IMAGE DENOISING AND SEGMENTATION VIA NONLINEAR DIFFUSION*

YUNMEI CHEN†, BABA C. VEMURI ‡, AND LI WANG §

Abstract.
Image dnoising and segmentation are fundamental problems in the field of image processing and computer vision with numerous applications. In this paper we present a novel nonlinear diffusion model augmented with reactive terms that yields quality denoising and segmentation results on a variety of images. We present a proof for the existence, uniqueness and stability of the viscosity solution of this PDE-based model. To achieve a faster implementation, we embed the the model in a scale space and the solution is achieved via a dynamic system governed by a coupled system of first order differential equations. The dynamic system finds the solution at a coarse scale and tracks it continuously to a desired fine scale. We implement this scale-space tracking using a multigrid technique and demonstrate the smoothing and segmentation results on several images.

Key words. Nonlinear Diffusion, Image Processing, Segmentation, PDEs, Scale-space

1. Introduction. Image denoising and segmentation are fundamental problems in the field of image processing and computer vision. Image denoising (or noise removal) is a technique that enhances images by reducing any degradations that may be present. The most common degradation source is the noise from the image acquisition system and can be modeled as Gaussian random noise in most cases. Another source of degradation is the so called salt-and-pepper noise that can occur due to a random bit error in a communication channel during transmission. In the context of image segmentation, the most important step is detection of region boundaries or edges. An *edge* in an image may be defined as a location in the image at which a significant change occurs in image intensity. Segmented images or edge maps contain very useful information and are used very often to convey the essential content of an image. Such image representations are useful in object recognition, low bit-rate image coding systems and various other applications [20].

The problem of image denoising and segmentation can be posed in either a deterministic or a stochastic framework. Stochastic methods are quite effective in achieving segmentation but are limited by their intense computational requirements [12, 11]. We will therefore limit ourselves to the deterministic formulations, specifically, partial differential equation (PDE) based methods that lend themselves to fast numerical implementations.

Image denoising and segmentation can be formulated using variational principles which in turn require solutions to PDEs. Recently, there has been a flurry of activity in the PDE-based segmentation schemes. In [29], Perona and Malik developed an *anisotropic diffusion* scheme for image smoothing and segmentation. The basic idea of this nonlinear smoothing scheme was to smooth the image while preserving the edges in it. This was done by using the following equation $I_t = div(c(\nabla I)\nabla I)$, where I is the image to be smoothed and I_t describes its evolution over time, and $c(\nabla I)$ is a decreasing function of ∇I. Segmentation was achieved by finding edges in this

*This work was supported by the Whitaker Foundation, the NIH under grant RO1-LM05944 and the National Science Foundation ERC under the grant EEC-94-02989 and its industrial partners.

†Department of Mathematics, University of Florida (yun@math.ufl.edu).

‡Department of Computer and Information Science and Engineering, University of Florida (vemuri@cise.ufl.edu).

§Department of Computer and Information Science and Engineering, University of Florida (liwang@cise.ufl.edu).

smoothed image. Catte et al., [6], Nitzberg and Mumford [26] and Alvarez *et al.* [2] recognized the ill-posedness of the Perona-Malik diffusion and proposed modifications to overcome the same. Since then, several nonlinear diffusion methods have been developed and a good account of these can be found in [27, 30, 33]. In [17], Kimia et al. proposed an elegant reaction-diffusion based theory which describes the shapes of objects in an entropy scale-space . This theory was later used by Tek and Kimia [31] for image segmentation applications.

Image segmentation can also be achieved by approaches based on *curve evolution*. Malladi et al. [23, 24] and Caselles et al. [4] used curve evolution for recovering shapes from 2-D and 3-D images. The curve evolution equation was implemented by embedding the initial curve as a level curve in a surface and allowing all the level curves of the surface to evolve simultaneously. This level-set method has the advantage of being able to elegantly represent topological changes during the evolution of the curve and thereby allows recovery of shapes without a priori knowledge of their topology. The evolution equation used in [4, 23, 24] was $\partial\phi/\partial t = g(\nabla I)\|\nabla\phi\|(c+\kappa)$, where ϕ is the embedding surface for the curve evolution, $g(\nabla I) = 1/(1+\|\nabla(G_\sigma * I)\|^2)$ is a stopping term applied on the curve evolution, $\kappa = div(\frac{\nabla\phi}{\|\nabla\phi\|})$ is the curvature of the level curves of ϕ, and c is a constant speed evolution term. This method was generalized more recently in Caselles et al. [5] and Kichenassamy et al. [16] who also established the link between the curve evolution based methods and the very popular elasticity-based *snakes* (active contour models) [15, 32, 25] used for segmentation in computer vision and image processing literature. In [22], Malladi and Sethian propose a unified approach to noise removal and image segmentation using the concept of min-max curvature flow. Based on the image data, a min/max switch was designed to select $min(\kappa, 0.0)$ or $max(\kappa, 0.0)$ so that the curvature based curve evolution smoothes out small oscillations, but maintains the essential properties of the shape. Results of implementation were shown on a variety of images yielding quality noise removal and image segmentation. Anisotropic diffusion filters that use a tensorial diffusivity parameter was introduced in Weickert [33]. These filters can be tailored to enhance image structures (edges, parallel lines, curves etc.) that occur in preferred directions. More recently, Kimmel et. al., [18] presented a very general flow called the Beltrami flow as a general framework for image smoothing and show that most flow-based smoothing schemes may be viewed as special cases in their framework.

In [30], Shah developed a common framework for curve evolution, image denoising and segmentation, and anisotropic diffusion. In this work, a new segmentation functional was developed which lead to a coupled system of PDEs, one of them performed nonlinear smoothing of the input image and the other smoothed an "edge strength" function. Shah [30] demonstrated that all the existing curve evolution and anisotropic diffusion schemes reported in literature can be viewed as special cases of his method.

Each of the methods we discussed above is in a variational form that minimizes an energy functional which in some cases is nonconvex. This nonconvex minimization is hard and most computationally feasible methods lead to suboptimal solutions. In [35, 34], a coarse-to-fine scale space tracking technique was proposed as a means to efficiently achieve a significant optimum for the nonconvex optimization. In this approach, the desired solution is found by first finding a solution at a significantly coarser scale and then tracking it down through finer scales (see Figure 1). It was demonstrated that this technique can find significant minima of practical interest that exist over a large range of scale [35, 34]. In our implementation, to achieve better computational performance, we use a similar approach to [35].

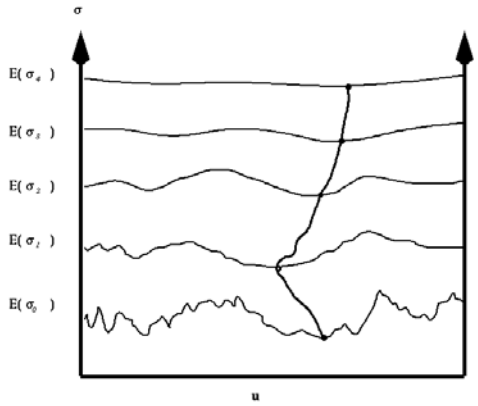


FIG. 1. *Scale space tracking: the desired solution is obtained by finding a solution at a coarser scale and then tracking it down through finer scales.*

In this paper, we present a new PDE-based image denoising and segmentation approach which is based on a nonlinear diffusion equation with additive reactive terms. We prove the existence, uniqueness and stability of the viscosity-solution for this equation and present implementation details along with several experiments demonstrating the effectiveness of the proposed denoising and segmentation scheme.

The rest of the paper is organized as follows: in the next section, we present our new nonlinear diffusion model which is used for selective smoothing of images. In section 3, we prove the existence, uniqueness and stability of the viscosity solution of the model. Section 4 contains a description of the numerical methods used to implement the model equation. In section 5, examples of results obtained using our new technique on a variety of image data are presented. We conclude the paper in section 6.

2. Nonlinear diffusion model. Numerous models of linear and nonlinear diffusion have been proposed in literature for achieving image smoothing and segmentation. A survey of various nonlinear methods is discussed in Weickert [33]. The linear models involved the standard heat equation $\partial_t u = \Delta u$ with $u(x,0) = f(x)$ where $f : \Re^2 \to \Re$ is a scalar valued image. This type of linear diffusion blurs important image features such as edges. In addition, it displaces the edges i.e., when moving from finer to coarser scales, it dislocates the edges. In general, relating structures across scales is nontrivial due to the presence of bifurcations.

Nonlinear anisotropic diffusion has been proposed by many researchers [29, 6, 2, 7]. All of these are nonlinear models and differ in the diffusivity coefficient and/or the diffusion term. Some of them are also supplemented with a reactive term. In the following we present a nonlinear diffusion equation supplemented with reactive terms for achieving edge preserving smoothing and segmentation.

