

Generalizing Mumford-Shah Model for Multiphase Piecewise Smooth Image Segmentation

Ying Gu, Wei Xiong, Li-Lian Wang, and Jierong Cheng

Abstract—This paper concerns multiphase piecewise smooth image segmentation with intensity inhomogeneities. Traditional methods based on the Mumford–Shah (MS) model require solving complicated diffusion equations evolving in irregular subdomains, leading to significant difficulties in efficient and accurate segmentation, especially in multiphase scenarios. In this paper, we propose a general framework to modify the MS model by using smoothing operators that can avoid the complicated implementation and inaccurate segmentation of traditional approaches. A detailed analysis connecting the smoothing operators and the diffusion equations is given to justify the modification. In addition, we present an efficient algorithm based on the direct augmented Lagrangian method, which requires fewer parameters than the commonly used augmented Lagrangian method. Typically, the smoothing operator in the general model is chosen to be Gaussian kernel, the bilateral kernel, and the directional diffusion kernel, respectively. Ample numerical results are provided to demonstrate the efficiency and accuracy of the modified model and the proposed minimization algorithm through various comparisons with existing approaches.

Index Terms—Mumford-Shah (MS) model, multiphase image segmentation, smoothing operator, Gaussian, bilateral, directional diffusion, augmented Lagrangian method.

I. INTRODUCTION

IMAGE segmentation aims to partition a given image into a number of constituent regions so as to extract objects of specific, distinctive features from image background. It finds ubiquitous applications (see, e.g., [1], [2]), where an important issue is how to segment real images with intensity inhomogeneities. It is however noteworthy that intensity-based segmentation based on the piecewise constant models [3]–[5] often leads to unsatisfactory segmentation in such occasions. A feasible approach is to assume an image consisting of piecewise smooth phases, so one can adopt the region-based piecewise smooth (PS) approximation model [6]–[9], which leads to better segmentation.

While intensive research has been devoted to piecewise constant image segmentation, limited works are available for

image segmentation with PS structures. In this work, we are committed to PS image segmentation using variational approaches. Under a variational framework, image segmentation is accomplished by solving a minimisation problem involving an energy functional that mathematically incorporates various necessary *a priori* or *a posteriori* information of a given image. Among those variational models for this purpose [4], [5], [10], [11], the Mumford-Shah (MS) model [4] becomes the most popular one. Basically, given an image I defined in $\Omega \subset \mathbb{R}^2$, we search for a piecewise smooth approximation u of I , and an edge set $\Gamma (\subset \Omega)$ such that

$$\min_{u, \Gamma} \left\{ E_{MS}(u, \Gamma) := |\Gamma| + \mu \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \frac{\nu}{2} \int_{\Omega} |u - I|^2 dx \right\}, \quad (1.1)$$

where $\mu, \nu > 0$ are parameters, and $|\Gamma|$ is the length of Γ . The third term in the energy functional $E_{MS}(u, \Gamma)$ is the data fidelity term that ensures u to be close to the input image I ; while the second term penalises strong variations of u that enhances the smoothness of u . Lastly, through penalising the length of the edges, the model also favours to find a “tight” enclosure of the smooth regions. It is seen that the model involves two unknowns $\{u, \Gamma\}$ in different natures, which presents a great challenge for both mathematical analysis and numerical solution. We refer to [12] and the references therein for a review of many insightful analysis and various attempts of approximating, relaxing or simplifying the MS model.

In order to solve (1.1), two important issues are: how to represent the edge set Γ and how to relate it to the unknown u . Hereafter, we treat the edges as rectifiable closed curves (cf. [13]). Accordingly, the edge set Γ separates Ω into m disjoint $\{\Omega_i\}_{i=1}^m$, i.e.

$$\Omega = \bigcup_{i=1}^m \Omega_i, \quad \Omega_i \cap \Omega_j = \emptyset \quad \text{for } i \neq j,$$

and I is approximated by a smooth function u_i in each Ω_i . Hence, we can rewrite the original MS model (1.1) as

$$\min_{\{u_i, \Omega_i\}_{i=1}^m} \left\{ E_{MS}(\{u_i, \Omega_i\}_{i=1}^m) := \sum_{i=1}^m \left(|\partial \Omega_i| + \mu \int_{\Omega_i} |\nabla u_i|^2 dx + \frac{\nu}{2} \int_{\Omega_i} |u_i - I|^2 dx \right) \right\}. \quad (1.2)$$

Some attempts have tackled this piecewise smooth image segmentation problem from a continuous perspective by minimizing the MS model (1.2) (see, e.g., [6]–[9], [14], [15] and the references therein). Most notably, Vese and Chan [7] used the classical level-set method [16] involving multiple phase functions $\{\phi_i\}$ to describe the edge Γ . It requires evolving

Manuscript received April 8, 2016; revised August 25, 2016 and November 8, 2016; accepted November 26, 2016. Date of publication December 6, 2016; date of current version January 5, 2017. The work of L.-L. Wang was supported in part by the Singapore MOE AcRF Tier 1 under Grant RG 15/12 and Grant RG27/15, and in part by the Singapore MOE AcRF Tier 2 under Grant MOE 2013-T2-1-095, ARC 44/13. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yonggang Shi.

Y. Gu, W. Xiong, and J. Cheng are with the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore 138632 (e-mails: guy@i2r.a-star.edu.sg; xiongw@i2r.a-star.edu.sg; chengjr@i2r.a-star.edu.sg).

L.-L. Wang is with the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371 (e-mail: lilian@ntu.edu.sg).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2016.2636450

the coupled system of $\{u_i, \phi_i\}$, where the Euler Lagrangian equation with respect to u_i of (1.2) is

$$u_i - \frac{2\mu}{v} \Delta u_i = I \quad \text{in } \Omega_i, \quad (1.3)$$

where $\Delta = \text{div}(\nabla)$ and Ω_i is determined by the level-set function ϕ_i for $i = 1, \dots, m$.

Observe that a set of PDEs for $\{u_i\}$ has to be solved in each subdomain Ω_i , which is time consuming and at times, inaccurate in segmentation for many phases. There are some other region-based methods using the MS model for PS image segmentation. For example, Piovano *et al.* [8] solved the segmentation problem by using local statistics formulation, and evolved the contour through local averaging of image intensities inside and outside the contour separately. Brox and Cremers [17] used the statistical interpretation of the MS functional where the piecewise smooth approximation was evolved locally as with [7]. In a nutshell, in the aforementioned typical PS MS models, the smooth approximation u_i was numerically solved in each irregular subdomain, and the algorithms were computationally time consuming. These motivated the efforts in developing models for solving the smooth functions in the whole domain. In particular, Li *et al.* [18] proposed a region-based active contour model with a classical level-set formulation for two phases. A more efficient minimisation algorithm was reported in Gu *et al.* [19] for an extended model for multiple phases. More recently, Zhang *et al.* [20] developed a statistical active contour model based on the MS model, which integrated a bias field with a classical level-set formulation. In addition, Ali *et al.* [21] proposed a variational model with two fitting terms of the region and edge enhanced quantities.

There has been much recent interest in processing images with intensity inhomogeneity (see, e.g., [22]–[26]). Under the assumption that the “smooth” bias field (i.e., intensity inhomogeneity) $b(\approx 1)$ as “small” perturbation of piecewise constant intensities, the following model is proposed (cf. [22], [23], [27], [28]):

$$I(\mathbf{x}) = b(\mathbf{x})c_i + n(\mathbf{x}), \quad \mathbf{x} \in \Omega_i, \quad (1.4)$$

where n is the noise. With a modification of the MS model, Li *et al.* [23] introduced the energy minimization problem:

$$\min_{b, \{c_i, \Omega_i\}_{i=1}^m} \left\{ \sum_{i=1}^m \left(\frac{v}{2} \int_{\Omega_i} (I - bc_i)^2 d\mathbf{x} + \frac{v\lambda}{2} \int_{\Omega_i} (I - c_i)^2 d\mathbf{x} + |\partial\Omega_i| \right) + \mu \int_{\Omega} |\nabla b|^2 d\mathbf{x} \right\}, \quad (1.5)$$

where $v, \lambda, \mu > 0$ are constants. Very recently, Spencer and Chen [26] further incorporated a stabilised term to enforce b to be close to 1. From a different perspective, we reported in [24] the decomposition model:

$$I(\mathbf{x}) = \lambda b(\mathbf{x}) + (1 - \lambda) \sum_{i=1}^m c_i \chi_i(\mathbf{x}) + n(\mathbf{x}), \quad (1.6)$$

a.e. in Ω , where χ_i is the indicator function of Ω_i (cf. (3.1)), and $\lambda > 0$ is a constant. Basically, a clear image $I - n$ is viewed as a combination of piecewise constant and smooth

components. The associated minimisation problem for this decomposition in [24] reads

$$\min_{b, \{c_i, \Omega_i\}_{i=1}^m} \left\{ v\lambda^2 \int_{\Omega} |I - b|^2 d\mathbf{x} + \sum_{i=1}^m \left(\mu\lambda^2 \int_{\Omega_i} |\nabla b|^2 d\mathbf{x} + (1 - \lambda)^2 v \int_{\Omega_i} |I - c_i|^2 d\mathbf{x} + |\partial\Omega_i| \right) \right\}. \quad (1.7)$$

The recent work of Jung [25] further modified the L^2 -terms involving I in (1.7) to L^1 -data fidelity terms.

