continuous dependence on the initial data is also guaranteed for a trajectory segment lying inside a region of continuity of the RHS. In order to show, however, that the solutions that we have defined are continuous with respect to initial conditions over *arbitrary* time intervals, we must take into account the presence of discontinuities on the RHS. In particular, what we must show is that trajectories that start very near a discontinuity surface remain close to one that starts on the surface. More precisely, we need to be able to show that a trajectory whose initial point is very close to S_{n_i} will, in fact, hit S_{n_i}. In the literature on differential equations and control theory [5], [19], the behavior that our differential equations exhibit is referred to as "sliding modes." Specifically, as proven in the Appendix, the behavior of our evolution near discontinuity hyperplanes satisfies the following:

Lemma on Sliding: Let σ be a permutation of $(n_1, \cdots, n_{p-1})$, and m an integer between 1 and $p-1$, and let

$$S = \bigcap_{q=1}^{m} S_{\sigma(q)} \setminus \left(\bigcup_{q=m+1}^{p-1} S_{\sigma(q)} \right).$$

Then, as $\boldsymbol{v}$ approaches S from any quadrant,[2] we have:

$$\lim(\dot{v}_{\sigma(q)} \operatorname{sign}(v_{\sigma(q)})) \leq 0 \qquad \text{for } q = 1, \cdots, m,$$

and for at least one q this inequality is strict. ■

Intuitively, and as illustrated in Fig. 6, this lemma states that the solution field of our equation near any discontinuity surface points toward that surface. As a consequence, a trajectory which hits such a surface may be continuously extended to "slide" along the surface, as shown in [5] and [19]. For this reason the discontinuity surfaces are commonly referred to as "sliding surfaces." In our case, a simple calculation verifies that the dynamics along such a surface, obtained through any of the three classical definitions in [5] and [19], correspond exactly to the definition given in the preceding section.

The lemma on sliding, together with the well-posedness of SIDE's inside their continuity regions, directly implies the overall well-posedness of 1-D SIDE's: for finite T, the trajectory from $t = 0$ to $t = T$ depends continuously on its initial point. As shown in Property 2 to follow, a SIDE reaches a steady state in finite time, which establishes its well-posedness for infinite time intervals.

We call u_{n_i}, with $i \in \{2, \cdots, p-1\}$ a local maximum (minimum) of the sequence $(u_{n_1}, \cdots, u_{n_p})$ if $u_{n_i} > u_{n_{i\pm1}}$ ($u_{n_i} < u_{n_{i\pm1}}$). The point u_{n_1} is a local maximum (minimum) if $u_{n_1} > u_{n_2}$ ($u_{n_1} < u_{n_2}$); u_{n_p} is a local maximum (minimum) if $u_{n_p} > u_{n_{p-1}}$ ($u_{n_p} < u_{n_{p-1}}$). Similarly, a region of a 2-D image is a local maximum (minimum) if its value is larger (smaller) than the values of its neighbors. Rephrasing this definition in terms of our spring-mass model, a maximum (minimum) is a particle with all its attached springs directed downward (upward). Therefore, we immediately have [as we saw in (14) and (15)] that the maxima (minima) are always pulled down (up):

Property 1—Maximum Principle: Every local maximum is decreased and every local minimum is increased by a SIDE. Therefore,

$$|u_i(t)| < \max_n |u_n(0)| \text{ for } t > 0. \tag{17}$$

Using this result, we can prove the following.

Property 2—Finite Evolution Time: A SIDE, started at $\boldsymbol{u}_0 = (u_{0,1}, \cdots, u_{0,N})^T$, reaches its equilibrium (i.e., the point $\boldsymbol{u} = (u_1, \cdots, u_N)^T$ where $u_1 = \cdots = u_N = 1/N \sum_{i=1}^{N} u_{0,i}$) in finite time.

Proof: The sum of the vertical positions of all unit-mass particles is equal to the sum of the vertical positions of the compound particles, weighted by their masses: $\sum_{n=1}^{N} u_n = \sum_{i=1}^{p} u_{n_i} m_{n_i}$. The time derivative of this quantity is zero, as verified by summing up the right-hand sides of (11). Therefore, the mean vertical position $(1/N) \sum_{n=1}^{N} u_n$ is constant throughout the evolution. Writing (11) for $i = 1$, $\dot{u}_{n_1} = (1/m_{n_1}) F(u_{n_2} - u_{n_1})$, we see that the leftmost compound particle is stationary only if $p = 1$, i.e., if all unit-mass particles have the same vertical position: $u_{n_1} = u_1 = u_2 = \cdots = u_N$. Since the mean is conserved, the unique steady state is $u_1 = \cdots = u_N = (1/N) \sum_{i=1}^{N} u_{0,i}$. To prove that it is reached in finite time, we use the fact that a SIDE force function assigns larger force to shorter springs. If we put $L = 2 \max_n |u_n(0)|$, then the maximum principle implies that in our system there cannot exist a spring with vertical extent larger than L at any time during the evolution. Therefore, the rate of decrease of the absolute maximum, according to (11), is at least $F(L)/N$ [because $F(L)$ is the smallest force possible in the system, and N is the largest mass]. Similarly, the absolute minimum always increases at least as quickly. They will meet no later than at $t = LN/(2F(L))$, at which point the sequence $\boldsymbol{u}(t)$ must be a constant sequence. ■