Our model takes the following form:

$$\frac{\partial u}{\partial t} = g(\nabla G_\sigma * u)|\nabla u| div(\frac{\nabla u}{|\nabla u|}) + \nabla g(\nabla G_\sigma * u) \cdot \nabla u \tag{1}$$

$$- \beta|\nabla u|(u - I) \tag{2}$$

$$\frac{\partial u}{\partial n}|_{\partial R} = 0, \quad u_0 = I \tag{3}$$

where $I(x,y)$ is the intensity image to be processed, $u(x,y,t)$ is it's smoothed version, β is a weighting parameter, $G_\sigma(x,y) = \frac{1}{4\pi\sigma} exp\{-(x^2+y^2)/(2\pi\sigma^2)\}$ is a Gaussian smoothing kernel with a pre-specified σ, and $g(s) = 1/(1+\|s\|^2/K)$ is a non-increasing real valued function (for $s > 0$) which tends to zero as $s \to \infty$ with a constant K.

This equation can be posed as the minimization of the following energy functional for fixed v, where $v = \nabla G_\sigma * u$. Note that the variation is taken with respect to u for a fixed v.

$$E_v(u) = \int\int_R \{g(v)|\nabla u| + \beta(u-I)^2\} dxdy \tag{4}$$

From this functional, it is easy to see that:

1. the coefficient of the first term namely, $g(\nabla G_\sigma * u)$ serves the purpose of selecting the locations in the image for smoothing. For instance, at image locations having large values of gradient, this coefficient takes on a small value thereby reducing the smoothing performed at these locations since $g(s)$ is a non-increasing function of s.
2. The term $|\nabla u|$ regularizes the solution u and is responsible for the nonlinear term $|\nabla u| div\left(\frac{\nabla u}{|\nabla u|}\right)$ in equation 2. It diffuses u in a direction away from that of ∇u.
3. The term $(u-I)^2$ forces u to be a close approximation to the data I.

Note that our model is different from the one reported in Alvarez *et. al.*, [2] in that it yields two additional terms in the Euler-Lagrange PDE 2 – a.k.a. the gradient descent equation – for the variation principle 4. These two reactive terms are responsible for forcing, the evolution-based image smoothing to stop at the edges and the resulting solution to be close approximation to the original image data respectively. Our model is also different from the one reported in Alvarez and Esclarin [1] where in the reactive term was used to accommodate the image quantization application. Our model is in spirit similar the smoothing equation in [30] however, the diffusivity coefficient is very different in our case.

For a *fast implementation* and to obtain quality smoothing and segmentation results, the scale parameter σ in the functional 4 can be varied from large (coarse) to small (fine) values using a scale space tracking scheme. The scale space tracking can be posed as the steady state solution of a differential equation obtained by differentiating the equation representing the equilibrium condition of the functional 4 with respect to σ. The resultant differential equation describing the scale space trajectory is however very complicated. A simpler coupled system of equations for tracking the solution across scale (from coarse-to-fine σ) for this nonlinear diffusion model can be obtained using the approach described in [35]. The coupled system of scale space equations are given by

$$\begin{aligned} \frac{\partial u}{\partial t} &= -\nabla E_v(u,\sigma) \\ \frac{\partial \sigma}{\partial t} &= -c_1 e^{-c_2|\nabla E_v(u,\sigma)|} \end{aligned} \tag{5}$$

where c_1 and c_2 are pre-specified constants. For a set of initial u_0 and σ_0 which is far away from the solution, $\nabla E_v(u,\sigma)$ is large while $-c_1 e^{-c_2|\nabla E(u,\sigma)|}$ is small, thus equation (5) is solved for u at a nearly constant scale until a solution at this scale is found. Once a solution is obtained, it satisfies $\frac{\partial u}{\partial t} = 0$ and $\frac{\partial \sigma}{\partial t} = -c_1$. Equation (5) can then be used to track the solution down to finer scales. This technique is

computationally efficient and yields quality smoothing and segmentation results as is demonstrated in our experiments. Note that the proof of existence, uniqueness and stability of a viscosity solution of our model presented in the following section does not hold for the above described coupled system of scale-space tracking equations.

3. Existence, Uniqueness and Stability. Since our model (3) is highly nonlinear and degenerate, we need the notion of so-called viscosity solution (see [9]). In this section we will prove the existence, uniqueness and stability for the viscosity solution to the equation (3).

The well-posedness of the viscosity solution of the mean curvature flow $u_t - |\nabla u| div(\frac{\nabla u}{|\nabla u|}) = 0, \quad in\ R^n \times R_+$ and for the generalized mean curvature flow $u_t - |\nabla u|(div(\frac{\nabla u}{|\nabla u|}) + \gamma) = 0, \quad in\ R^n \times R_+,\ \gamma \in R$, were studied by Evans and Spruck [10] and by Chen-Giga-Goto [8], respectively. In [2], the existence, uniqueness and stability for the viscosity solution of the following highly nonlinear diffusion equation: $u_t - g(\nabla G_\sigma * u)|\nabla u|(div(\frac{\nabla u}{|\nabla u|}) = 0, \quad in\ R^n \times R_+$, where $g(p) = 1/1 + |p|^2$ was established. Viscosity analysis for the level set form of the active contour model $u_t - g(\nabla G_\sigma * I)|\nabla u|(div(\frac{\nabla u}{|\nabla u|}) + \gamma) + \nabla g \cdot \nabla u = 0, \quad in\ R^2 \times R_+,\ \gamma \in R$, where $I(x, y)$ is the initial surface embedding the initial contour, was briefely disscussed in [5] and [16].

Our model has a similar structure, but more nonlinear terms than in the models mentioned above. The proof of well posedness for our model is inspired by the idea of [2] and the definition of the viscosity solutions for our model follows the notion established in [9]. We first prove the existence of a solution. Our solution is obtained by applying limits to the approximate solutions of the penalized uniformly parabolic equations. We then prove the uniqueness and the stability of our solution. The proof follows the ideas in [2]. However, since our model has more nonlinear factors or terms than the models mentioned above, more careful and delicate estimates are required, especially in getting the uniform L^∞-norm estimate for the gradient of the approximate solutions and in establishing the estimate $\sup_{\Omega\times[0,T]} |u - v| \le C \sup_{\Omega\times\{t=0\}} |u - v|$, where u and v are two viscosity solutions of (3).

Our model equation (2) is in two dimensions, mathematically we can study this problem for n-dimensional cases. σ, β and K are constants in (2), and they do not effect the proof of well-posedness. To simplify the presentation, we shall consider $\sigma = \beta = K = 1$ and work with periodic boundary conditions similar to the presentation in [2]. Then, by periodic extension we consider the following Cauchy problem

$$
\begin{array}{ll}
\frac{\partial u}{\partial t} = g(\nabla G * u) a_{ij}(\nabla u) u_{x_i x_j} & \\
\quad + \frac{\partial g}{\partial l}(\nabla G * u)[(\nabla G_{x_l} * u) \cdot \nabla u] - |\nabla u|(u - I), & x \in R^n,\ t \in R_+ \\
u(x, 0) = I(x), & x \in R^n
\end{array} \tag{6}
$$

where $G = \frac{1}{4\pi} exp\{-(x^2+y^2)/4\}$, $g(p) = \frac{1}{1+|p|^2}$, $a_{ij}(p) = \delta_{ij} - \frac{p_i p_j}{|p|^2}$, and the summation convention is used.

First, let us recall the definition of viscosity subsolution of (6). A function $u \in C([0, T] \times R^n)$ for some $T > 0$ is said to be a viscosity solution of (6), if for all $\phi \in C^2(R^2 \times R)$, the following condition holds at any point $(x_0, t_0) \in R^n \times (0, T]$, at which $(u - \phi)$ attains a local maximum.

$$
\begin{aligned}
\frac{\partial \phi}{\partial t}(x_0, t_0) - g((\nabla G * u)(x_0, t_0)) a_{ij}(\nabla \phi(x_0, t_0)) \phi_{x_i x_j}(x_0, t_0) \\
+ \frac{\partial g}{\partial l}((\nabla G * u)(x_0, t_0))[(\nabla G_{x_l} * u)(x_0, t_0) \cdot \nabla \phi(x_0, t_0)]
\end{aligned} \tag{7}
$$

$$-|\nabla\phi(x_0,t_0)|(u-I)(x_0,t_0)\le 0,\quad if \nabla\phi(x_0,t_0)\neq 0$$

$$\frac{\partial\phi}{\partial t}(x_0,t_0)-g((\nabla G*u)(x_0,t_0))\limsup_{p\to 0}a_{ij}(p)\phi_{x_ix_j}(x_0,t_0)\le 0, \quad \text{if } \nabla\phi(x_0,t_0)=0 \tag{8}$$

A viscosity supersolution is similarly defined by substituting "local maximum" for "local minimum", "≤ 0" for "≥ 0", and "limsup" for "liminf" in equations (7) and (8) respectively. A viscosity solution is a continuous function which is both a subsolution and a supersolution. We now state and prove the main theorem of this paper.