With the understanding that the gradient descent flow related to (1.3) can be expressed as an integral form of a diffusion operator, we reported in [29] the idea of replacing the second term in E_{MS} of the MS model (1.2) by an integral operator, where only preliminary results were presented and some advantages were observed.

Inspired by [29], we explore the idea of using smoothing operators to modify the MS model for fast and accurate multiphase PS image segmentation. Based on the principle, we can derive new variational models, e.g., for segmentation of images with density inhomogeneity (see (2.9) below), and for image decomposition (see (2.11) below). We intend to provide a general framework together with basic principles for selecting the involved smoothing operators and some mathematical justifications. We outline the main contributions of this paper as follows:

- We provide some insights into the connections between the diffusion equations and the modified MS models with smoothing operators. The analysis leads to some basic criteria for the choice of smoothing operators and justifies the proposed simplifications.
- With the aid of the smoothing operators, the complicated implementation of (1.3) can be replaced by some closed-form formulas. Meanwhile, the convolution can be solved by the fast Fourier transform (FFT). Moreover, the use of the indicator functions for multiple phases can be solved by the direct approach in [30] without convex relaxation. With all these at our disposal, we can develop very efficient minimisation algorithms for the modified MS model for PS image segmentation with many phases.
- The proposed general model together with the algorithm outperforms many existing methods, and can further incorporate additional *a priori* and *a posteriori* information of images for other related image processing tasks.

The rest of the paper is organised as follows. In Section II, we present the basic idea of the approach in [29], and then introduce two smoothing kernels. With some analysis of the connection between the diffusion equations and these kernels, we provide a general framework using the smoothing operators to modify the MS model. We develop efficient algorithms for the general model in Section III based on a splitting technique and the direct augmented Lagrangian method. In Section IV, we focus on the modified MS models with three different smoothing operators: the Gaussian kernel, bilateral kernel and the directional diffusion kernel, and then show their performances of the models and minimisation algorithms for different images, and compare them with other existing

methods. Some important notation used in this paper is listed as follows:

- Boldface letters for vectors, e.g., $\mathbf{x} = (x_1, x_2)$ and $\mathbf{u} = (u_1, \dots, u_m)$;
- ∂ : partial derivative operator;
- $*$: the convolution;
- ∇ : the gradient operator;
- Δ : the Laplacian operator;
- v^0 : initial condition of v .

II. A GENERAL MODEL

We first present the basic idea reported in [29]. Note that the gradient descent flow of the MS model (1.2) is given by

$$\frac{\partial u_i}{\partial t} = \frac{2\mu}{v} \Delta u_i - (u_i - I) \quad \text{in } \Omega_i, \quad (2.1)$$

for $1 \leq i \leq m$. Vese and Chan [7] proposed to solve (2.1) directly by using explicit methods through an iteration process in each subdomain. As the regions are time-dependent, i.e., $\Omega_i = \Omega_i(t)$, the whole numerical algorithm is time consuming and at times, unstable for many phases. It is important to observe that (2.1) can be associated with the diffusion process

$$\begin{cases} \frac{\partial u_i}{\partial t} = \frac{2\mu}{v} \Delta u_i & \text{in } \Omega_i, \\ u_i|_{t=0} = I. \end{cases} \quad (2.2)$$

If $\Omega_i = \mathbb{R}^2$, the solution of (2.2) is

$$u_i = G_\sigma * I = \int_{\mathbb{R}^2} G_\sigma(\mathbf{x} - \mathbf{y}) I(\mathbf{x}) d\mathbf{x}, \quad (2.3)$$

where

$$G_\sigma(\mathbf{x}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{|\mathbf{x}|^2}{2\sigma^2}\right), \quad \sigma = \sqrt{\frac{4\mu t}{v}}. \quad (2.4)$$

Within the image domain of $I(\mathbf{x})$, each subregion Ω_i is bounded. In view of this, we can propose a normalized approximate solution to (2.2):

$$u_i(\mathbf{y}) = \frac{\int_{\Omega_i} G_\sigma(\mathbf{x} - \mathbf{y}) I(\mathbf{x}) d\mathbf{x}}{\int_{\Omega_i} G_\sigma(\mathbf{x} - \mathbf{y}) d\mathbf{x}}, \quad 1 \leq i \leq m. \quad (2.5)$$

Remark 1: It is worthwhile to point out that

- Formula (2.5) can be viewed as a normalised and locally weighted average of the image within Ω_i (cf. [31]).
- We propose to use (2.5) to replace the complicated evolution of u_i in (2.1) and the convolution can be computed by the fast Fourier transform. This is essential for the efficient algorithm to be introduced in the forthcoming section.

Note that we can rewrite (2.5) as

$$\iint_{\Omega_i \times \Omega_i} G_\sigma(\mathbf{x} - \mathbf{y}) (I(\mathbf{x}) - u_i(\mathbf{y})) d\mathbf{y} d\mathbf{x} = 0, \quad (2.6)$$

which turns out to the equation resulted from the optimality condition with respect to u_i of the minimisation problem:

$$\min_{\{u_i, \Omega_i\}_{i=1}^m} \sum_{i=1}^m \left(|\partial\Omega_i| + \frac{v}{2} \iint_{\Omega_i \times \Omega_i} G_\sigma(\mathbf{x} - \mathbf{y}) |u_i(\mathbf{y}) - I(\mathbf{x})|^2 d\mathbf{y} d\mathbf{x} \right). \quad (2.7)$$

In contrast to (1.2), the smooth term and fidelity term therein are replaced by the new fidelity term involving the Gaussian smoothing operator.

We elaborate more on this idea and revisit the model (1.4) for density inhomogeneity. Note that the gradient descent flow of (1.5) with respect to b (for fixed c_i) is

$$\frac{\partial b}{\partial t} = \frac{2\mu}{vc_i^2} \Delta b - \left(b - \frac{I}{c_i}\right) \quad \text{in } \Omega_i. \quad (2.8)$$

Following (2.2)-(2.3), we can formally write $b = G_\sigma * (I/c_i)$ with $\sigma = \sqrt{4\mu t/(vc_i^2)}$. Like (2.7), we can derive the following model alternative to (1.5):

$$\min_{b, \{c_i, \Omega_i\}_{i=1}^m} \left\{ \sum_{i=1}^m \left(\frac{v\lambda}{2} \int_{\Omega_i} (I - c_i)^2 d\mathbf{x} + |\partial\Omega_i| \right) + \frac{v}{2} \iint_{\Omega_i \times \Omega_i} G_\sigma(\mathbf{x} - \mathbf{y}) (I(\mathbf{x}) - b(\mathbf{y})c_i)^2 d\mathbf{y} d\mathbf{x} \right\}. \quad (2.9)$$

We see that the segmentation of images with density inhomogeneity can be modelled by this framework.

In fact, the same argument applies to the decomposition model (1.6). The counterpart of (2.8) is

$$\frac{\partial b}{\partial t} = \frac{\mu}{v} \Delta b - (b - I) \quad \text{in } \Omega_i. \quad (2.10)$$

Thus alternative to (1.7), we have

$$\min_{b, \{c_i, \Omega_i\}_{i=1}^m} \left\{ \sum_{i=1}^m \left(v(1 - \lambda)^2 \int_{\Omega_i} (I - c_i)^2 d\mathbf{x} + |\partial\Omega_i| \right) + v\lambda^2 \iint_{\Omega_i \times \Omega_i} G_\sigma(\mathbf{x} - \mathbf{y}) (I(\mathbf{x}) - b(\mathbf{y}))^2 d\mathbf{y} d\mathbf{x} \right\}. \quad (2.11)$$

To the best of our knowledge, the models (2.9) and (2.11) appear new. Our general algorithm to be presented can be directly applied to their minimisation.

In view of the above, we further consider the modification of the original MS model (1.2) by using a more general smoothing operator with K in place of G_σ in (2.7), that is,

$$\min_{\{u_i, \Omega_i\}_{i=1}^m} \sum_{i=1}^m \left(|\partial\Omega_i| + \frac{v}{2} \iint_{\Omega_i \times \Omega_i} K(\mathbf{x}, \mathbf{y}) |u_i(\mathbf{y}) - I(\mathbf{x})|^2 d\mathbf{y} d\mathbf{x} \right). \quad (2.12)$$

A. General Properties

In general, smoothing operation on a function f can be formulated as

$$T_K f(\mathbf{x}) = \int K(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}, \quad (2.13)$$

where K is a nonnegative kernel function $K : \mathbb{R}^2 \rightarrow [0, +\infty)$. Such a smoothing operator should have the following properties.