The above property allows us immediately to state the well-posedness results as follows.

Property 3—Well-Posedness: For any initial condition $\boldsymbol{u}_0^*$, a SIDE has a unique solution $\boldsymbol{u}^*(t)$ satisfying $\boldsymbol{u}^*(0) = \boldsymbol{u}_0^*$. Moreover, for any such $\boldsymbol{u}_0^*$ and any $\varepsilon > 0$, there exists a $\delta > 0$ such that $|\boldsymbol{u}_0 - \boldsymbol{u}_0^*| \leq \delta$ implies $|\boldsymbol{u}(t) - \boldsymbol{u}^*(t)| \leq \varepsilon$ for $t \geq 0$, where $\boldsymbol{u}(t)$ is the solution of the SIDE with the initial condition $\boldsymbol{u}_0$. ■

As we pointed out in the introductory section, a SIDE evolution defines a natural set of hitting times which intuitively should be of use in characterizing features in an image. For this

[2] In $\mathbb{R}^{p-1}$, a quadrant containing a vector $\boldsymbol{a} = (a_1, \cdots, a_{p-1})^T$ such that $a_i \neq 0$ for $i = 1, \cdots, p-1$ is the set $Q = \{\boldsymbol{b} \in \mathbb{R}^{p-1}: b_i a_i > 0 \text{ for } i = 1, \cdots, p-1\}$.

to be true, however, we would need some type of continuity of this hitting time sequence. Specifically, let $t_n(\boldsymbol{u}_0)$ denote the "nth hit time," i.e., the time when the solution starting at $\boldsymbol{u}_0$ reaches the sliding hyperplane S_n. By Property 2, this is a finite number. Let $\boldsymbol{u}(t)$ be "a typical solution" if it never reaches two different sliding hyperplanes at the same time: $t_i(\boldsymbol{u}(0)) \neq t_j(\boldsymbol{u}(0))$ if $i \neq j$. One of the consequences of the Lemma on Sliding is that a trajectory that hits a single hyperplane S_n does so transversally (that is, cannot be tangent to it). Since trajectories vary continuously, this means that nearby solutions also hit S_n. Therefore, for typical solutions the following holds.

Property 4—Stability of Hit Times: If $\boldsymbol{u}(t)$ is a typical solution, all solutions with initial data sufficiently close to $\boldsymbol{u}(0)$ get onto surfaces S_n in the same order as $\boldsymbol{u}(t)$. ■

The sequence in which a trajectory hits surfaces S_n is an important characteristic of the solution. Property 4 says that, for a typical solution $\boldsymbol{u}(t)$, the (strict) ordering of hit times $t_n(\boldsymbol{u}(0))$ is stable with respect to small disturbances in $\boldsymbol{u}(0)$:

$$t_{n_1}(\boldsymbol{u}(0)) < t_{n_2}(\boldsymbol{u}(0)) < \cdots < t_{n_{N-1}}(\boldsymbol{u}(0)) \qquad (18)$$

where $(n_1, \cdots, n_{N-1})$ is a permutation of $(1, \cdots, N-1)$. For the purposes of segmentation and edge detection, the only interesting output occurs at these $N-1$ time points, since they are the only instants when the segmentation of the initial signal changes (i.e., when regions are merged and edges are erased). While a thorough investigation of how to use these hitting times and in particular how to stop a SIDE so as to obtain the best segmentation is an open one, the fact that our choice is limited to a finite set of time points provides us with both a natural sequence of segmentations of increasing granularity and with, at the very least, some simple stopping rules. For example, if the number of "useful" regions, r, is known or bounded *a priori*, a natural candidate for a stopping time would be $t_{n_{N-r}}$, i.e., the time when exactly r regions remain. In the next section we illustrate the effectiveness of such a rule in the simplest case, namely when $r = 2$ so that we are seeking a partition of the field of interest into two regions. These results together with the properties described here provide ample motivation for a more detailed examination of the properties of the sequence of segmentations produced by a SIDE flow. Such an investigation is currently ongoing.