THEOREM 3.1. *The Cauchy problem (6) has a unique viscosity solution* $u\in C(R^n\times[0,T])\cap L^\infty(0,T;W^{1,\infty}(R^n))$ *for any* $T\in[0,\infty)$, *and* $\inf_{R^n} I\le u(x,t)\le \sup_{R^n} I$, *provided that* I *is Lipschitz continuous and continuous on* R^n.

Moreover, if $v\in C(R^n\times R_+)$ *is a viscosity solution of (6) with* I *replaced by a Lipschitz continuous function* I_1, *then for all* $T\in[0,+\infty)$, *there exists a constant* $C>0$, *depending only on* I, I_1, *and* T, *such that*

$$\sup_{0\le t\le T}\|u(x,t)-v(x,t)\|_{L^\infty(R^n)}\le C\|I-I_1\|_{L^\infty(R^n)}$$

Proof. In this proof of the theorem, we use the technique of viscosity solutions theory discussed in [1, 2, 8, 9, 10]. The proof is in several stages.

Step 1. We first show that if u is a viscosity solution of (6) on $R^n\times R_+$, then,

$$\inf_{R^n} I\le u\le \sup_{R^n} I,\qquad \text{on } R^n\times[0,\infty) \tag{9}$$

Let $\phi=\sup_{R^n} I+\delta t$ (where $\delta>0$) in (7) and assume that $u-\phi$ attains a local maximum at (x_0,t_0) with $t_0>0$, then $\nabla\phi(x_0,t_0)=0$ and from equation (7), $\frac{\partial\phi}{\partial t}(x_0,t_0)\le 0$. This contradicts $\frac{\partial\phi}{\partial t}\equiv\delta>0$ on $R^n\times[0,\infty)$. Therefore, $u-\phi$ must attain its maximum at $t_0=0$. So,

$$u-\phi\le\sup(I-\sup_{R^n} I)$$

$$u\le\sup_{R^n} I+\delta t.$$

Similarly we have (from the definition of supersolution)

$$u\ge\inf_{R^n} I-\delta t.$$

Leting $\delta\to 0$ proves (9).

Step 2. Next we prove the gradient estimate for the approximate solution. Consider the following Cauchy problem:

$$\begin{aligned}
&\tfrac{\partial u^\varepsilon}{\partial t}=g^\varepsilon(\nabla G*u^\varepsilon)a^\varepsilon_{ij}(\nabla u^\varepsilon)u^\varepsilon_{x_ix_j}\\
&\quad+\tfrac{\partial g^\varepsilon}{\partial l}(\nabla G*u^\varepsilon)[(\nabla G_{x_l}*u^\varepsilon)\cdot\nabla u^\varepsilon]\\
&\quad-b^\varepsilon(\nabla u^\varepsilon)(u^\varepsilon-I^\varepsilon), && x\in R^n,\ t\in R_+\\
&u^\varepsilon(x,0)=I^\varepsilon(x), && x\in R^n
\end{aligned} \tag{10}$$

where

$$
\begin{aligned}
&0 < \varepsilon < 1\\
&g^\varepsilon(s) = g(s) + \varepsilon\\
&a_{ij}^\varepsilon(p) = (\varepsilon+1)\delta_{ij} - \tfrac{P_iP_j}{|P|^2+\varepsilon^2}\\
&b^\varepsilon(p) = \sqrt{|p|^2+\varepsilon}\\
&I^\varepsilon \in C^\infty(R^n) \text{ (periodic) such that } I^\varepsilon \to I \text{ uniformly and}\\
&\|\nabla I^\varepsilon\|_{L^\infty(R^n)} \le \|\nabla I\|_{L^\infty(R^n)}, \quad \|I^\varepsilon\|_{L^\infty(R^n)} \le \|I\|_{L^\infty(R^n)}.
\end{aligned}
$$

From the theory of quasilinear uniformly parabolic equations ([19] §6, thm. 4.4), the problem (10) admits a smooth solution $u^\varepsilon \in C^\infty(R^n \times R_+)$. Since any smooth solution is a viscosity solution, by an argument similar to that in step 1, we know that,

$$
|u^\varepsilon| \le M \text{ for } (x,t) \in R^n \times [0,\infty) \tag{11}
$$

where $M > 0$ is a constant depending only on I. Now we shall show a uniform estimate for $|\nabla u^\varepsilon|_{L^\infty(R^n)}$.

Differentiating (10) with respect to x_k yields,

$$
\begin{aligned}
\frac{\partial u^\varepsilon_{x_k}}{\partial t} &= g^\varepsilon(\nabla G * u^\varepsilon)a_{ij}^\varepsilon(\nabla u^\varepsilon)u^\varepsilon_{x_ix_jx_k} + g^\varepsilon(\nabla G * u^\varepsilon)\frac{\partial a_{ij}^\varepsilon}{\partial l}(\nabla u^\varepsilon)u^\varepsilon_{x_lx_k}u^\varepsilon_{x_ix_j}\\
&+ \frac{\partial g^\varepsilon}{\partial l}(\nabla G * u^\varepsilon)(G_{x_lx_k} * u^\varepsilon)a_{ij}^\varepsilon(\nabla u^\varepsilon)u^\varepsilon_{x_ix_j}\\
&+ \frac{\partial^2 g^\varepsilon}{\partial l \partial m}(\nabla G * u^\varepsilon)(G_{x_mx_k} * u^\varepsilon)(\nabla G_{x_l} * u^\varepsilon \cdot \nabla u^\varepsilon)\\
&+ \frac{\partial g^\varepsilon}{\partial l}(\nabla G * u^\varepsilon)[(\nabla G_{x_lx_k} * u^\varepsilon)\cdot \nabla u^\varepsilon] + \frac{\partial g^\varepsilon}{\partial l}(\nabla G * u^\varepsilon)[(\nabla G_{x_l} * u^\varepsilon)\cdot \nabla u^\varepsilon_{x_k}]\\
&- \frac{\partial b^\varepsilon(\nabla u^\varepsilon)}{\partial m}u^\varepsilon_{x_mx_k}(u^\varepsilon - I^\varepsilon)\\
&- b^\varepsilon(\nabla u^\varepsilon)(u^\varepsilon_{x_k} - I^\varepsilon_{x_k}), \qquad x \in R^n,\ t > 0.
\end{aligned} \tag{12}
$$

Multiplying by $2u^\varepsilon_{x_k}$ on both sides of (12) and taking summation w.r.t. k, we get

$$
\begin{aligned}
\frac{\partial |\nabla u^\varepsilon|^2}{\partial t} &- g^\varepsilon(\nabla G * u^\varepsilon)a_{ij}^\varepsilon(\nabla u^\varepsilon)\frac{\partial^2 |\nabla u^\varepsilon|^2}{\partial x_i \partial x_j}\\
&- g^\varepsilon(\nabla G * u^\varepsilon)\frac{\partial a_{ij}^\varepsilon}{\partial l}(\nabla u^\varepsilon)u^\varepsilon_{x_ix_j}\frac{\partial |\nabla u^\varepsilon|^2}{\partial x_l}\\
&- \frac{\partial g^\varepsilon}{\partial l}(\nabla G * u^\varepsilon)(\nabla G_{x_l} * u^\varepsilon)\cdot(\nabla |\nabla u^\varepsilon|^2) + \frac{\partial b^\varepsilon(\nabla u^\varepsilon)}{\partial m}(u^\varepsilon - I^\varepsilon)\frac{\partial |\nabla u^\varepsilon|^2}{\partial x_m}\\
&= 2\frac{\partial g^\varepsilon}{\partial l}(\nabla G * u^\varepsilon)(G_{x_lx_k} * u^\varepsilon)a_{ij}^\varepsilon(\nabla u^\varepsilon)u^\varepsilon_{x_ix_j}u^\varepsilon_{x_k}\\
&+ 2\frac{\partial^2 g^\varepsilon}{\partial l \partial m}(\nabla G * u^\varepsilon)(G_{x_mx_k} * u^\varepsilon)(\nabla G_{x_l} * u^\varepsilon \cdot \nabla u^\varepsilon)u^\varepsilon_{x_k}\\
&+ 2\frac{\partial g^\varepsilon}{\partial l}(\nabla G * u^\varepsilon)[(\nabla G_{x_lx_k} * u^\varepsilon)\cdot \nabla u^\varepsilon]u^\varepsilon_{x_k}\\
&- 2g^\varepsilon(\nabla G * u^\varepsilon)a_{ij}^\varepsilon(\nabla u^\varepsilon)u^\varepsilon_{x_kx_i}u^\varepsilon_{x_kx_j}u^\varepsilon_{x_k}\\
&- 2b^\varepsilon(\nabla u^\varepsilon)(u^\varepsilon_{x_k} - I^\varepsilon_{x_k})u^\varepsilon_{x_k}
\end{aligned} \tag{13}
$$