- Normalization:

$$\int K(\mathbf{x}) d\mathbf{x} = 1;$$

- Symmetry

$$T_{K(x,y)}f = T_{K(y,x)}f;$$

- For f and g in $C_0^\infty(\mathbb{R}^2)$, we have

$$\int T_K f(x)g(x)dx = \int f(x)T_K g(x)dx;$$

- The composition of two smoothing operators is still a smoothing operator.

$$\int K_3(x, y)f(y)dy = \int K_1(x, y)K_2(z, y)f(z)dzdy;$$

- Localization

$$K(x) \geq K(y), \text{ if } |x| \leq |y|$$

and

$$\lim_{|x| \rightarrow \infty} K(x) = 0.$$

Our basic idea is to use the smoothing operators K to replace the smooth term and the data fidelity term, that avoids solving complicated diffusion equations (1.3). There are some smoothing methods proposed to smooth u_i , e.g., those in [32]. In this work, we mainly consider the Gaussian-type smoothing operators. In the following, we introduce two popular Gaussian-type kernels, a bilateral kernel and a directional diffusion kernel, and study the relationship between these two kernels and the diffusion equations. Such relationship helps us justify the condition of the smoothing operators K to modify the MS model.

B. Bilateral Kernel and Directional Diffusion Kernel

- **Bilateral kernel.** Gaussian kernel is a low-pass filtering. It may blur the image edges in the smoothing effect. Hence, some nonlinear filters are constructed to prevent from the blurring problem, for example, [33] and [34] use the bilateral kernel. The bilateral kernel combines both the position domain and range domain filtering;

$$\begin{aligned} G_B(x, y) &= \frac{1}{C_B(x)} G_{\sigma_d}(x - y) G_{\sigma_r}(I(x) - I(y)) \\ &= \frac{1}{C_B(x)} \exp\left(-\frac{|x - y|^2}{2\sigma_d^2} - \frac{|I(x) - I(y)|^2}{2\sigma_r^2}\right), \end{aligned} \quad (2.14)$$

where

$$C_B(x) = \int_{\Omega} \exp\left(-\frac{|x - y|^2}{2\sigma_d^2} - \frac{|I(x) - I(y)|^2}{2\sigma_r^2}\right) dy,$$

and $\sigma_d, \sigma_r > 0$ are control parameters, and $G_{\sigma_r}(I(x) - I(y))$ is the Gaussian function for intensity domain, which can reduce the influence of the intensity dissimilarities. The bilateral kernel can distinguish well the flow influence from different regions [35], [36]. The bilateral kernel plays the role of standard domain filter, producing faster diffusion inside the homogeneous region and slower diffusion across the edge of the region. Hence, image edges can be preserved.

- **Directional diffusion kernel.** The bilateral kernel is a nonlinear filter and has the difficulty in reproducing linear functions [37]. In order to handle such difficulties, Wang [38] constructed a new directional diffusion kernel. They replace the term $I(x) - I(y)$ in the bilateral kernel (2.14) with $\langle \nabla I(x), (x - y) \rangle$, where $\langle \cdot, \cdot \rangle$ is the inner product in \mathbb{R}^2 . Hence, the directional diffusion kernel is given as

$$G_D(x, y) = \frac{1}{C_D(x)} \exp\left(-\frac{|x - y|^2}{4\sigma} - \frac{\langle \nabla I(x), (x - y) \rangle^2}{4\sigma^2}\right), \quad (2.15)$$

where

$$C_D(x) = \int_{\Omega} \exp\left(-\frac{|x - y|^2}{4\sigma} - \frac{\langle \nabla I(x), (x - y) \rangle^2}{4\sigma^2}\right) dy,$$

and $\sigma > 0$ is a parameter. Note that the term $\langle \nabla I(x), (x - y) \rangle^2 / (4\sigma^2)$ equals to 0, when $y - x$ points the horizontal direction of the level curve of I , and has the maximum value in the gradient direction. Hence, the directional diffusion kernel acts like Gaussian diffusion in the level curve direction and blocks the diffusion in the gradient direction. This kernel has been studied in [38].

C. Relation Between the Kernels and Diffusion Equations

We next provide some insights into the above kernels. Barash [36] pointed out the connection between the neighbourhood filters and the diffusion equations, and Buades *et al.* [37] provided rigorous justifications. Such a relationship was also studied in [32], [33], [38], and [39] from various perspectives. Here, we first give a detailed description of the relation between these Gaussian-type smoothing kernels (Gaussian kernel, bilateral kernel and the directional diffusion kernel) and the diffusion equations through the argument in these relevant references.

Proposition 1: Let $v \in C^2(\Omega)$, and define the output functional of the Gaussian kernel as

$$\mathcal{G}(v)(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\Omega} \exp\left(-\frac{|x - y|^2}{2\sigma^2}\right) v(y) dy.$$

Then $\mathcal{G}(v)$ has the asymptotic expansion

$$\mathcal{G}(v)(x) - v(x) \approx 2\sigma^2 \Delta v(x), \quad \sigma \rightarrow 0.$$

We can find that the asymptotic expansion is consistent with the diffusion equations (1.3).

Now, we consider the bilateral kernel (2.14) and the directional diffusion kernel (2.15) following [37]–[39]. We first introduce some notation for the study of these two kernels. At each $x = (x_1, x_2) \in \Omega$, let

$$\eta = \frac{\nabla v}{|\nabla v|} = \frac{(v_{x_1}, v_{x_2})}{\sqrt{v_{x_1}^2 + v_{x_2}^2}}, \quad \xi = \frac{\nabla v^\perp}{|\nabla v|} = \frac{(-v_{x_2}, v_{x_1})}{\sqrt{v_{x_1}^2 + v_{x_2}^2}},$$

for $\nabla v \neq 0$. Recall the second-order directional derivatives:

$$v_{\eta\eta} = \eta^T H(v) \eta, \quad v_{\xi\xi} = \xi^T H(v) \xi. \quad (2.16)$$

where T means the transpose and $H(v)$ is the Hessian matrix of u in the form

$$H(v) = \begin{pmatrix} v_{x_1 x_1} & v_{x_1 x_2} \\ v_{x_1 x_2} & v_{x_2 x_2} \end{pmatrix}.$$

Denote the local orthogonal basis associated with v by

$$\mathbb{B} = \{b_1, b_2\} = \begin{cases} \{\eta, \xi\}, & |\nabla v(\mathbf{x})| \neq 0, \\ \{e_1, e_2\}, & |\nabla v(\mathbf{x})| = 0, \end{cases}$$

where $\{e_1, e_2\}$ is the standard basis on \mathbb{R}^2 . In the following analysis, we use $\mathbb{B} = \{b_1, b_2\}$ as the local Cartesian coordinate system. Due to the similarity between the bilateral kernel (2.14) and the Yaroslavsky kernel [40], the asymptotic expansion of the bilateral kernel can be studied through the Yaroslavsky kernel, which is expressed in [37]. Note that the output functional of the Yaroslavsky kernel is

$$Y(v)(\mathbf{x}) = \frac{1}{C_Y(\mathbf{x})} \int_{\Omega} \exp\left(-\frac{|v(\mathbf{x}) - v(\mathbf{y})|^2}{\sigma_r^2}\right) v(\mathbf{y}) d\mathbf{y},$$

where σ_r is a positive constant and

$$C_Y(\mathbf{x}) = \int_{\Omega} \exp\left(-\frac{|v(\mathbf{x}) - v(\mathbf{y})|^2}{\sigma_r^2}\right) d\mathbf{y}.$$

The corresponding asymptotic expansion of the Yaroslavsky kernel is stated as follows. We refer to [37] for the detailed proof.

Proposition 2: Let $E(\omega) = 2 \int_0^\omega e^{-s^2} ds$,

$$g(\omega) = \begin{cases} \frac{1}{3E(\omega)} \omega e^{-\omega^2}, & \omega > 0, \\ 1/6, & \omega = 0, \end{cases}$$

and

$$f(\omega) = \begin{cases} \frac{1}{2\omega^2} [6g(\omega)(\omega^2 + 1) - 1], & \omega > 0, \\ 1/6, & \omega = 0. \end{cases}$$

Suppose that $v \in C^2(\Omega)$, and let $\sigma_d, \sigma_r > 0$ such that $\sigma_d, \sigma_r \rightarrow 0$ and $\sigma_r = O(\sigma_d)$. Then we have

$$Y(v)(\mathbf{x}) - v(\mathbf{x}) \approx \sigma_d^2 \left[f\left(\frac{\sigma_d}{\sigma_r} |\nabla v(\mathbf{x})|\right) v_{\eta\eta}(\mathbf{x}) + g\left(\frac{\sigma_d}{\sigma_r} |\nabla v(\mathbf{x})|\right) v_{\xi\xi}(\mathbf{x}) \right].$$

The output functional of the directional diffusion kernel (2.15) is

$$D(v)(\mathbf{x}) = \frac{1}{C_D(\mathbf{x})} \int_{\Omega} \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{4\sigma} - \frac{\langle \nabla v(\mathbf{x}), (\mathbf{x} - \mathbf{y}) \rangle^2}{4\sigma^2}\right) v(\mathbf{y}) d\mathbf{y},$$

where $\sigma > 0$ is a constant and

$$C_D(\mathbf{x}) = \int_{\Omega} \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{4\sigma} - \frac{\langle \nabla v(\mathbf{x}), (\mathbf{x} - \mathbf{y}) \rangle^2}{4\sigma^2}\right) d\mathbf{y}.$$

Here, we have the asymptotic expansion of this directional diffusion kernel (cf. [38], [39]).