We already mentioned that our definition of solutions on sliding surfaces for SIDE's in one spatial dimension coincides with all three classical definitions of solutions for a general equation with discontinuous right-hand side, which are presented in Filippov's book [5, pp. 50–56]. We use a result of [5, p. 95] to infer the following.

Property 5—Continuous Dependence on the RHS: Let us consider a SIDE force function $F_S(v)$, and let $p_K(v)$ be a smoothing kernel of width K:

$$p_K(v) \geq 0, \qquad \text{supp}(p_K) = [-K; K], \qquad \int p_K(v)\, dv = 1.$$

Let $F_K(v) = \int F_S(w) p_K(v-w)\, dw$ be a regularized version of $F_S(v)$. Consider system (4) with $m_n = 1$ and $F(v) = F_K(v)$. Then for any ε, there is a K such that the solution of this system stays closer than ε to the solution of the SIDE with the same initial condition and force $F_S(v)$. ■

We note that if the smoothing kernel $p_K(v)$ is appropriately chosen, then the resulting $F_K(v)$ will be a Perona–Malik force function of thickness K. (For example, one easy choice for $p_K(v)$ is a multiple of the indicator function of the interval $[-K; K]$.) Thus, semi-discrete Perona–Malik evolutions with small K are regularizations of SIDE's, and consequently a SIDE in 1-D can be viewed as a limiting case of a Perona–Malik-type evolution. However, as we will see in the next section, the SIDE evolutions appear to have some advantages over such regularized evolutions even in 1-D.

As we stated at the start of this section, the analysis and properties we have just derived have focused on SIDE's for 1-D signals. Let us close this section by commenting on the properties of SIDE's in 2-D. The existence and uniqueness of solutions again follow easily from our construction of solutions. Property 1 (the maximum principle) is easily inferred from the 2-D spring-mass model. Property 2 (finite evolution time) also carries over, with the same proof. There is, however, no analog of the lemma on sliding in 2-D: it is easy to show that the solutions in the vicinity of a discontinuity hyperplane of (12) do not necessarily slide onto that hyperplane. Therefore, there is no global continuous dependence on the initial data. In particular, the sequence of hitting times and associated discontinuity planes does not depend continuously on initial conditions, and our SIDE evolution does not correspond to a limiting form of a Perona–Malik evolution in 2-D but in fact represents a decidedly different type of evolutionary behavior. Several factors, however, indicate the value of this new evolution and also suggest that a weaker stability result can be proven. First of all, as shown in the experimental results in the next section, SIDE's can produce excellent segmentations in 2-D images even in the presence of considerable noise. Moreover, thanks to the maximum principle, excessively wild behavior of solutions is impossible, something that is again confirmed by the experiments of the next section. Consequently, the sequence of hit times (18) does not seem to be very sensitive to the initial condition in that the presence of noise, while perhaps perturbing the ordering of hitting times and the sliding planes that are hit, seem to introduce perturbations that are, in some sense, "small." We are currently working on defining an appropriate metric on such hitting plane/time sequences that captures this behavior and that allows us to characterize stability that SIDE's display in the experiments described next.

VI. Experiments

Choosing a SIDE force function best suited for a particular application is an open research question. For the examples below, we use a very simple, piecewise-linear force function $F(v) = \text{sign}(v) - v/L$. Note that, formally, this function does not satisfy our definition (3) of a force function, since it is negative for $v > L$. Therefore, in our experiments we always make sure that L is larger than the dynamic range of the signal or image to be processed. In that case, thanks to the maximum principle, we will have $|u_i(t) - u_j(t)| < L$ for

any pair of pixels at any time t during evolution, and therefore $F(|u_i(t) - u_j(t)|) > 0$.

A. Experiment 1: 1-D Unit Step in High Noise Environment

We first test this SIDE on a unit step function corrupted by additive white Gaussian noise whose standard deviation is equal to the amplitude of the step, and which is depicted in Fig. 7(a). The remaining parts of this figure display snapshots of the SIDE evolution starting with the noisy data in Fig. 7(a), i.e., they correspond to the evolution at a selected set of hitting times. The particular members of the scale space which are illustrated are labeled according to the number of remaining regions. Note that the last remaining edge, i.e., the edge in Fig. 7(d) for the hitting time at which there are only two regions left, is located between samples 96 and 97, which is quite close to the position of the original edge (between the 100th and 101st samples). In this example, the step in Fig. 7(d) also has amplitude that is close to that of the original unit step. In general, thanks to the stability of SIDE's, the sizes of discontinuities will be diminished through such an evolution, much as they are in other evolution equations. However, from the perspective of segmentation this is irrelevant, i.e., the focus of attention is on detecting and locating the edge, not on estimating its amplitude.