¿From the definitions of a_{ij}^{ε}, b^{ε}, g^{ε}, and G, we have

$$|a_{ij}^{\varepsilon}(\nabla u^{\varepsilon})u_{x_i x_j}^{\varepsilon}|^2 \le 2a_{ij}^{\varepsilon}(\nabla u^{\varepsilon})u_{x_k x_i}^{\varepsilon}u_{x_k x_j}^{\varepsilon}$$
$$|\tfrac{\partial g}{\partial l}(\nabla G * u^{\varepsilon})|^2 \le 2g^{\varepsilon}(\nabla G * u^{\varepsilon})$$

and for any multi-index α with $|\alpha| \le 2$,

$$\sup_{R^n \times R_+} |\nabla^{\alpha} G * u^{\varepsilon}| \le C$$
$$\sup_{R^n \times R_+} |\nabla^{\alpha} g^{\varepsilon}(\nabla G * u^{\varepsilon})| \le C,$$

where $C > 0$ is a constant depending only on M in (11).

Inserting these estimates into (13) and using Cauchy's inequality we have

$$\text{RHS of (13)} \le C(|\nabla u^{\varepsilon}|^2 + 1) \qquad \text{in } R^n \times R_+ \tag{14}$$

where $C > 0$ is a constant depending only on M in (11), hence C depends only on I. Applying the maximum principle [3] to (14) yields for all $t \in [0, T]$ (for any $T < \infty$),

$$\begin{aligned} \|\nabla u^{\varepsilon}(\cdot, t)\|_{L^{\infty}(R^n)} &\le e^{ct}\|\nabla I^{\varepsilon}\|_{L^{\infty}(R^n)} \\ &\le e^{ct}\|\nabla I\|_{L^{\infty}(R^n)} \le C_T, \end{aligned} \tag{15}$$

where $C_T > 0$ depends only on T, and I. This implies that

$$|u^{\varepsilon}(x,t) - u^{\varepsilon}(y,t)| \le C_T|x - y|, \text{ for } \forall x, y \in R^n \text{ and } \forall t \in [0,T].$$

By the same argument used in [9], we have

$$|u^{\varepsilon}(x,s) - u^{\varepsilon}(x,t)| \le C_T|t - s|^{\frac{1}{2}}, \text{ for } \forall x \in R^n \text{ and } \forall s, t \in [0,T].$$

Then, by the Ascoli-Arzela theorem, there exists a subsequence u^{ε_k} of u^{ε}, and a function $u \in C(R^n \times [0,T]) \cap L^{\infty}(0,T; W^{1,\infty}(R^n))$ such that as $\varepsilon_k \to 0$,

$$u^{\varepsilon_k} \to u \text{ locally uniformly in } R^n \times R_+. \tag{16}$$

Step 3. Existence of viscosity solution (16).

We assert now that u obtained in (16) is a viscosity solution of (6) in the sense of (7) and (8).

Let $\phi \in C^2(R^n \times R)$ and assume $u - \phi$ has a strict local maximum at a point $(x_0, t_0) \in R^n \times R_+$. As $u^{\varepsilon_k} \to u$ uniformly near (x_0, t_0), $u^{\varepsilon_k} - \phi$ has a local maximum at a point (x_k, t_k) with

$$(x_k, t_k) \to (x_0, t_0) \text{ as } k \to \infty \tag{17}$$

and at (x_k, t_k)

$$\nabla u^{\varepsilon_k} = \nabla \phi, \quad u_t^{\varepsilon_k} = \phi_t, \quad a_{ij}^{\varepsilon_k}(\nabla u^{\varepsilon_k})u_{x_i x_j}^{\varepsilon_k} \le a_{ij}^{\varepsilon_k}(\nabla \phi)\phi_{x_i x_j} \tag{18}$$

Therefore, (10) implies that at (x_k, t_k),

$$\begin{aligned} &\frac{\partial \phi}{\partial t} - g^{\varepsilon_k}(\nabla G * u^{\varepsilon_k})a_{ij}^{\varepsilon_k}(\nabla \phi)\phi_{x_i x_j} - \frac{\partial g^{\varepsilon_k}}{\partial l}(\nabla G * u^{\varepsilon_k})[\nabla G_{x_l} * u^{\varepsilon_k} \cdot \nabla \phi] \\ &\quad + b^{\varepsilon_k}(\nabla \phi)(u^{\varepsilon_k} - I^{\varepsilon_k}) \le 0 \end{aligned} \tag{19}$$

If $\nabla\phi(x_0,t_0)\neq 0$, from to (17), for sufficiently large k, $\nabla\phi(x_k,t_k)\neq 0$. One may apply limits in (19) to obtain (recalling the definitions of a_{ij}^{ε}, g^{ε}, b^{ε}, and (16) - (17))

$$\frac{\partial\phi}{\partial t}-g(\nabla G*u)a_{ij}(\nabla\phi)\phi_{x_ix_j}-\frac{\partial g}{\partial l}(\nabla G*u)[(\nabla G_{x_l}*u)\cdot\nabla\phi]+b(\nabla\phi)(u-I)\leq 0 \quad \text{at} \quad (x_0,t_0)$$

which is the same as (7).

If $\nabla\phi(x_0,t_0)=0$, let

$$h^k=\frac{\nabla\phi(x_k,t_k)}{\sqrt{|\nabla\phi(x_k,t_k)|^2+\varepsilon^2}}$$

then (19) reduces to

$$\begin{aligned}\frac{\partial\phi}{\partial t}&-g^{\varepsilon_k}(\nabla G*u^{\varepsilon_k})((\varepsilon_k+1)\delta_{ij}-h_i^kh_j^k)\phi_{x_ix_j}\\&-\frac{\partial g^{\varepsilon_k}}{\partial l}(\nabla G*u^{\varepsilon_k})[(\nabla G_{x_l}*u^{\varepsilon_k})\cdot\nabla\phi]\\&+b^{\varepsilon_k}(\nabla\phi)(u^{\varepsilon_k}-I^{\varepsilon_k})\leq 0 \qquad \text{at} \quad (x_k,t_k)\end{aligned} \tag{21}$$

Since $\nabla\phi(x_k,t_k)\to 0$, $\varepsilon_k\to 0$ as $k\to\infty$, hence, $b^{\varepsilon_k}(\nabla\phi(x_k,t_k))\to 0$. Moreover, becuase $|h^k|\leq 1$, there is a sub-sequence of h^k, also denoted by h^k, such that as $k\to\infty$, $h^k\to h$ in $R^n\times R$ with $|h|\leq 1$. Applying limits to (21), we get

$$\frac{\partial\phi}{\partial t}-g(\nabla G*u)(\delta_{ij}-h_ih_j)\phi_{x_ix_j}\leq 0 \qquad \text{at} \quad (x_0,t_0) \tag{22}$$

This is the same as (8). If $u-\phi$ has a local maximum, but not necessarily a strict local maximum at (x_0,t_0), we just need to repeat the argument above with $\phi(x,t)$ replaced by $\tilde{\phi}(x,t)=\phi(x,t)+|x-x_0|^4+(t-t_0)^4$. Therefore, u is a subsolution of (6). Similarly, we can show that u is a supersolution. Hence, u is a viscosity solution of (6).

Step 4. Uniqueness.

We will now prove the uniqueness and stability for the viscosity solution of (6). This proof is based on Theorem 8.3 in [9]. Let u be a viscosity solution of (6) with Lipschitz continuous initial data I and v be a viscosity solution of (6) with I replaced by a Lipschitz continuous function I_1. Let

$$\omega(x,y,t)=u(x,t)-v(y,t)-(4\delta)^{-1}|x-y|^4-\lambda t,\quad t\in[0,T],\ x,y\in R^n$$

where $\delta>0$ and $\lambda>0$ are constants to be determined later.

Claim: $\omega(x,y,t)$ attains maximum at $t=0$ for an arbitrary positive constant λ.