Proposition 3: For $v \in C^2(\Omega)$, we have

$$D(v)(\mathbf{x}) - v(\mathbf{x}) \approx \sigma(L_{v,\mathbf{x}}v)(\mathbf{x}), \quad \sigma \rightarrow 0,$$

where

$$L_{v,\mathbf{x}}v = \begin{cases} \frac{\partial^2 v}{\partial \eta^2}, & \text{if } \nabla v(\mathbf{x}) \neq 0, \\ \Delta v, & \text{if } \nabla v(\mathbf{x}) = 0. \end{cases}$$

From Propositions 2-3, we find the corresponding diffusion equations of the bilateral kernel (2.14) and the directional diffusion kernel (2.15) can be written as

$$v_t = c_1 v_{\xi\xi} + c_2 v_{\eta\eta}. \quad (2.17)$$

It is known from Buades *et al.* [37] that the neighborhood filter can be linked to the Perona-Malik model [41]–[43].

$$v_t = \text{div}(g(|\nabla v|^2) \nabla v), \quad (2.18)$$

where g is decreasing such that

$$g(0) = 1, \quad \lim_{s \rightarrow +\infty} g(s) = 0, \quad s = |\nabla v|^2.$$

Note that the evolution of u_i of the smooth term in the Mumford-Shah model (1.2) can be expressed as

$$\frac{\partial u_i}{\partial t} = \text{div}(c(\mathbf{x}) \nabla u_i), \quad (2.19)$$

where $c(\mathbf{x})$ can be taken as a nonlinear diffusion coefficient $c(\mathbf{x}) = c(|\nabla u_i|)$. This coefficient $c(|\nabla u_i|) \rightarrow 0$ when $|\nabla u_i| \rightarrow \infty$. Hence, it has a small weightage when the gradient is very large (i.e., across the edges). From the above analysis, (2.19) can be formally written as

$$\frac{\partial u_i}{\partial t} = \hat{c}_1 \frac{\partial^2 u_i}{\partial \xi \partial \xi} + \hat{c}_2 \frac{\partial^2 u_i}{\partial \eta \partial \eta}. \quad (2.20)$$

Observe from (2.17) and (2.20) that the diffusion equations of the kernels (2.14) and (2.15) are in the same form as that of the equation reduced from the MS model. As such, these two kernels can play the role as the smoothing operators to modify the MS model.

III. MINIMIZATION ALGORITHM

Next, we will provide the minimization algorithms for the general minimisation problem (2.12).

Denote the indicator function of Ω_i by

$$\chi_i(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \Omega_i, \\ 0, & \text{if } \mathbf{x} \in \Omega \setminus \Omega_i. \end{cases} \quad (3.1)$$

The length of Γ can be characterized by the total variation (TV) [44], that is,

$$|\partial \Omega_i| = TV(\chi_i) = \int_{\Omega} |\nabla \chi_i| d\mathbf{x}.$$

Then we can reformulate (2.12) as

$$\min_{\mathbf{u}, \chi} \left\{ E(\mathbf{u}, \chi) := \sum_{i=1}^m \int_{\Omega} |\nabla \chi_i(\mathbf{x})| d\mathbf{x} + \frac{\nu}{2} \iint_{\Omega \times \Omega} K(\mathbf{x}, \mathbf{y}) |u_i(\mathbf{y}) - I(\mathbf{x})|^2 \chi_i(\mathbf{x}) d\mathbf{y} d\mathbf{x} \right\}, \quad (3.2)$$

where $\mathbf{u} = (u_1, \dots, u_m)$ and $\chi = (\chi_1, \dots, \chi_m)$.

Next we use the splitting technique as follows:

(a). Fixing χ , find \mathbf{u} such that

$$\min_{\mathbf{u}} \sum_{i=1}^m \iint_{\Omega \times \Omega} K(\mathbf{x}, \mathbf{y}) |u_i(\mathbf{y}) - I(\mathbf{x})|^2 \chi_i(\mathbf{x}) d\mathbf{y} d\mathbf{x}. \quad (3.3)$$

(b). Fixing \mathbf{u} , find χ such that

$$\min_{\chi} \left\{ \sum_{i=1}^m \int_{\Omega} |\nabla \chi_i(\mathbf{x})| d\mathbf{x} + \frac{\nu}{2} \iint_{\Omega \times \Omega} K(\mathbf{x}, \mathbf{y}) |u_i(\mathbf{y}) - I(\mathbf{x})|^2 \chi_i(\mathbf{x}) d\mathbf{y} d\mathbf{x} \right\}, \quad (3.4)$$

with the constraint

$$\chi_i \in \{0, 1\}, \quad 1 \leq i \leq m; \quad \sum_{i=1}^m \chi_i \equiv 1.$$

In the following, we describe the algorithms for (3.3) and (3.4).

A. Minimization With Respect to \mathbf{u}

Using calculus of variation, we can obtain the Euler-Lagrange equations of the functions $\mathbf{u} = (u_1, u_2, \dots, u_m)$ that minimize (3.3) satisfying

$$\int_{\Omega} K(\mathbf{x}, \mathbf{y}) (u_i(\mathbf{y}) - I(\mathbf{x})) \chi_i(\mathbf{x}) d\mathbf{x} = 0. \quad (3.5)$$

It admits the closed-form formula

$$u_i = \frac{\int_{\Omega} K(\mathbf{x}, \mathbf{y}) (\chi_i(\mathbf{x}) I(\mathbf{x})) d\mathbf{x}}{\int_{\Omega} K(\mathbf{x}, \mathbf{y}) \chi_i(\mathbf{x}) d\mathbf{x}}, \quad 1 \leq i \leq m. \quad (3.6)$$

The original complicated implementation of the smooth function \mathbf{u} in (1.3) is replaced with the closed-form formulation, and the implementation is easy and fast. Meanwhile, the convolution in (3.6) can be handled efficiently through the fast Fourier transform method (see, e.g., [45]).

B. Minimization With Respect to χ

Now we study the minimization problem (3.4) with respect to χ . With known u_i , (3.4) can be seen as the Potts-type model [46]. The essential methods for such models are to implement the convexification followed by the thresholding. Here we use our direct approach proposed in [30], where the variables can be explicitly expressed through the direct augmented Lagrangian method. Moreover, the iterative algorithm is based on a shrinkage in the intermediate unknowns from the augmented Lagrangian formulation. Consequently, the algorithms involve fewer number of parameters, resulting in fast and easy implementation. Below we provide detailed algorithms.

We introduce a new variable \mathbf{q}_i and rewrite (3.4) into the following constrained problem:

$$\min_{\mathbf{q}, \chi} \left\{ \sum_{i=1}^m \int_{\Omega} |\mathbf{q}_i| d\mathbf{x} + \sum_{i=1}^m \int_{\Omega} f_i \chi_i d\mathbf{x} \right\},$$

subject to $\mathbf{q}_i = \nabla \chi_i$, $\chi_i^2 = \chi_i$, $\sum_{i=1}^m \chi_i = 1$,

where for $1 \leq i \leq m$,

$$f_i(\mathbf{x}) = \frac{\nu}{2} \int_{\Omega} K(\mathbf{x}, \mathbf{y}) |u_i(\mathbf{y}) - I(\mathbf{x})|^2 d\mathbf{y}. \quad (3.7)$$

The unconstrained problem can be attained by using the augmented Lagrangian method:

$$\min_{\chi, \mathbf{q}} \max_{\vec{\lambda}_1, \vec{\lambda}_2, \lambda_3} \left\{ \mathcal{L}(\chi, \vec{\mathbf{q}}, \vec{\lambda}_1, \vec{\lambda}_2, \lambda_3) = \sum_{i=1}^m \int_{\Omega} (|\mathbf{q}_i| d\mathbf{x} + f_i \chi_i d\mathbf{x} + \lambda_{1i} (\mathbf{q}_i - \nabla \chi_i) d\mathbf{x} + \frac{r_1}{2} (\mathbf{q}_i - \nabla \chi_i)^2 d\mathbf{x} + \lambda_{2i} (\chi_i^2 - \chi_i) d\mathbf{x} + \frac{r_2}{2} (\chi_i^2 - \chi_i)^2 d\mathbf{x}) + \int_{\Omega} \lambda_3 \left(\sum_{i=1}^m \chi_i - 1 \right) d\mathbf{x} + \frac{r_3}{2} \int_{\Omega} \left(\sum_{i=1}^m \chi_i - 1 \right)^2 d\mathbf{x} \right\}, \quad (3.8)$$

where $r_1, r_2, r_3 > 0$ are penalty weights,

$$\vec{\lambda}_1 = (\lambda_{11}, \dots, \lambda_{1m}), \quad \lambda_2 = (\lambda_{21}, \dots, \lambda_{2m})$$

(with $\lambda_{2i} \geq 0$) and scalar λ_3 are Lagrange multipliers. The optimality conditions imply