This example also provides us with the opportunity to contrast the behavior of a SIDE evolution with a Perona–Malik evolution and in fact to describe the behavior that originally motivated our work. Specifically, as we noted in the discussion of Property 5 of the previous section, a SIDE in 1-D can be approximated with a Perona–Malik equation of a small thickness K. Observe that a Perona–Malik equation of a large thickness K will diffuse the edge before removing all the noise. Consequently, if the objective is segmentation, the desire is to use as small a value of K as possible. Following the procedure prescribed by Perona, Shiota, and Malik in [15], we computed the histogram of the absolute values of the gradient throughout the initial signal, and fixed K at 90% of its integral. The resulting evolution is shown in Fig. 8. In addition to its good denoising performance, it also blurs the edge, which is clearly undesirable if the objective is a sharp segmentation. The comparison of Figs. 7 and 8 strongly suggests that the smaller K the better. It was precisely this observation that originally motivated the development of SIDE's. However, while in 1-D a SIDE evolution can be viewed precisely as a limit of a Perona–Malik evolution as K goes to 0, there is still an advantage to using the form of the evolution that we have described rather than a Perona–Malik evolution with a very small value of K. Specifically, the presence of explicit reductions in dimensionality during the evolution makes a SIDE implementation more efficient than that described in [15]. Even for this simple example the Perona–Malik evolution that produced the result comparable to that in Fig. 7 evolved approximately five times more slowly than our SIDE evolution. Although a SIDE in 2-D cannot be viewed as a limit of Perona–Malik evolutions, the same comparison in speed of evolution is still true, although in this case the difference in computation time can be orders of magnitude.

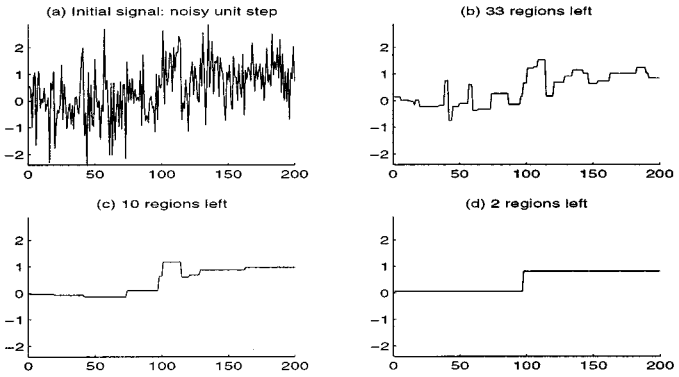


Fig. 7. Scale space of a SIDE for a noisy unit step at location 100: (a) the original signal and (b)–(d) representatives of the resulting SIDE scale space.

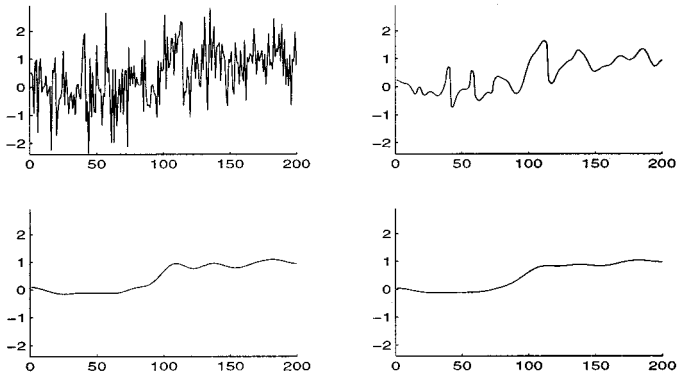


Fig. 8. Scale space of a Perona–Malik equation with a large K for the noisy step of Fig. 7.

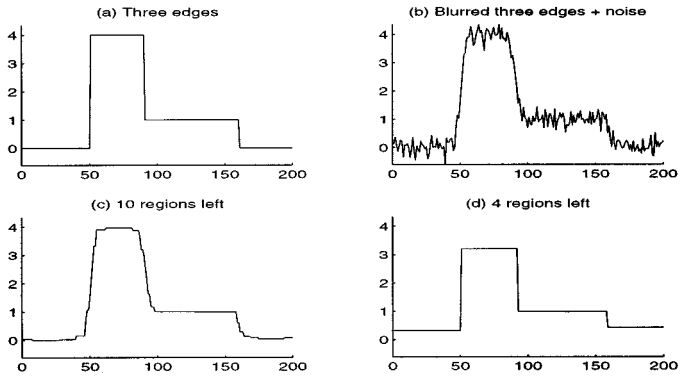


Fig. 9. Scale space of a SIDE for a noisy blurred three-edge staircase: (a) noise-free original signal, (b) its noisy realization, and (c), (d) representatives of the resulting SIDE scale space.