Indeed, if $\omega(x,y,t)$ attains its maximum at some point (x_0,y_0,t_0) with $t_0>0$, by theorem 8.3 in [9], for each $\mu>0$, there exist X and Y, $(n\times n)-$symmetric matrices, and $\alpha,\beta\in R$, such that

$$\alpha-\beta=\lambda, \tag{23}$$

$$\begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix}\leq A+\mu A^2 \tag{24}$$

and

$$\begin{aligned}(25)\qquad &\alpha - g((\nabla G * u)(x_0,t_0))a_{ij}(\delta^{-1}|x_0-y_0|^2(x_0-y_0))X_{ij}\\ &- \frac{\partial g}{\partial l}((\nabla G * u)(x_0,t_0))[(\nabla G_{x_l} * u)(x_0,t_0)\cdot\delta^{-1}|x_0-y_0|^2(x_0-y_0)]\\ &+\delta^{-1}|x_0-y_0|^3(u(x_0,t_0)-I(x_0))\le 0\end{aligned}$$

$$\begin{aligned}(26)\qquad &\beta - g((\nabla G * v)(y_0,t_0))a_{ij}(\delta^{-1}|x_0-y_0|^2(x_0-y_0))Y_{ij}\\ &- \frac{\partial g}{\partial l}((\nabla G * v)(y_0,t_0))[(\nabla G_{x_l} * v)(y_0,t_0)\cdot\delta^{-1}|x_0-y_0|^2(x_0-y_0)]\\ &+\delta^{-1}|x_0-y_0|^3(v(y_0,t_0)-I_1(y_0))\ge 0\end{aligned}$$

where

$$\begin{aligned}(27)\qquad &A=(A_{ij})_{n\times n} \text{ and}\\ &A_{ij}=\delta^{-1}|x_0-y_0|^2\delta_{ij}+2\delta^{-1}(x_0-y_0)_i(x_0-y_0)_j,\end{aligned}$$

here $(x_0-y_0)_i$ stands for the $i-th$ component of x_0-y_0.

We observe that $x_0\neq y_0$. Indeed, if $x_0=y_0$, then from (27), $A=0$, hence $X\le 0$ and $Y\ge 0$ from (24). Thus, (25) - (26) leads to $\alpha\le 0$ and $\beta\ge 0$, which contradicts $\alpha-\beta=\lambda>0$. We now choose

$$(28)\qquad \mu=\delta|x_0-y_0|^{-2}$$

¿From (27) and (24) after some algebra we have,

$$(29)\qquad \begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix}\le 2\delta^{-1}\begin{pmatrix} B & -B\\ -B & B\end{pmatrix},$$

where

$$B_{ij}=|x_0-y_0|^2\delta_{ij}+5(x_0-y_0)_i(x_0-y_0)_j,\qquad 1\le i,j\le n.$$

Let

$$\begin{aligned}&U=(\nabla G*u)(x_0,t_0),\qquad V=(\nabla G*v)(y_0,t_0),\\ &D=(a_{ij}(\delta^{-1}|x_0-y_0|^2(x_0-y_0))_{1\le i.j\le n},\end{aligned}$$

and

$$Q=\begin{pmatrix} g(U)D & \sqrt{g(U)g(V)}D\\ \sqrt{g(U)g(V)}D & g(V)D\end{pmatrix}$$

Noting that Q is a nonnegative symmetric matrix, from (29) we have

$$Q\begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix}\le 2\delta^{-1}Q\begin{pmatrix} B & -B\\ -B & B\end{pmatrix}.$$

Taking the trace we get

$$\begin{aligned}(30)\qquad g(U)D_{ij}X_{ij}-g(V)D_{ij}Y_{ij}&\le 2\delta^{-1}(\sqrt{g(U)}-\sqrt{g(V)})^2\mathrm{trace}(DB)\\ &\le 4\delta^{-1}(\sqrt{g(U)}-\sqrt{g(V)})^2|x_0-y_0|^2.\end{aligned}$$

Then, from (23) and (25) - (26),

$$\lambda = \alpha - \beta \le I + II + III, \tag{31}$$

where

$$\begin{aligned} I &= g(U)D_{ij}X_{ij} - g(V)D_{ij}Y_{ij} \\ &\le 4\delta^{-1}(\sqrt{g(U)} - \sqrt{g(V)})^2|x_0 - y_0|^2 \end{aligned} \tag{32}$$

$$\begin{aligned} II = [&\frac{\partial g}{\partial l}(U)(\nabla G_{x_l} * u)(x_0, t_0) \\ &- \frac{\partial g}{\partial l}(V)(\nabla G_{x_l} * v)(y_0, t_0)] \cdot \delta^{-1}|x_0 - y_0|^2(x_0 - y_0) \end{aligned} \tag{33}$$

$$III = \delta^{-1}|x_0 - y_0|^3[(u(x_0, t_0) - v(y_0, t_0)) + (I(x_0) - I_1(y_0))]. \tag{34}$$

We now estimate (32)–(34). First,

$$\begin{aligned} |U - V| &\le |(\nabla G * u)(x_0, t_0) - (\nabla G * v)(x_0, t_0)| \\ &+ |(\nabla G * v)(x_0, t_0) - (\nabla G * v)(y_0, t_0)| \\ &\le C(\sup_{R^n \times [0,T]} |u - v| + |x_0 - y_0|) \end{aligned} \tag{35}$$

where constant $C > 0$ depends only on I and I_1 (as evident from (9)) and the Lipschitz constants for u and v. Then, by the mean value theorem and the fact that u and v are Lipschitz continuous, we have

$$\sqrt{g(U)} - \sqrt{g(V)} \le C|U - V| \le C(\sup_{R^n \times [0,T]} |u - v| + |x_0 - y_0|). \tag{36}$$

Similarly,

$$\begin{aligned} &|\frac{\partial g}{\partial l}(U)(\nabla G_{x_l} * u)(x_0, t_0) - \frac{\partial g}{\partial l}(V)(\nabla G_{x_l} * v)(y_0, t_0)| \\ &\le |\frac{\partial g}{\partial l}(U)(\nabla G_{x_l} * u)(x_0, t_0) - \frac{\partial g}{\partial l}(V)(\nabla G_{x_l} * v)(x_0, t_0)| \\ &+ |\frac{\partial g}{\partial l}(V)(\nabla G_{x_l} * v)(x_0, t_0) - \frac{\partial g}{\partial l}(V)(\nabla G_{x_l} * v)(y_0, t_0)| \\ &\le C(\sup_{R^n \times [0,T]} |u - v| + |x_0 - y_0|), \end{aligned} \tag{37}$$

here we used (35) and the facts that $|\frac{\partial g}{\partial l}| \le 2$ and v are Lipschitz continuous.

It is easy to see that

$$|u(x_0, t_0) - v(y_0, t_0)| + |I(x_0) - I_1(y_0)| \tag{38}$$

$$\le |u(x_0, t_0) - v(x_0, t_0)| + |v(x_0, t_0) - v(y_0, t_0)| \tag{39}$$

$$\begin{aligned} &+ |I(x_0) - I_1(x_0)| + |I_1(x_0) - I_1(y_0)| \\ &\le C(\sup_{R^n \times [0,T]} |u - v| + |x_0 - y_0|). \end{aligned} \tag{40}$$

The constants $C > 0$ in (35) - (37) depend only on I, I_1 and the Lipschitz constants for u and v.

Inserting (35)-(38) into (31) - (34) yields

$$\lambda \leq C\delta^{-1}[(\sup_{R^n\times[0,T]}|u-v|)^2|x_0-y_0|^2+|x_0-y_0|^4 + (\sup_{R^n\times[0,T]}|u-v|)|x_0-y_0|^3]. \tag{41}$$

On the other hand, since (x_0, y_0, t_0) is the maximum point of $\omega(x, y, t)$,

$$u(x_0,t_0)-v(x_0,t_0)-(4\delta)^{-1}|x_0-y_0|^4-\lambda t_0 \geq u(y_0,t_0)-v(y_0,t_0)-\lambda t_0.$$

This leads to

$$|x_0-y_0| \leq (4\delta L)^{\frac{1}{3}} \tag{42}$$

where L is a Lipschitz constant for u in $R^n \times [0, T]$. Combining this with (41) yields

$$\begin{aligned}\lambda &\leq C\delta^{-1}\{(\sup_{R^n\times[0,T]}|u-v|)^2(4\delta L)^{\frac{2}{3}}+(\sup_{R^n\times[0,T]}|u-v|)4\delta L+(4\delta L)^{\frac{4}{3}}\} \\ &\leq C_0\{\delta^{-\frac{1}{3}}(\sup_{R^n\times[0,T]}|u-v|)^2+\sup_{R^n\times[0,T]}|u-v|+\delta^{\frac{1}{3}}\}\end{aligned} \tag{43}$$

where $C_0 > 0$ depends only on I, I_1 and the Lipschitz constants of u and v. We now set

$$\delta = L^{-4}(\sup_{R^n\times[0,T]}|u-v|)^3 \tag{44}$$

and from (43) we obtain

$$\lambda \leq C_0(L^{\frac{4}{3}}+L^{-\frac{4}{3}}+1)\sup_{R^n\times[0,T]}|u-v|. \tag{45}$$