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \chi_i} &= f_i + \text{div} \lambda_{1i} + r_1 \text{div}(\mathbf{q}_i - \nabla \chi_i) + \lambda_{2i} (2\chi_i - 1) \\ &\quad + r_2 (\chi_i^2 - \chi_i) (2\chi_i - 1) + \lambda_3 + r_3 \left(\sum_{i=1}^m \chi_i - 1 \right) = 0, \end{aligned} \quad (3.9a)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} = \frac{\mathbf{q}_i}{|\mathbf{q}_i|} + \lambda_{1i} + r_1 (\mathbf{q}_i - \nabla \chi_i) = 0, \quad (3.9b)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_{1i}} = \mathbf{q}_i - \nabla \chi_i = 0, \quad (3.9c)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_{2i}} = \chi_i^2 - \chi_i = 0, \quad (3.9d)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_3} = \sum_{i=1}^m \chi_i - 1 = 0. \quad (3.9e)$$

By (3.9b) and (3.9c),

$$\lambda_{1i} = -\frac{\nabla \chi_i}{|\nabla \chi_i|}, \quad 1 \leq i \leq m. \quad (3.10)$$

Following from (3.9c), (3.9d) and (3.9e), Equation (3.9a) can be simplified into

$$\lambda_{2i} (2\chi_i - 1) = -(f_i + \text{div} \lambda_{1i} + \lambda_3), \quad (3.11)$$

which implies

$$\lambda_{2i}^2 (4\chi_i^2 - 4\chi_i + 1) = (f_i + \text{div} \lambda_{1i} + \lambda_3)^2.$$

Then by (3.9d) (note: $\lambda_{2i} \geq 0$),

$$\lambda_{2i} = |f_i + \text{div} \lambda_{1i} + \lambda_3| = |h_i + \lambda_3|,$$

where $h_i := \text{div} \lambda_{1i} + f_i$. Plugging it into (3.11) leads to

$$(2\chi_i - 1) |h_i + \lambda_3| = (h_i + \lambda_3) = 0. \quad (3.12)$$

Then the expression of λ_3 can be written as

$$-\lambda_3 = h_k + \frac{h_j - h_k}{2}, \quad (3.13)$$

Algorithm 1 The implementation steps for the general model (3.2).

1. Initialization: set $\chi_i^0, \bar{\lambda}_1^0 = \mathbf{0}$ and choose $\beta, \tau, \nu, \sigma > 0$;
2. For $n = 0, 1, \dots$

(i) Compute u_i^n by

$$u_i^n = \frac{\int_{\Omega} K(\mathbf{x}, \mathbf{y}) (\chi_i^n(\mathbf{x}) I(\mathbf{x})) d\mathbf{x}}{\int_{\Omega} K(\mathbf{x}, \mathbf{y}) \chi_i^n(\mathbf{x}) d\mathbf{x}}; \quad 1 \leq i \leq m.$$

(ii) Compute λ_3 by (3.13);

(iii) Compute χ^n by

$$\chi_i^n = \frac{1}{2} - \frac{\text{div} \lambda_{1i}^n + f_i^n + \lambda_3}{2|\text{div} \lambda_{1i}^n + f_i^n + \lambda_3|_{\beta}};$$

where

$$f_i^n(\mathbf{x}) = \frac{\nu}{2} \int_{\Omega} K(\mathbf{x}, \mathbf{y}) |u_i^n(\mathbf{y}) - I(\mathbf{x})|^2 d\mathbf{y}.$$

(iv) Compute $\bar{\lambda}_1^{n+1}$ by the Chambolle-type algorithm:

$$\lambda_{1i}^{n+1} = \frac{\lambda_{1i}^n - \tau \nabla \chi_i^n}{1 + \tau |\nabla \chi_i^n|}, \quad 1 \leq i \leq m;$$

3. Endfor till some stopping rule meets;

where h_k is the smallest and h_j is the second smallest value of $\{h_i\}_{i=1}^m$ respectively.

Therefore, in the computation, the evolution of χ_i can be regularized by

$$\chi_i = \frac{1}{2} - \frac{h_i + \lambda_3}{2|h_i + \lambda_3|_{\beta}}, \quad (3.14)$$

where λ_3 is defined in (3.13), and

$$|h_i + \lambda_3|_{\beta} = |\text{div} \lambda_{1i} + f_i + \lambda_3| + \beta,$$

for sufficiently small $\beta > 0$.

Next, for fixed χ , the equation (3.10) satisfies

$$\nabla \chi_i + |\nabla \chi_i| \lambda_{1i} = \mathbf{0}, \quad 1 \leq i \leq m. \quad (3.15)$$

The iterative semi-implicit Chambolle's dual scheme [47] can be applied. With time step size τ , we can approximate λ_{1i} at $t = n\tau$. The evolution of λ_{1i} is given by

$$\begin{aligned} \frac{\lambda_{1i}^{n+1} - \lambda_{1i}^n}{\tau} &= -\nabla \chi_i^n - |\nabla \chi_i^n| \lambda_{1i}^{n+1}, \\ \Rightarrow \lambda_{1i}^{n+1} &= \frac{\lambda_{1i}^n - \tau \nabla \chi_i^n}{1 + \tau |\nabla \chi_i^n|}. \end{aligned}$$

Now, we provide the full algorithm in the top of the page.

In the computation, the selection of σ is related to the nature of the image; the general model (2.12) can handle a piecewise smooth image with intensity inhomogeneity when σ is small. On the contrary, the model performs as a piecewise constant model when σ is large.

IV. NUMERICAL RESULTS

We demonstrate the performance of the proposed general framework and algorithms by testing them on typical images

and comparing with some related methods. In the comparison experiments, we choose three typical smoothing operators: $K(\mathbf{x}, \mathbf{y}) = \{G_{\sigma}(\mathbf{x}, \mathbf{y}), G_B(\mathbf{x}, \mathbf{y}), G_D(\mathbf{x}, \mathbf{y})\}$ as the modified Gaussian model, modified bilateral model and modified directional diffusion model respectively.

A. A Comparison Study

We compare the modified models with the Vese-Chan (CV) model [7] (code downloadable at <http://www.shawnlankton.com/2007/05/active-contours/>), the global Chan-Vese (GCV) model [48] (code downloadable at personal homepage of the first author), our previous decomposition work in [24] (cf. (1.7)), the nonlocal based segmentation algorithm (code downloadable at <http://guygilboa.eew.technion.ac.il/publications/>), and the local statistical model [20] (code is downloadable at <http://www4.comp.polyu.edu.hk/~cslzhang/>). For the sake of clarity, we briefly review these models.

1) *Vese-Chan Model (CV) [7] and Global Chan-Vese Model (GCV) [48]*: These two methods use the classical level set formulations [7]. They label the phases through the combinations of the signs of the signed distance functions. Hence, n signed distance functions $\{\phi_j\}_{j=1}^n$ are used to label $m = 2^n$ phases, and the Heaviside functions $H(\phi_j)$ can be used to express the corresponding indicator functions. The minimization problems solved in [7] and [48] are the same. Here, we present two-phase optimization model for the sake of comparison:

$$\begin{aligned} \min_{u_1, u_2, \phi} \left\{ \frac{\nu}{2} \int_{\Omega_1} (u_1 - I)^2 H(\phi) d\mathbf{x} + \mu \int_{\Omega_1} |\nabla u_1|^2 d\mathbf{x} \right. \\ \left. + \frac{\nu}{2} \int_{\Omega \setminus \Omega_1} (u_2 - I)^2 (1 - H(\phi)) d\mathbf{x} \right. \\ \left. + \mu \int_{\Omega \setminus \Omega_1} |\nabla u_2|^2 d\mathbf{x} + \int_{\Omega} \delta(\phi) |\nabla \phi| d\mathbf{x} \right\}, \end{aligned}$$

where $\mu, \nu > 0$ are parameters, δ is the Dirac Delta function, and u_1, u_2 are the piecewise smooth function in Ω_1 and $\Omega \setminus \Omega_1$. The Vese-Chan model [7] and the global Chan-Vese model [48] solve for the smooth functions u_1 and u_2 through the iteration process as shown in (1.3). The difference between these two models is the evolution of ϕ . One can refer to [7] and [48] for details.

2) *Nonlocal Based Segmentation (Nonlocal, in Short) [49]*: The nonlocal regularization is introduced in [50], which is later more popular to be used in image denoising (e.g. nonlocal means [51]). Here we present the nonlocal segmentation algorithm proposed in [49]. Let I be the input image and $w(\mathbf{x}, \mathbf{y})$ a simple symmetric window. Firstly, we set the initial condition for u to be u_0 (1 inside the region and -1 outside the region). The evolution is

$$u_t = \mathcal{L}u, \quad u|_{t=0} = u_0.$$

and the linear operator of nonlocal regularization is expressed as

$$\mathcal{L}u(\mathbf{x}) = \int_{\Omega} (u(\mathbf{y}) - u(\mathbf{x})) w(\mathbf{x}, \mathbf{y}) d\mathbf{y},$$

Finally, we choose the region of object to be $u > 0$.