B. Experiment 2: Edge Enhancement in 1-D

Our second 1-D example shows that SIDE's can stably enhance edges. The staircase signal in the upper left-hand corner of Fig. 9 was convolved with a Gaussian and corrupted by additive noise. The evolution was stopped when there were only four regions (three edges) left. The locations of the edges are very close to those in the original signal. (As in the previous example, the amplitudes of the final signal are quite different

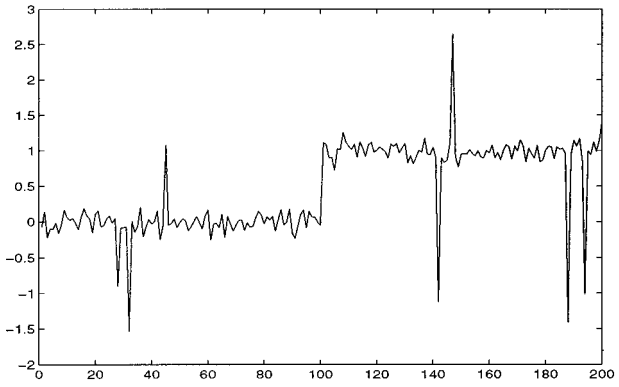


Fig. 10. A unit step with heavy-tailed noise.

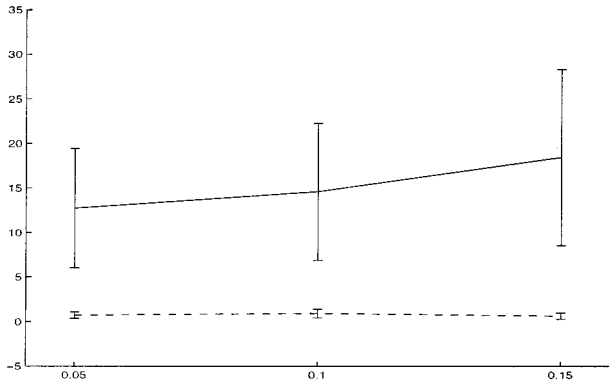


Fig. 11. RMS errors for Monte-Carlo runs. (Koepfler–Lopez–Morel: solid line; SIDE: broken line).

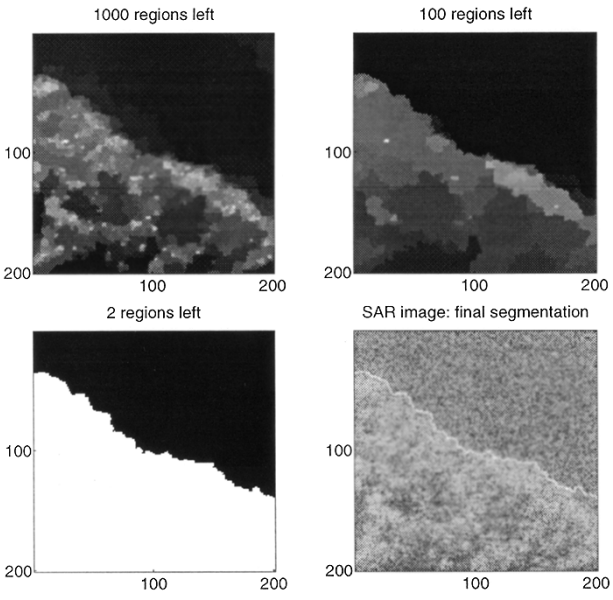


Fig. 12. Scale space of a SIDE for the SAR image of trees and grass, and the final boundary superimposed on the initial image.

from those of the initial condition. This is immaterial, since we are interested in segmentation, not in restoration.)

C. Experiment 3: Robustness in 1-D

We now compare robustness of our algorithm to that of Koepfler, Lopez, and Morel's [9] minimization of Mumford–Shah functional [11], as implemented in MegaWave2 software package from CEREMADE [6]. For that purpose, we use Monte Carlo simulations on a unit step signal corrupted by "heavy-tailed" noise which is, with high probability $1-\varepsilon$, normally distributed with $\sigma = 0.1$, and, with low probability ε, normally distributed with $\sigma = 1$. A typical sample path, for $\varepsilon = 0.05$, is shown in Fig. 10. During each trial, each algorithm was stopped when only two regions remained, and the resulting jump location was taken as the output. The root mean squared errors in locating the jump for $\varepsilon = 0.05$, $\varepsilon = 0.1$, and $\varepsilon = 0.15$ are shown in Fig. 11 (the solid line is Koepfler–Lopez–Morel, the broken line is SIDE). The error bars are ± two standard deviations.

As anticipated, the quadratic term of the Mumford–Shah energy makes it nonrobust to heavy-tailed noise, and the performance degrades considerably as the contamination probability increases. At the same time, our method is very robust, even if the outlier probability is as high as 0.15.