Let

$$\lambda = C_0(L^{\frac{4}{3}}+L^{-\frac{4}{3}}+2)\sup|u-v| \tag{46}$$

This leads to a contradiction with (45). Therefore for the choice (46) of λ, $\omega(x, y, t)$ attains its maximum at $t = 0$, which is our claim. Hence,

$$\begin{aligned}&u(x,t)-v(y,t)-(4\delta)^{-1}|x-y|^4-\lambda t \\ &\leq \sup_{x,y\in R^n}(u(x,0)-v(y,0)-(4\delta)^{-1}|x-y|^4) \\ &\leq \sup_{x,y\in R^n}(I(y)-I_1(y)+I(x)-I_1(x)-(4\delta)^{-1}|x-y|^4) \\ &\leq \sup_{R^n}|I-I_1|+\sup_{|x-y|\geq 0}(I(x)-I(y)-(4\delta)^{-1}|x-y|^4) \\ &\leq \sup_{R^n}|I-I_1|+\sup_{|x-y|\geq 0}L|x-y|-(4\delta)^{-1}|x-y|^4)\end{aligned} \tag{47}$$

Noticing that $\sup_{r\geq 0} Lr-(4\delta)^{-1}r^4)$ is achieved at $r = (\delta L)^{1/3}$, and letting $x = y$ in (47), from (44) and (46) we get

$$\begin{aligned}\sup_{R^n\times[0,T]}|u-v| \leq \sup_{R^n}|I-I_1|+\frac{3}{4}\sup_{R^n\times[0,T]}|u-v| \\ +C_0(L^{\frac{4}{3}}+L^{-\frac{4}{3}}+2)T\sup_{R^n\times[0,T]}|u-v|\end{aligned} \tag{48}$$

Therefore, there exists $T_0 > 0$, sufficiently small ($T_0 < \frac{1}{8C_0(L^{\frac{4}{3}}+L^{-\frac{4}{3}}+2)}$) such that from (48) we have

$$\sup_{R^n\times[0,T_0]} |u-v| \leq 8 \sup_{R^n} |I - I_1|. \tag{49}$$

For large t, by iteration, we esily obtain

$$\sup_{R^n\times[0,T]} |u-v| \leq C(T) \sup_{R^n} |I - I_1|$$

This proves the uniqueness and stability for u. □

4. Numerical implementation. The numerical implementation of the nonlinear diffusion equation (2) is based on the *upwind finite difference scheme* developed by Osher and Sethian [28] for curve evolution via level sets. Implementing the time derivative $\frac{\partial u}{\partial t}$ and the diffusive term $g(\nabla G_\sigma * I)|\nabla u| div(\frac{\nabla u}{|\nabla u|})$ presents no difficulties and is straight forward. $\frac{\partial u}{\partial t}$ is approximated by forward differences

$$\frac{u^{n+1}[i,j] - u^n[i,j]}{\triangle t}$$

and the diffusive term is approximated using the usual central differences, with

$$|\nabla u| div(\frac{\nabla u}{|\nabla u|}) = \frac{u_y^2 u_{xx} - 2u_x u_y u_{xy} + u_x^2 u_{yy}}{u_x^2 + u_y^2}.$$

What does require special care is the implementation of the data term (in fact, an inflation term i.e., constant speed of expansion) $\beta|\nabla u|\frac{(u-I)}{|u-I|}$ and the "doublet" term $\nabla g \cdot \nabla u$ [5]. The inflation/ballooning term permits the development of first-order shocks, i.e., discontinuities in orientation of the boundary of a shape, where the derivative is not defined. Thus, we have to approximate the spatial derivative using the upwind finite difference scheme [28]. The "doublet" term permits the development of discontinuities which indicate the presence of object boundaries. In this case, we can not use central finite differences but have to use forward or backward finite differences adaptively so that their directions are always away from the discontinuities. Let

$$\begin{aligned}
D_i^+\phi^n[i,j] &= \phi^n[i+1,j] - \phi^n[i,j] \\
D_i^-\phi^n[i,j] &= \phi^n[i,j] - \phi^n[i-1,j] \\
D_j^+\phi^n[i,j] &= \phi^n[i,j+1] - \phi^n[i,j] \\
D_j^-\phi^n[i,j] &= \phi^n[i,j] - \phi^n[i,j-1] \\
D_i\phi^n[i,j] &= (\phi^n[i+1,j] - \phi^n[i-1,j])/2 \\
D_j\phi^n[i,j] &= (\phi^n[i,j+1] - \phi^n[i,j-1])/2
\end{aligned}$$

then,

$$\begin{aligned}
(\nabla g \cdot \nabla u)[i,j] &= max(D_i g^n[i,j],0) D_i^- u^n[i,j] \\
&+ min(D_i g^n[i,j],0) D_i^+ u^n[i,j] \\
&+ max(D_j g^n[i,j],0) D_j^- u^n[i,j] \\
&+ min(D_j g^n[i,j],0) D_j^+ u^n[i,j]
\end{aligned}$$

and

$$\begin{aligned}
|\nabla u| = \{&(max(D_i^- u^n[i,j],0))^2 + (min(D_i^+ u^n[i,j],0))^2 \\
&+ (max(D_j^- u^n[i,j],0))^2 + (min(D_j^+ u^n[i,j],0))^2\}^{\frac{1}{2}}
\end{aligned}$$

For a detailed discussion on this scheme, we refer the reader to Osher and Sethian [28], and Malladi et al. [21].

The discrete equations for the scale-space tracking are realized using an explicit Euler step in which time derivatives are approximated using forward difference formulas in time with a step size Δt. This gives us the following discrete equations:

$$\sigma^{t+1} = \sigma^t - (\Delta t) c_1 \exp -c_2 |\nabla E_v^t(u, \sigma)| \tag{50}$$

$$u^{t+1} = u^t - (\Delta t) \nabla E_v^t(u^t, \sigma^t). \tag{51}$$

These equations were implemented using a multi-grid technique [35, 13, 34]. The multi-grid implementation speeds up the solution and leads to quality smoothing and segmentation results.

5. Experimental results. In this section we present the application of our model for smoothing and segmentation of several image data sets. To test the effectiveness of this model, we first choose a synthetic image. Figure 2(a) shows the noiseless image which contains a triangle, a rectangle, a circle and a thin ellipse. Figure 2(b) is obtained by adding Gaussian noise to (a), and the signal-to-noise ratio (SNR: the ratio between the variance of the noise-free image and the variance of noise [20]) is 1:2. Figure 2(c)-(e) shows the smoothing and segmentation results of the noisy synthetic image using our nonlinear diffusion model, where (c) is the smoothed image u, (d) shows the magnitude of $g(\nabla u)$, and (e) depicts the edge map which is the local minima in $g(\nabla u)$ obtained using the non-minima suppression method [14]. The various parameter settings used to obtain these results were, $\beta = 0.01$, $K = 200.0$ (note that $\sqrt{K}$ is the contrast parameter), number of iterations is 250, and the total computing time on an Ultra Sparc–1 is 180 seconds. Note that our smoothing model has preserved the edges with minimal rounding of the sharp corners. The local maxima in $g(\nabla u)$ yields a very good quality segmentation considering the amount of noise in the input data. In this and all the following examples, the stopping criterion for our nonlinear smoothing algorithm was a user specified tolerance (10^{-3}) on the norm of the difference between two consecutive iterates of u.

The second example is the smoothing and segmentation results of the popular Lenna image wherein we artificially added Gaussian noise ($SNR = 3 : 1$) to the image. The original Lenna image is shown in figure 3(a), (b) is the noisy version of (a) obtained by adding Gaussian noise, (c) depicts the smoothed image u, (d) and (e) shows the magnitude of $g(\nabla u)$ and the edge map. The parameter settings used to obtain these results were, $\beta = 0.01$, $K = 200.0$, number of iterations is 120, and the total computing time on an Ultra Sparc–1 is 85 seconds. Once again, the results of smoothing and segmentation are of high quality and the later may be used as a caricature of the original image leading to considerable image data compression.

Figure 4 presents the smoothing and segmentation of a CT chest scan. In this figure, (a) is the CT chest scan, (b) depicts the smoothed image u, (c) shows the magnitude of $g(\nabla u)$ and (d) is the edge map. The parameter settings used to obtain these results were, $\beta = 0.005$, $K = 200.0$, number of iterations is 10, and the total compute time on an Ultra Sparc–1 is 15 seconds. We can see that some of the "important" details which are unclear in the original image are enhanced in the smoothing and segmentation results.