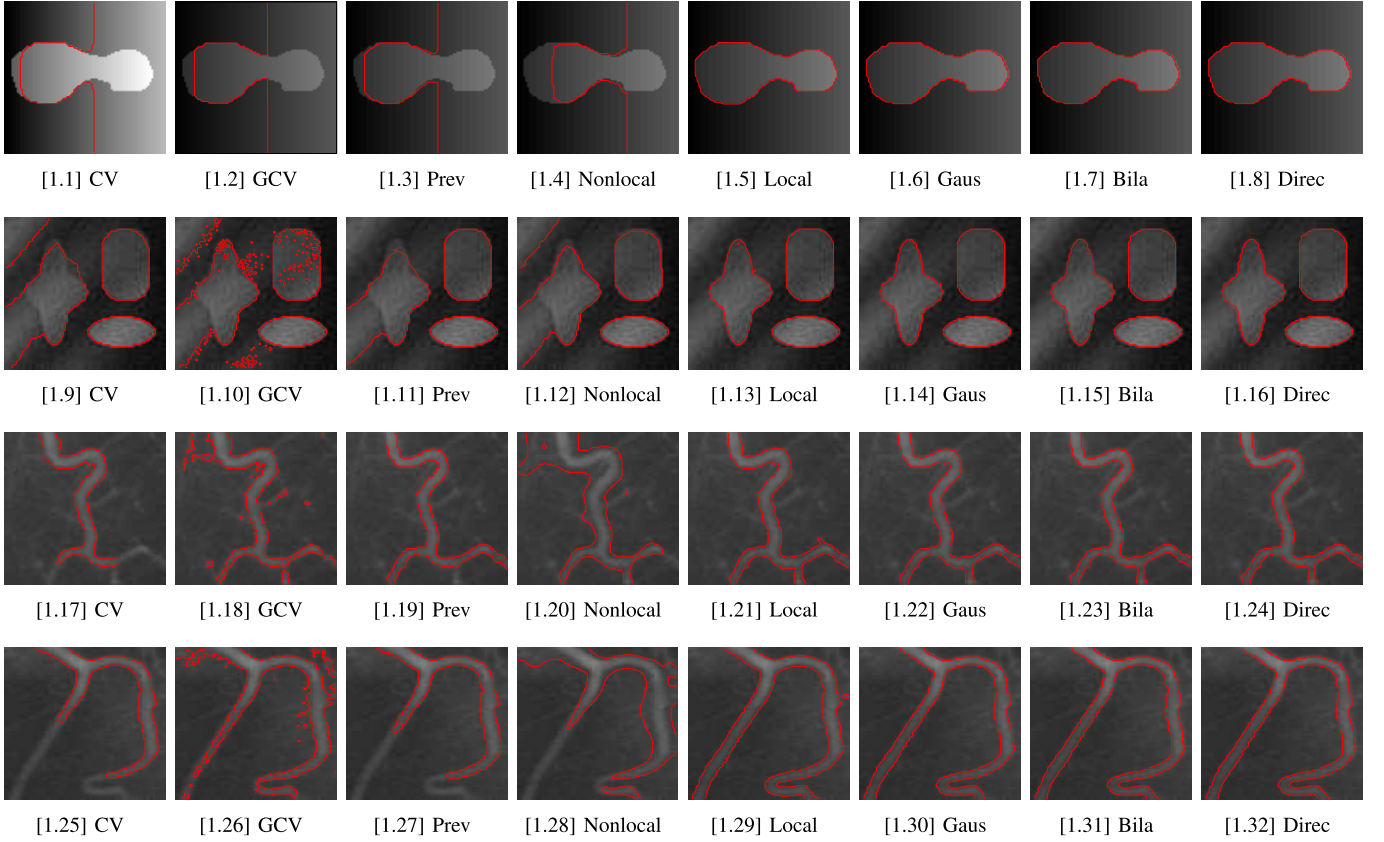


Fig. 1. Results obtained by eight different models. Column 1: segmented results of Chan-Vese model (CV). Column 2: segmented results of global Chan-Vese model (GCV). Column 3: segmented results of our previous decomposition model (Prev). Column 4: segmented results of nonlocal based segmentation method (Nonlocal). Column 5: segmented results of local statistical model (Local). Column 6: segmented results of modified Gaussian model (Gaus). Column 7: segmented results of modified bilateral model (Bila). Column 8: segmented results of modified directional diffusion model (Direc). The latter four models can produce more accurate results than the former four models.

3) *Local Statistical Method (Local, in Short)* [20]: The local statistical method proposed in [20] used the bias field (1.4), assuming the object is approximated as a combination of the biased field, the classical level set formulation, and the constant. Therefore, the two-phase minimization problem based on MS model is rewritten in the statistical form with the classical level set formulation

$$\min_{b, c_1, c_2, \phi} \left\{ \int_{\Omega} \int_{\Omega} G_{\sigma}(\mathbf{x}, \mathbf{y}) \left(\log(\sigma_1) H(\phi) + \frac{(I(\mathbf{x}) - b(\mathbf{y})c_1)^2}{2\sigma_1^2} H(\phi) + \log(\sigma_2)(1 - H(\phi)) + \frac{(I(\mathbf{x}) - b(\mathbf{y})c_2)^2}{2\sigma_2^2} (1 - H(\phi)) \right) d\mathbf{y} d\mathbf{x} \right\},$$

where σ_1, σ_2 are the parameters. G_{σ} is defined in (2.4). The evolution with respect to b, c_1, c_2, ϕ is calculated by using the gradient descent methods directly.

In the first experiment, we compare these five algorithms with our modified Gaussian model (2.12) ($K = G_{\sigma}$, denoted as Gaus), the modified bilateral model (2.12) ($K = G_B$, denoted as Bila), and the modified directional diffusion model (2.12) ($K = G_D$, denoted as Direc). All experiments were run under Matlab R2010a on a PC with Dual 3.2 GHz processor. We test in Fig. 1 two synthetic images and two real

TABLE I
COMPARISON OF COMPUTATIONAL TIME (SECONDS)
OF DIFFERENT METHODS IN FIG. 1

Methods (in Fig. 1)	Fig. 1 [1.1]-[1.8]	Fig. 1 [1.9]-[1.16]	Fig. 1 [1.17]-[1.24]	Fig. 1 [1.25]-[1.32]
CV	4.66	7.11	4.83	5.20
GCV	2.93	3.90	3.2	3.7
Previous	2.43	2.50	3.02	4.60
Non-local	2.17	2.78	3.19	3.63
Local	8.07	9.93	8.60	17.18
Gaussian	1.59	1.78	2.09	2.18
Bilateral	1.20	1.48	1.95	2.26
Directional	1.62	1.40	1.97	2.01

vessel images. These four image are PS images with intensity inhomogeneity. We tabulate in Table I the computational time (in seconds) obtained by eight algorithms, and we depict in Fig. 1 the final segmented images. We find that the former four models fail to get good segmented results, whereas the methods using the local statistical model and our new proposed models perform better. This is due to the reason that the latter four methods utilize local information in the global optimization framework. Moreover, we plot the decay of numerical energy functional (3.2) of these four images in Fig.1 with respect to the iteration for $K = G_{\sigma}, G_B$ and

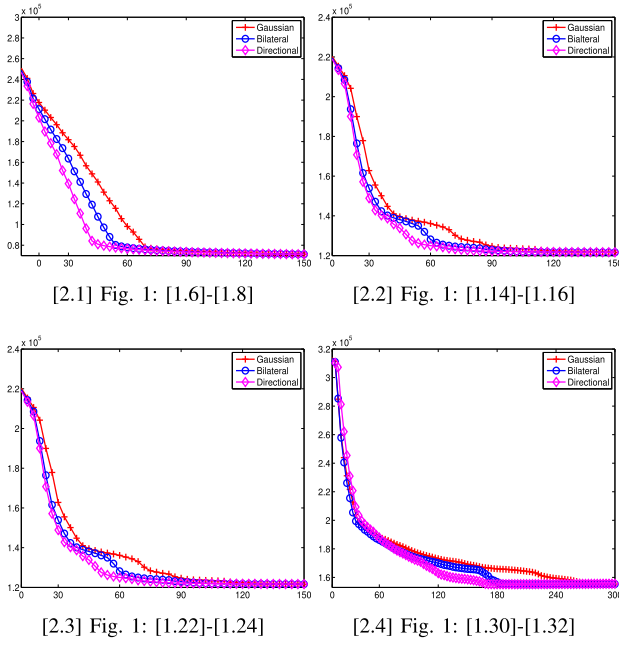


Fig. 2. Decay of the energy functional in (3.2) for three smoothing kernels and four images in Fig. 1, against the iteration steps.

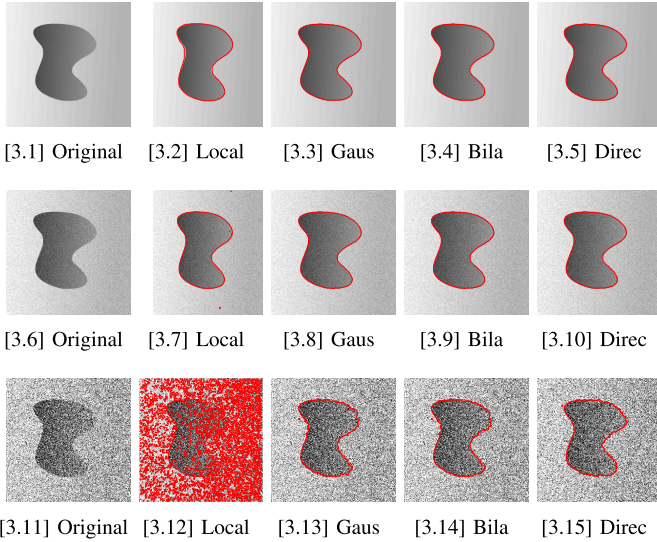


Fig. 3. Results obtained by four different models. Column 1: Images with different noise level. Column 2: segmented results of local statistical model. Column 3: segmented results of the modified Gaussian model. Column 4: segmented results of the modified bilateral model. Column 5: segmented results of the modified directional diffusion model. All of our three modified models can handle large noisy images, and are more robust to the noise.