Note that neither of the two algorithms is optimal for this simple 1-D problem, for which near-perfect results can be achieved in a computationally efficient manner by very simple procedures. The purpose of including this example is to provide statistical evidence for our claim of robustness of SIDE's. This becomes very important for complicated 2-D problems, such as the one considered in the next example, where simple techniques no longer work.

D. Experiment 4: SIDE Evolutions in 2-D

Both the sharpness of boundaries and robustness are also evident in the image experiments. These properties are used to advantage in segmenting the SAR image of Fig. 1 in which only two textures are present (forest and grass). The scale space is shown in Fig. 12 (with the intensity values of each image scaled so as to take up the whole grayscale range), as well as the resulting boundary superimposed onto the original log-magnitude image. SAR imagery, such as the example shown here, are subject to the phenomenon known as speckle, which is present in any coherent imaging system and which leads to the large amplitude variations and noise evident in the original image. Consequently, the accurate segmentation of such imagery can be quite challenging and in particular cannot be accomplished using standard edge detection algorithms. For example, the scale-space of the region-merging algorithm of [9], as implemented in [6] and discussed above, is depicted in Fig. 13. If evolved until two regions remain, it will find the boundary around a burst of noise. In contrast, the two-region segmentation displayed in Fig. 12(d) is very accurate.

Finally we note, that, as mentioned in Experiment 1, the SIDE evolutions require far less computation time than Perona–Malik-type evolutions. Since in 2-D a SIDE evolution is not a limiting form of a Perona–Malik evolution, the comparison is not quite as simple. However, in experiments that we have performed in which we have devised Perona–Malik evolutions that produce results as qualitatively similar to those in Fig. 12 as possible, we have found that the resulting computational effort is roughly 130 times slower for this (201×201) image than our SIDE evolution.

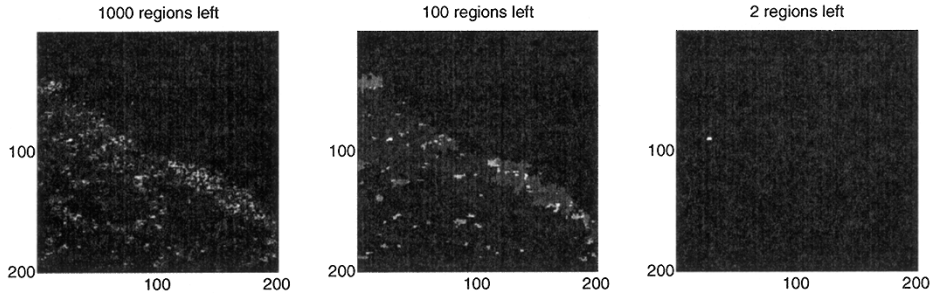


Fig. 13. Segmentations of the SAR image via the region-merging method [9].

VII. Conclusion

In this paper we have presented a new approach to edge enhancement and segmentation, and demonstrated its successful application to signals and images with very high levels of noise, as well as to blurry signals. Our approach is based on a new class of evolution equations for the processing of imagery and signals which we have termed stabilized inverse diffusion equations or SIDE's. These evolutions, which have discontinuous right-hand sides, have conceptual and mathematical links to other evolution-based methods in signal and image processing, but they also have their own unique qualitative characteristics and properties that, together with the promising results presented here, suggest the merit of several further lines of investigation—such as proving stability in 2-D; developing methods for choosing the force function F best suited to various applications; investigating improved ways of extracting information from the sequence of hitting times and corresponding images; applying SIDE's to image restoration (e.g., by filtering the initial condition within each region found by the SIDE); and understanding the properties of the PDE (16) and its interpretations.

Appendix
Proof of Lemma on Sliding

To simplify notation, we replace n_i with i in (11) and rewrite the system in terms of $v_i = u_{i+1} - u_i$:

$$\dot{v}_i = \frac{1}{m_{i+1}}\left(F(v_{i+1}) - F(v_i)\right) - \frac{1}{m_i}\left(F(v_i) - F(v_{i-1})\right) \qquad i = 1, \cdots, p-1. \tag{19}$$

We need to prove that if $(i_1, \cdots, i_{p-1})$ is any permutation of $(1, \cdots, p-1)$, then, as $\boldsymbol{v}$ approaches $S = \cap_{k=1}^{m} S_{i_k} \backslash (\cup_{k=m+1}^{p-1} S_{i_k})$, $\lim(\dot{v}_{i_q} \operatorname{sign}(v_{i_q})) \leq 0$ for all integers q between 1 and m, and for at least one such q the inequality is strict (i.e., the trajectories enter S transversally). Note that for every point $s \in S$ and every quadrant Q, we only need to find one sequence of $\boldsymbol{v}$'s approaching s from Q and satisfying these inequalities. This is because the solutions vary continuously inside each quadrant.