Each of the following examples in figures $5 - 7$ presents four images: (a) is the original image, (b) shows the smoothing result, and (c) and (d) are the segmentation results from our nonlinear diffusion model. These images are infra-red images for

automatic target tracking applications and some of them are the Texas Instruments (TI) high value target acquisition images.

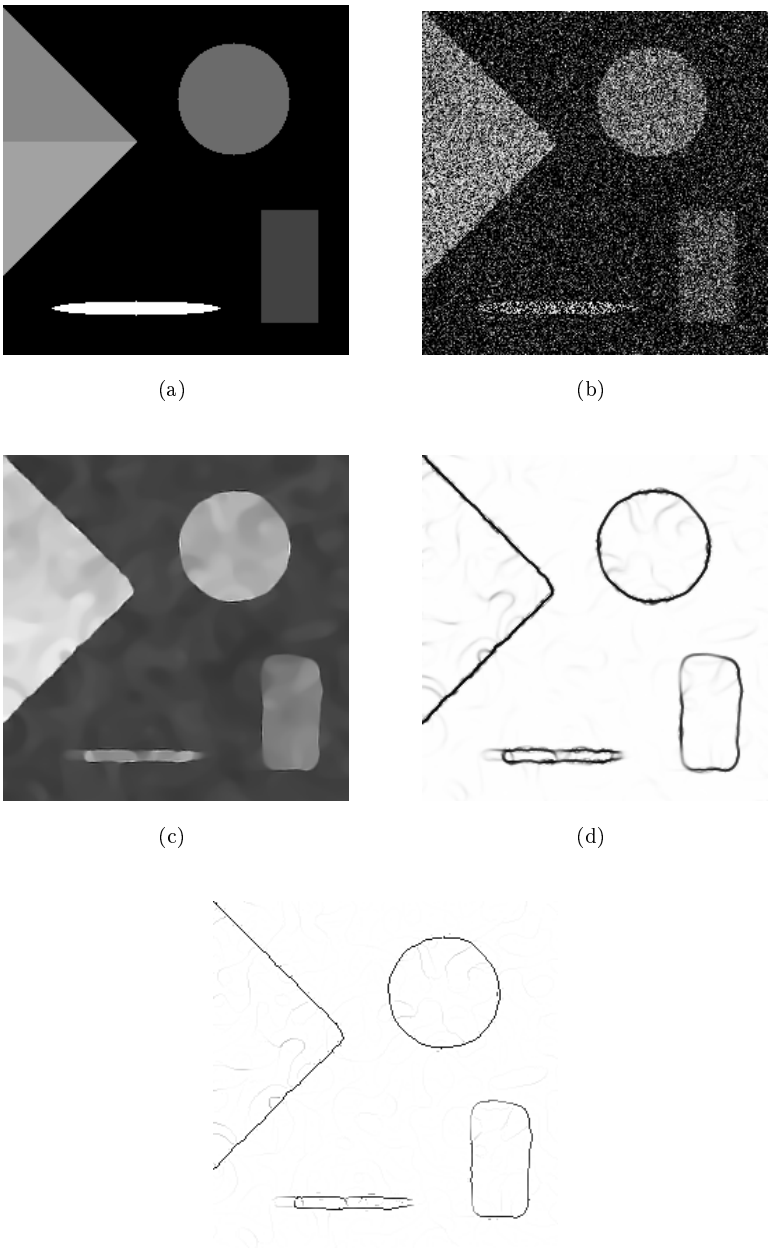
6. Conclusion. We have proposed a new PDE-based image denoising and segmentation algorithm based on nonlinear diffusion augmented by reactive terms. We prove the existence, uniqueness and stability of the viscosity solution of this model. For a fast implementation, we embed this model in a scale space and achieve scale space tracking via a dynamical system of coupled differential equations. The scale space tracking is implemented using a multi-grid scheme. We present experiments depicting the performance of our model on several image data sets. Our future efforts will be focused on a 3D implementation of our model and testing on volume images e.g., magnetic resonance (MR) and CT images.

Acknowledgments. Authors would like to thank Dr. P. Mergo for providing the CT images, and Dr. J. N. Wilson for providing the infra-red images.

REFERENCES

[1] L. Alvarez and J. Esclarín. Image quantization using reaction-diffusion equations. *SIAM J. Appl. Math.*, 57(1):153–175, 1997.
[2] L. Alvarez, P-L. Lions, and J-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. ii. *SIAM J. Numer. Anal.*, 29(3):845–866, June 1992.
[3] H. Brezis. *Analyse Fonctionnelle, Théorie et Applicatons.* Masson, Paris, 1987.
[4] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric method for active contours. *Numerische Mathematik*, 66, 1993.
[5] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Fifth International Conference on Computer Vision*, 1995.
[6] F. Catte, P. L. Lions, J. M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal of Numerical Analysis*, 29:182–193, 1992.
[7] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Two deterministic half-quadratic regularization algorithms for computed imaging,. In *in Proc. of the IEEE Intl. Conf. on Image Processing (ICIP),*, volume 2, pages 168–172. IEEE Computer Society Press, 1994.
[8] Y-G. Chen, Y. Giga, and S. Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *Journal of Differential Geometry*, 33:749–786, 1991.
[9] M. G. Crandall, H. Ishii, and P. L. Lions. User's guide to viscosity solution of second order partial differential equation. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992.
[10] L. C. Evans and J. Spruck. Motion of level sets by mean curvature. i. *J. Differential Geometry*, 33:635–681, 1991.
[11] D. Geiger and A. Yuille. A common framework for image segmentation,. *International Journal of Computer Vision,*, 6:227–243, 3 1991.
[12] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and bayesian restoration of images,. *IEEE Trans. on Pattern Analysis and Machine Intelligence,*, 6:721–741, 1984.
[13] W. Hackbusch. *Multi-Grid Methods and Applications,*. Springer-Verlag,, Berlin, 1985,.
[14] R. Jain, R. Kasturi, and B. G. Schunck. *Machine Vision.* McGraw-Hill, Inc., 1995.
[15] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. Journal of Computer Vision*, 1(4):321–332, 1988.
[16] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. In *Fifth International Conference on Computer Vision*, 1995.
[17] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker. Shapes, shocks, and deformations i: The components of shape and the reaction-diffusion space. Technical Report LEMS-105, LEMS, Division of Engineering, Brown University, 1992.
[18] R. Kimmel, N. Sochen, and R. Malladi. Images as embedding maps and minimal surfaces:movies, color and volumetric medical images,. In *in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition,*, pages 350–355, June 1997.
[19] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Uraltseva. *Linear and Quasilinear Equations of Parabolic Type.* American Mathematical Society, Providence, RI, 1968.

[20] J. S. Lim. *Two-Dimensional Signal and Image Processing*. Prentice Hall, Englewood Cliffs, NJ, 1990.
[21] R. Malladi, J. Sethian, and B. C. Vemuri. A fast level set based algorithm for topology-independent shape modeling. *J. Mathematical Imaging and Vision*, 6:269–289, 1996.
[22] R. Malladi and J. A. Sethian. A unified approach to noise removal, image enhancement and shape recovery. *IEEE Trans. on Image Processing*, 5(11):1554–1568, 1996.
[23] R. Malladi, J. A. Sethian, and B. C. Vemuri. Evolutionary fronts for topology independent shape modeling and recovering. In *Proc. ECCV*, 1995.
[24] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Trans. PAMI*, 17(2):158–175, 1995.
[25] T. Mclnerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1(2), 1996.
[26] M. Nitzberg and T. Shiota. Nonlinear image filtering with edge and corner enhancement,. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8):826–832, 1992.
[27] P. J. Olver, G. Sapiro, and A. Tannenbaum. Invariant geometric evolutions of surfaces and volumetric smoothing. Preprint, 1995.
[28] S. J. Osher and J. A. Sethian. Fronts propagation with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.
[29] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. PAMI*, 12(7):629–639, 1990.
[30] J. Shah. A common framework for curve evolution, segmentation and anisotropic diffusion. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 1996.
[31] H. Tek and B. B. Kimia. Image segmentation by reaction-diffusion bubbles. In *Fifth International Conference on Computer Vision*, 1995.
[32] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: Recovering 3d shape and nonrigid motion. *Artificial Intelligence*, 36:91–123, 1988.
[33] J. Weickert. A review of nonlinear diffusion filtering,. In (Eds.) B. ter Haar Romney, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-space theory in computer vision,*, volume 1252, of *Lecture Notes in Computer Science,*, pages 3–28. Springer-Verlag, 1997.
[34] G. Whitten. Scale space tracking and deformable sheet models for computational vision. *IEEE Trans. Patt. Anal. Machine Intell.*, 15(7):697–706, July 1993.
[35] A. Witkin, D. Terzopoulos, and M. Kass. Signal matching through scale space. *Int. J. Comput. Vision*, 1:133–144, 1987.