G_D respectively in Fig. 2. The decay of energy functional can justify the convergence. We visualize that the modified directional diffusion model converges faster than the other two models.

The previous comparison showed the advantages of latter four models. Next, we compare these four models in the second experiment. We test on a synthetic image of size 85×88 with different noise levels (from low to high). In all tests, noisy images are generated by adding Gaussian noise with zero mean and different variances

TABLE II
EVALUATION RESULTS OF THE METHODS IN FIG. 3

	Fig. 3 [3.1]		Fig. 3 [3.6]		Fig. 3 [3.11]	
	Precision	Recall	Precision	Recall	Precision	Recall
Local	1	0.9911	0.9991	0.9991	0.9479	0.7234
Gaussian	0.9987	1	0.9987	0.9998	0.9902	0.9970
Bilateral	0.9988	1	0.9982	0.9999	0.9891	0.9973
Directional	0.9989	1	0.9981	0.9999	0.9939	0.9974

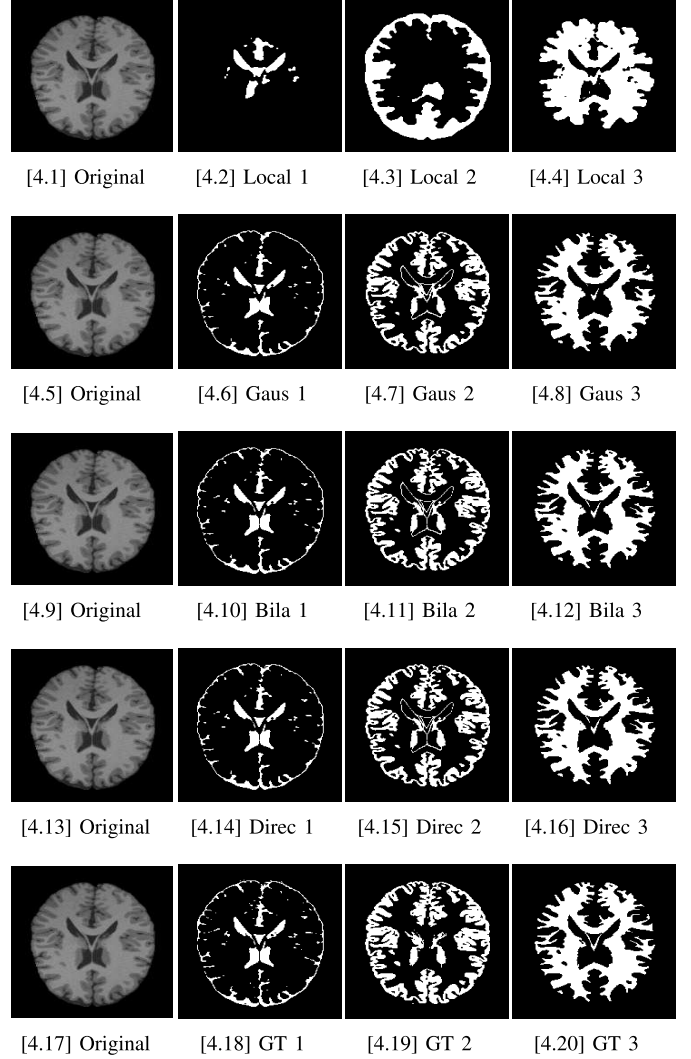


Fig. 4. Results obtained by local statistical model, our modified Gaussian model, bilateral model and directional diffusion model. Row 1: segmented results of local statistical model. Row 2: segmented results of modified Gaussian model. Row 3: segmented results of modified bilateral model. Row 4: segmented results of modified directional diffusion model. Row 5: ground truth. Our proposed three models produce resembling results as the ground truth, and all three phases match the ground truth better than local statistical model.

$d = 0.002, 0.02, 0.2$. In Fig. 3, we present the noisy images and segmentation results of four different models. We observe from Fig. 3: [3.1]-[3.5] that the four models perform similarly for low noise images. However, the modified Gaussian model, bilateral model and directional diffusion model yield more accurate results than the local statistical model for higher noise-level images. We also provide the comparison results in

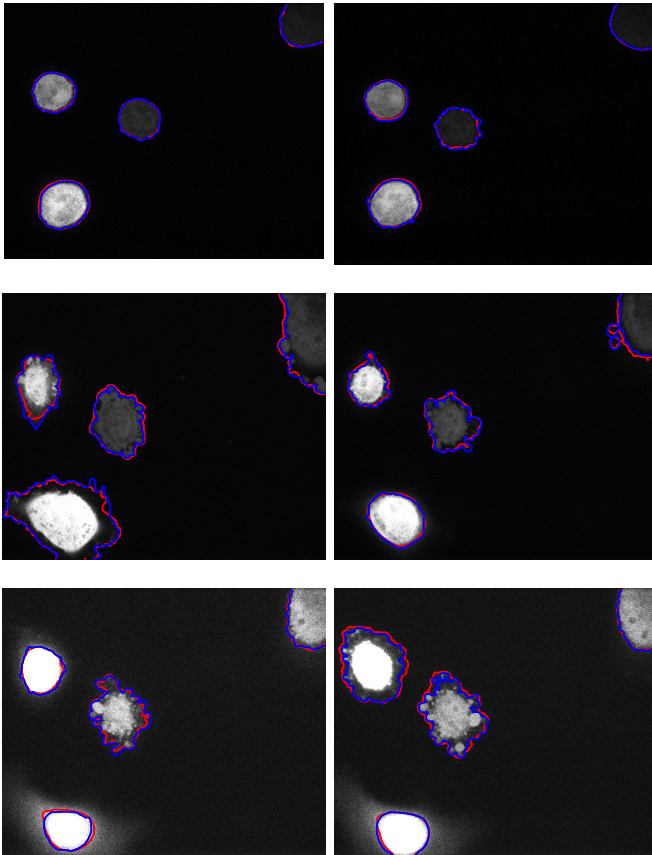


Fig. 5. Segmented results (in red) on cell dataset C3DH-H157 by using the modified directional diffusion model and the ground truth (in blue). Our modified model can obtain satisfactory results.

terms of precision and recall in Table II. Precision and recall are the basic measures used in evaluating results. Precision is the fraction of the true positive pixels in the segmented results, while recall (or sensitivity) is the fraction of the true positive pixels in the ground truth of the foreground. We also observe from the evaluation results in Table II that the modified directional diffusion model and the modified bilateral model produce slightly better results than the modified Gaussian model. Therefore, the three proposed modified MS models are more robust to noise.

In medicine area, many research applications using magnetic resonance (MR) images involve segmenting PS images with intensity inhomogeneity. Hence in the third experiment, we test on MRI brain images. This dataset comes from <http://www.bic.mni.mcgill.ca/brainweb/>, and we choose a noise level of 7% and non-uniformity (“RF”) of 20%. There are three classes that should be segmented: cerebrospinal fluid (CF), gray matter (GM) and white matter (WM). Fig. 4 provides the segmented results of the local statistical model, the modified Gaussian model, the modified bilateral model, the modified directional diffusion model and the ground truth (GT). We also compare the evaluation results in terms of precision and recall for each phase in Table III. We find that our three proposed modified models can produce much better results than local statistical model.

TABLE III
EVALUATION RESULTS OF THE METHODS IN FIG. 4

	CF		GM		WM	
	Precision	Recall	Precision	Recall	Precision	Recall
Local	0.7080	0.3268	0.6680	0.5990	0.7800	0.9142
Gaussian	0.9306	0.8598	0.9503	0.8519	0.9621	0.9710
Bilateral	0.9698	0.9167	0.9622	0.8663	0.9742	0.9822
Directional	0.9736	0.9109	0.9636	0.8645	0.9720	0.9853

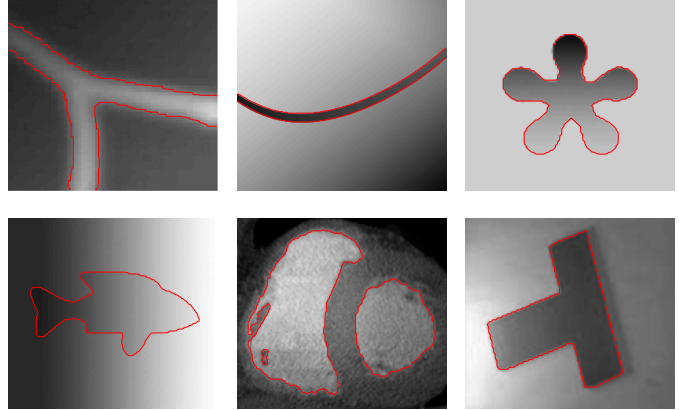


Fig. 6. More tests on images segmentation with intensity inhomogeneities using the modified directional diffusion model. Our proposed model can achieve accurate results for the image with low contrast.