Fix $v_{i_{m+1}}, \cdots, v_{i_{p-1}}$ at nonzero values, let

$$\varepsilon = \tfrac{1}{2} \min_{m+1 \leq j \leq p-1} |v_{i_j}|,$$

let initially $\delta = \varepsilon$, set $|v_{i_1}| = \cdots = |v_{i_m}| = \delta$, and drive $\boldsymbol{v}$ toward S by letting δ go to zero. Take an arbitrary index q between 1 and m. By our construction, v_{i_q} is approaching zero, and either $v_{i_q} = \delta > 0$ or $v_{i_q} = -\delta < 0$. If $v_{i_q} = \delta$, then, by construction, $v_{i_q} \leq |v_{i_{q\pm1}}|$, implying $F(v_{i_q}) \geq F(v_{i_{q\pm1}})$, which makes the RHS of (19) for $i = i_q$ nonpositive: $\lim_{\delta \to 0} \dot{v}_{i_q} \leq 0$. If $m < p-1$, then there is a j between 1 and m such that at least one of the two neighbors of v_{i_j} is in the set $\{v_{i_{m+1}}, \cdots, v_{i_{p-1}}\}$, and whose absolute value is therefore staying above ε: $|v_{i_{j+1}}| > \varepsilon$ or $|v_{i_{j-1}}| > \varepsilon$. Without loss of generality, suppose it is the left neighbor: $|v_{i_{j-1}}| > \varepsilon$. If $m = p-1$, define $j = 1$. If our arbitrary q happens to be equal to this j, then

$$F(v_{i_q}) - F(v_{i_{q-1}}) = F(v_{i_j}) - F(v_{i_{j-1}}) > F(v_{i_j}) - F(\varepsilon),$$

and hence (19) for $i = i_j$ has a strictly negative limit: $\lim_{\delta \to 0} \dot{v}_{i_j} < 0$. Similar reasoning for the case $v_{i_q} = -\delta$ leads to $\lim_{\delta \to 0} \dot{v}_{i_q} \geq 0$, and, if it happens that $q = j$, then $\lim_{\delta \to 0} \dot{v}_{i_q} > 0$. ■

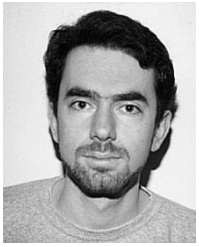
Acknowledgment

The authors would like to thank C. Bouman, O. Faugeras, S. Mallat, J.-M. Morel, and V. Utkin for stimulating discussions and many suggestions that greatly improved this paper.

References

[1] L. Alvarez, P. L. Lions, and J.-M. Morel, "Image selective smoothing and edge detection by nonlinear diffusion, II," *SIAM J. Numer. Anal.*, vol. 29, no. 3, 1992.
[2] C. Bouman and K. Sauer, "An edge-preserving method for image reconstruction from integral projections," in *Proc. Conf. Info. Sci. and Syst.*, Baltimore, MD, Mar. 1991, pp. 383–387.
[3] C. Brice and C. Fennema, "Scene analysis using regions," *Artif. Intell.*, vol. 1, 1970.
[4] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *Proc. ICCV*, Cambridge, MA, 1995, pp. 694–699.
[5] A. F. Filippov, *Differential Equations with Discontinuous Righthand Sides*. Norwell, MA: Kluwer, 1988.
[6] J. Froment and S. Parrino, MegaWave2 software package, CEREMADE, URA CNRS 749 de l'Université Paris-IX Dauphine, 1995
[7] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 3, 1992.
[8] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comp. Vis.*, vol. 1, pp. 321–331, 1988.
[9] G. Koepfler, C. Lopez, and J.-M. Morel, "A multiscale algorithm for image segmentation by variational method," *SIAM J. Numer. Anal.*, vol. 31, no. 1, 1994.
[10] J.-M. Morel and S. Solimini, *Variational Methods in Image Segmentation*. Boston, MA: Birkhauser, 1995.
[11] D. Mumford and J. Shah, "Boundary detection by minimizing functionals, I," in *Proc. CVPR*, San Francisco, CA, June 1985, pp. 22–26.
[12] S. Osher and L. I. Rudin, "Feature-oriented image enhancement using shock filters," *SIAM J. Numer. Anal.*, vol. 27, no. 4, 1990.
[13] T. Pavlidis, "Segmentation of pictures and maps through functional approximation," *Comput. Graph. Image Processing*, vol. 1, 1972.