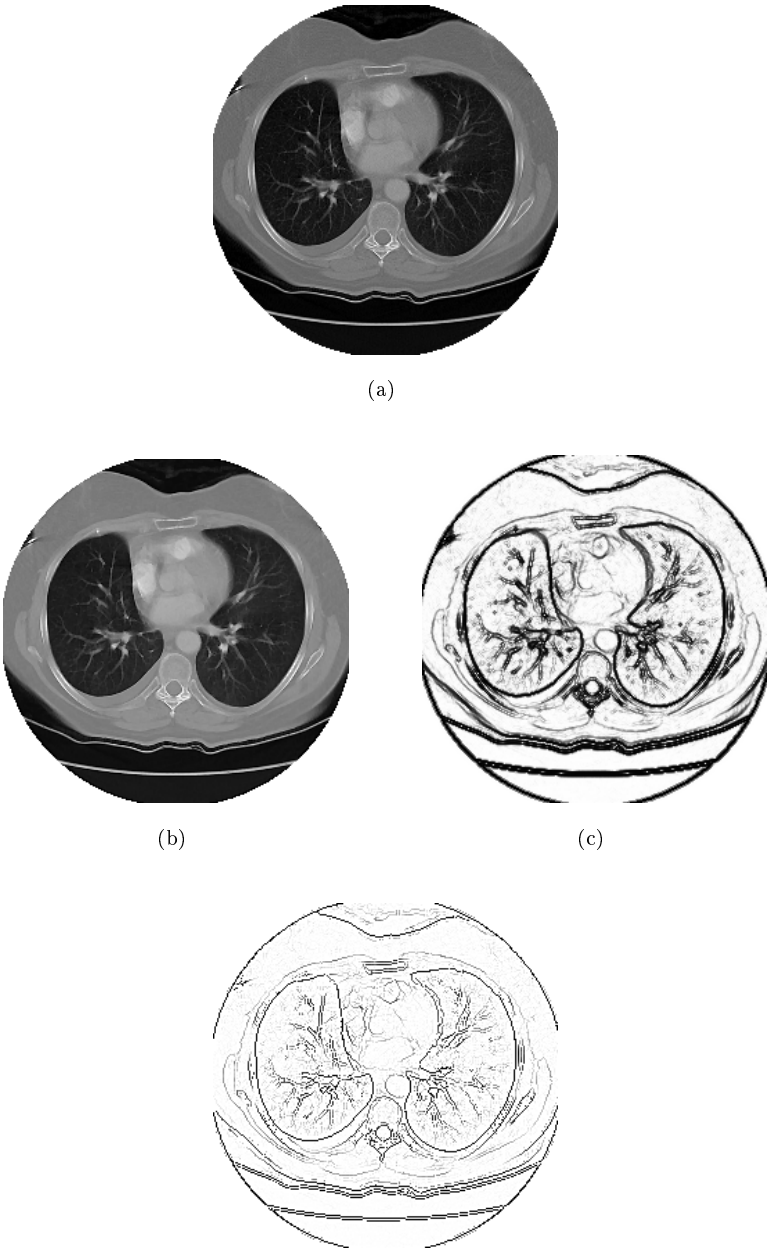
(a) (b)

(c) (d)

(e)

FIG. 2. *Smoothing and segmentation results of a noisy synthetic image* (256×256). *(a) Noise-free image; (b) noisy image; (c) smoothed image* u*; (d) magnitude of* $g(\nabla u)$*; (e) edge map.*

FIG. 3. *Smoothing and segmentation results of the noisy Lenna image* (256×256) *using single-PDE-based model. (a) Original Lenna image; (b) Lenna image with Gaussian noise; (c) smoothed image u; (d) magnitude of* $g(\nabla u)$*; (e) edge map.*



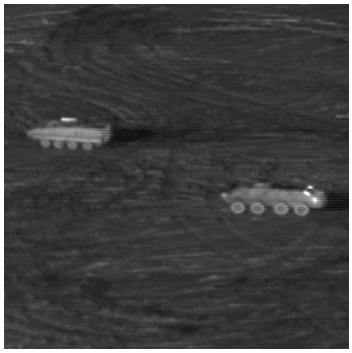
(a)

(b)

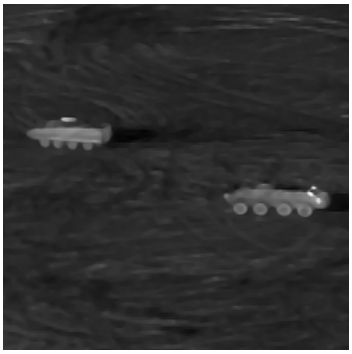
(c)

(d)

FIG. 4. *CT chest scan smoothing and segmentation. (a) Original image (*512×512*); (b) smoothed image* u*; (c) magnitude of* $g(\nabla u)$*; (d) edge map.*

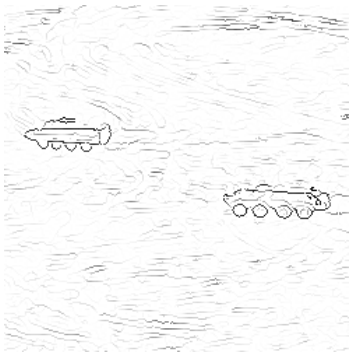


(a)



(b)

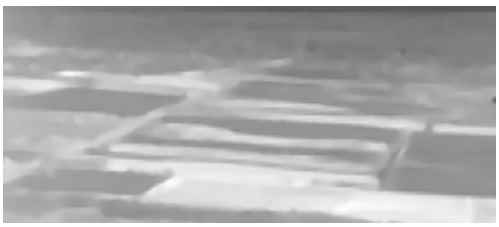
(c)



(d)

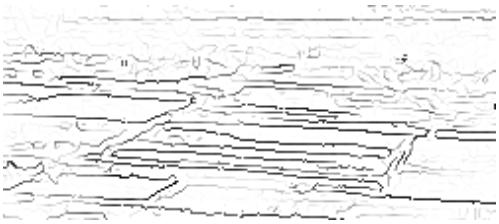
FIG. 5. *Smoothing and segmentation results of tanks image. (a) Original image (*256×256*); (b) smoothed image* u*; (c) magnitude of* $g(\nabla u)$*; (d) edge map. Parameters:* $\beta = 0.05$*,* $K = 200.0$*, number of iterations is 20, and the total computing time on an Ultra Sparc–1 is 7 seconds.*

(a)



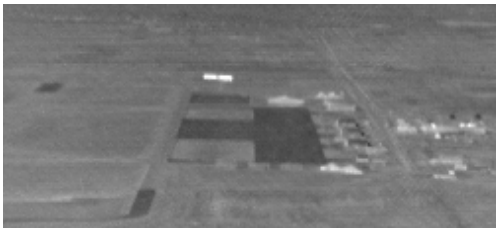
(b)

(c)

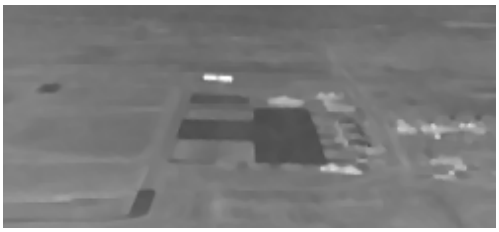


(d)

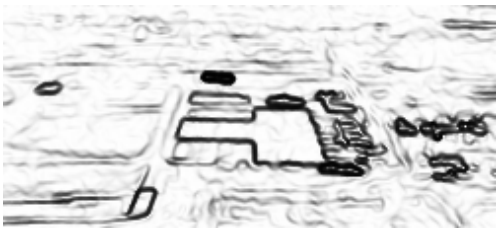
FIG. 6. *Smoothing and segmentation results of one of the TI Infra-Red High-Value Target Acquisition (IRHVTA) Data. (a) Original image* (256×256)*; (b) smoothed image* u*; (c) magnitude of* $g(\nabla u)$*; (d) edge map. Parameters:* $\beta = 0.05$, $K = 100.0$, *number of iterations is 20, and the total computing time on an Ultra Sparc–1 is 7 seconds.*



(a)



(b)



(c)

(d)

FIG. 7. *Smoothing and segmentation results of one of TI InfraRed High-Value Target Acquisition (IRHVTA) Data. (a) Original image* (256×256)*; (b) smoothed image u; (c) magnitude of* $g(\nabla u)$*; (d) edge map. Parameters:* $\beta = 0.05$, $K = 200.0$, *number of iterations is 10, and the total computing time on an Ultra Sparc–1 is 4 seconds.*