In the forth experiment, we perform experiments on human squamous lung carcinoma cell images. This dataset comes from C3DH-H157 [52]. In Fig. 5, we present the segmented results (in red) by using the modified directional diffusion model and the ground truth (in blue). Meanwhile, we compute the average precision and the average recall for the cells detected. The average value is 0.9340 and 0.9617, respectively. We find that our proposed models can detect the cells well.

Finally, we test some typical images used as examples in the related references. We present the segmented results by using the modified directional diffusion model in Fig. 6. We see that in all tests, the modified directional diffusion model can provide satisfactory segmentation results.

V. CONCLUSION

In this work, we proposed a general framework to modify the MS model for multiphase PS image segmentation, where the essential idea was to use the smoothing operators to replace the regularisation term. Some insights and analysis were provided to justify the modification and to understand the general principles for selecting smooth kernels. Using the closed-form formula for the targeted smooth approximation and the direct augmented Lagrangian method, we were able to come up with a fast minimisation algorithm for the general model. Through ample numerical results and comparison with other relevant methods, we demonstrated the efficiency and accuracy of our proposed model and algorithm.

REFERENCES

- [1] M. K. Agoston, *Computer Graphics and Geometric Modelling: Implementation and Algorithms*. Berlin, Germany: Springer, 2005.
- [2] N. Paragios, Y. Chen, and O. Faugeras, *Handbook of Mathematical Models in Computer Vision*. Berlin, Germany: Springer, 2006.

- [3] J. Morel and S. Solimini, *Variational Methods in Image Segmentation*. Basel, Switzerland: Birkhäuser, 1994.
- [4] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.
- [5] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
- [6] A. Tsai, A. Yezzi, and A. Willsky, "Curve evolution implementation of the Mumford–Shah functional for image segmentation, denoising, interpolation, and magnification," *IEEE Trans. Image Process.*, vol. 10, no. 8, pp. 1169–1186, Aug. 2001.
- [7] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 271–293, Dec. 2002. [Online]. Available: <http://dx.doi.org/10.1023/A:1020874308076>
- [8] J. Piovanio, M. Rousson, and T. Papadopoulos, "Efficient segmentation of piecewise smooth images," in *Proc. Scale Space Variational Methods Comput. Vis.*, Jun. 2007, pp. 709–720.
- [9] Y. Zhang, "Fast segmentation for the piecewise smooth Mumford–Shah functional," *Int. J. Signal Process.*, vol. 2, no. 4, pp. 1140–1145, 2006.
- [10] N. Paragios and R. Deriche, "Geodesic active contours and level sets for detection and tracking of moving objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 3, pp. 266–280, Mar. 2000.
- [11] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.
- [12] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. New York, NY, USA: Springer, 2006.
- [13] J. Morel and S. Solimini, *Variational Methods in Image Segmentation: With Seven Image Processing Experiments*. Cambridge, MA, USA: Birkhäuser, 1995.
- [14] A. Chambolle, "Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations," *SIAM J. Appl. Math.*, vol. 55, no. 3, pp. 827–863, 1995.
- [15] J. Lie, M. Lysaker, and X.-C. Tai, "A binary level set model and some applications to Mumford–Shah image segmentation," *IEEE Trans. Image Process.*, vol. 15, no. 5, pp. 1171–1181, May 2006.
- [16] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi formulations," *J. Comput. Phys.*, vol. 79, no. 1, pp. 12–49, 1988.
- [17] T. Brox and D. Cremers, "On the statistical interpretation of the piecewise smooth Mumford–Shah functional," in *Proc. SSVM*, 2007, pp. 203–213.
- [18] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, Oct. 2008.
- [19] Y. Gu, W. Xiong, L.-L. Wang, J. Cheng, W. Huang, and J. Zhou, "A new approach for multiphase piecewise smooth image segmentation," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 4417–4421.
- [20] K. Zhang, L. Zhang, K. M. Lam, and D. Zhang, "A level set approach to image segmentation with intensity inhomogeneity," *IEEE Trans. Cybernetics*, vol. 46, no. 2, pp. 546–557, 2016.
- [21] H. Ali, N. Badshah, K. Chen, and G. Khan, "A variational model with hybrid images data fitting energies for segmentation of images with intensity inhomogeneity," *Pattern Recognit.*, vol. 51, pp. 27–42, Mar. 2016.
- [22] C. Li, R. Huang, Z. Ding, C. Gatenby, and D. Metaxas, "A variational level set approach to segmentation and bias correction of images with intensity inhomogeneity," in *Medical Image Computing and Computer-Assisted Intervention MICCAI* (Series Lecture Notes in Computer Science). Berlin, Germany: Springer, 2008, pp. 1083–1091.
- [23] F. Li, M. Ng, and C. Li, "Variational fuzzy Mumford–Shah model for image segmentation," *SIAM J. Appl. Math.*, vol. 70, no. 7, pp. 2750–2770, 2010.
- [24] Y. Gu, L.-L. Wang, W. Xiong, J. Cheng, W. Huang, and J. Zhou, "Efficient and robust image segmentation with a new piecewise-smooth decomposition model," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 2718–2722.
- [25] A. Jung, "Piecewise-smooth image segmentation models with L^1 data-fidelity terms," *J. Sci. Comput.*, pp. 1–33, Aug. 2016. doi: 10.1007/s10915-016-0280-z.
- [26] J. Spencer and K. Chen, "Stabilised bias field: Segmentation with intensity inhomogeneity," *J. Algorithms Comput. Technol.*, vol. 10, no. 4, pp. 302–313, 2016.
- [27] W. M. Wells, III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 15, no. 8, pp. 429–442, Aug. 1996.
- [28] D. Pham and J. Prince, "An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities," *Pattern Recognit. Lett.*, vol. 20, no. 1, pp. 57–68, 1998.
- [29] Y. Gu *et al.*, "A new Mumford–Shah type model involving a smoothing operator for multiphase image segmentation," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 1990–1994.
- [30] Y. Gu, L. Wang, and X. Tai, "A direct approach toward global minimization for multiphase labeling and segmentation problems," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2399–2411, May 2012.
- [31] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2029–2039, Nov. 2008.
- [32] B. B. Kimia and K. Siddiqi, "Geometric heat equation and nonlinear diffusion of shapes and images," *Comput. Vis. Image Understand.*, vol. 64, no. 3, pp. 305–322, 1996.
- [33] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
- [34] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.
- [35] J. Weickert, *Anisotropic Diffusion in Image Processing*. Stuttgart, Germany: Teubner, 1998.
- [36] D. Barash, "Fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 6, pp. 844–847, Jun. 2002.
- [37] A. Buades, B. Coll, and J. Morel, "The staircasing effect in neighborhood filters and its solution," *IEEE Trans. Image Process.*, vol. 6, no. 15, pp. 1499–1505, Jun. 2006.
- [38] J. Wang, "Construction of local nonlinear filter without staircase effect in image restoration," *Int. J. Applicable Anal.*, vol. 90, no. 8, pp. 1257–1273, 2010.
- [39] J. Wang, "Approximation of weighted local mean operators," *Int. J. Applicable Anal.*, vol. 93, no. 4, pp. 823–839, 2014.
- [40] L. Yaroslavsky, *Digital Picture Processing—An Introduction*. Berlin, Germany: Springer, 1985.
- [41] L. Alvarez, P. Lions, and J. Morel, "Image selective smoothing and edge detection by nonlinear diffusion. II," *SIAM J. Numer. Anal.*, vol. 29, no. 3, pp. 845–866, 1992.
- [42] R. Whitaker and S. Pizer, "A multi-scale approach to nonuniform diffusion," *CVGIP, Image Understand.*, vol. 57, no. 1, pp. 99–110, 1993.
- [43] F. Catté, P. Lions, J. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM J. Numer. Anal.*, vol. 29, no. 1, pp. 182–193, 1992.
- [44] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*. Berlin, Germany: Springer, 2002.
- [45] J. Shen, T. Tang, and L. Wang, *Spectral Methods: Algorithms, Analysis and Applications*. Berlin, Germany: Springer, 2011.
- [46] R. Potts, "Some generalized order-disorder transformations," in *Proc. Cambridge Philos. Soc.*, 1952, vol. 48, no. 1, pp. 106–109.
- [47] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imag. Vis.*, vol. 20, no. 1, pp. 89–97, 2004.
- [48] X. Bresson, S. Esedoglu, P. Vanderghynst, J. Thiran, and S. Osher, "Fast global minimization of the active contour/snake model," *J. Math. Imag. Vis.*, vol. 28, no. 2, pp. 151–167, 2007.
- [49] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Model. Simul.*, vol. 7, no. 3, pp. 1005–1028, 2008.
- [50] G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation," *Multiscale Model. Simul.*, vol. 6, no. 2, pp. 595–630, 2007.
- [51] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
- [52] *ISBI: The First Cell Tracking Challenge*. accessed on 2013. [Online]. Available: <http://www.codesolorzano.com/celltrackingchallenge>