[14] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, 1990.
[15] P. Perona, T. Shiota, and J. Malik, "Anisotropic diffusion," in *Geometry-Driven Diffusion in Computer Vision*, B. M. ter Haar Romeny, Ed. Norwell, MA: Kluwer, 1994.
[16] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, 1992.
[17] G. Sapiro, "From active contours to anisotropic diffusion: Connections between basic PDE's in image processing," in *Proc. ICIP*, Lausanne, Switzerland, 1996.
[18] B. M. ter Haar Romeny, Ed., *Geometry-Driven Diffusion in Computer Vision*. Norwell, MA: Kluwer, 1994.
[19] V. I. Utkin, *Sliding Modes in Control and Optimization*. Berlin, Germany: Springer-Verlag, 1992.
[20] J. Weickert, "Nonlinear diffusion scale-spaces: From the continuous to the discrete setting," in *Proc. ICAOS: Images, Wavelets, PDE's*, Paris, France, 1996, pp. 111–118.
[21] A. Witkin, "Scale-space filtering," in *Int. Joint Conf. Artificial Intelligence*, Karlsruhe, Germany, 1983, pp. 1019–1022.
[22] S. C. Zhu and D. Mumford, "Prior learning and Gibbs reaction-diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 11, 1997.



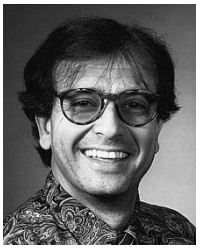
Ilya Pollak received the B.S. and M.Eng. degrees in 1995 and the Ph.D. degree in 1999, all from the Massachusetts Institute of Technology, Cambridge.

He held a National Science Foundation Graduate Fellowship from 1994 to 1997. In the Summer of 1996, he held a visiting position at the Institut National de Recherche en Informatique et en Automatique, Sophia Antipolis, France. He is currently a Postdoctoral Fellow in the Division of Applied Mathematics, Brown University, Providence, RI. He will be an Assistant Professor of Electrical and Computer Engineering at Purdue University, West Lafayette, IN, starting in August 2000. His research interests are in signal and image processing.

Alan S. Willsky (S'70–M'73–SM'82–F'86) received the S.B. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, in 1969 and 1973, respectively.

He joined the MIT faculty in 1973; his present position is Professor of Electrical Engineering. From 1974 to 1981, he was Assistant Director of the MIT Laboratory for Information and Decision Systems. He is also a founder and member of the board of directors of Alphatech, Inc., and is currently a member of the U.S. Air Force Scientific Advisory Board. He has held visiting positions at Imperial College, London, U.K., L'Université de Paris-Sud, and the Institut de Recherche en Informatique et Systémes Aléatoires, Rennes, France. He is the author of the research monograph *Digital Signal Processing and Control and Estimation* and is the coauthor of the undergraduate text *Signals and Systems*. His research interests are in development and application of advanced methods of estimation and statistical signal and image processing. Methods he has developed have been successfully applied in a wide variety of applications including failure detection in high-performance aircraft, advanced surveillance and tracking systems, electrocardiogram analysis, computerized tomography, and remote sensing.

Dr. Willsky received the Donald P. Eckman Award from the American Automatic Control Council in 1975. He was Program Chairman for the 17th IEEE Conference on Decision and Control, has been an Associate Editor of several journals and guest editor for several special issues, and has served as a member of the Board of Governors and Vice President for Technical Affairs of the IEEE Control Systems Society. In 1988, has was made a Distinguished Member of the IEEE Control Systems Society. He has given plenary and keynote lectures at a number of major scientific meetings including the 20th IEEE Conference on Decision and Control, the 1991 IEEE International Conference on Systems Engineering, the 1991 SIAM Conference on Applied Linear Algebra, the 1992 Inaugural Workshop for the National Centre for Robust and Adaptive Systems, Canberra, Australia, the 1992 INRIA 25th Anniversary Symposium in Paris, the 1993 IEEE Symposium on Image and Multidimensional Signal Processing in Cannes, France, and the 1997 Wavelet Applications in Signal and Image Processing Conference. He was awarded the 1979 Alfred Noble Prize by the ASCE and the 1980 Browder J. Thompson Memorial Prize Award by the IEEE for a paper excerpted from his monograph.



Hamid Krim (M'80–SM'98–F'99) received all his degrees in electrical engineering.

As a Member of Technical Staff at AT&T Bell Labs, he has worked in the area of telephony and digital communication systems/subsystems. Following an NSF postdoctoral scholar fellowship at Foreign Centers of Excellence, LSS/University of Orsay, Paris, France, he became a Research Scientist at the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, performing and supervising research. He is presently on the faculty in the ECE Department, North Carolina State University, Raleigh. His research interests are in statistical signal and image analysis and mathematical modeling with a keen emphasis on applied problems.