A LOCAL NONPARAMETRIC MODEL FOR SIMULTANEOUS IMAGE SEGMENTATION AND ADAPTIVE SMOOTHING

YUNMEI CHEN * AND WEIHONG GUO†

Abstract.
Parametric and nonparametric region based active contour models have been widely used in image segmentation and showed promising results. However, since segmentation processing in these models are driven by intensity probability density functions (p.d.f.), intensity inhomogeneity and higher level of noise are always challenging problems that need to be addressed. In this paper we present a novel local nonparametric model for simultaneous image segmentation and adaptive smoothing. We treat the smoothed image intensity at each point as a random variable, whose realizations are the intensities of the observed noisy image in a neighborhood of this point. The neighborhood size varies from point to point depending on image gradients. A nonparametric p.d.f. estimation is applied to the smoothed image to get likelihood estimations for both object and background. Then, the simultaneous smoothing and segmentation is achieved by minimizing the negative log-likelihood estimations together with total length of the region boundaries. By the choice of the local adaptive neighborhoods the smoothing does not across the boundaries and is less at the locations where image gradient is higher.

The proposed model is implemented using its level set formulation. The experimental results on synthetic data, T1 weighted human brain MRI images, FLAIR MRI brain images, and echocardiographic images indicate the advantages of the proposed model in dealing with higher level noise and intensity inhomogeneity. The existence of a solution to the proposed model is also discussed.

Key words. Segmentation, smoothing, nonparametric probability density function, level set method, active contour

AMS subject classifications. 94A08, 68U10, 15A29

1. Introduction. Parametric and nonparametric region based active contour models have been widely used in image segmentation and showed promising results. These models incorporate region intensity statistics into segmentation to overcome the difficulty in edge based segmentation models that require strong edge information. However, since the segmentation process in these models is driven by intensity probability density functions, the difficulties of the task caused by intensity inhomogeneity and higher level of noise still need to be addressed. This is the motivation of this work.

Let Ω be a bounded open subset of R^N, and $I_0 : \Omega \to R$ be an observed noisy image. The celebrated work from Mumford and Shah [1] provides the following model for simultaneous smoothing and segmentation:

$$\min_{I,C} \sum_i \int_{\Omega_i} (I_i - I_0)^2 dX + \alpha \int_{\Omega \setminus C} |\nabla I|^2 dX + \beta |C| \tag{1.1}$$

where C separates Ω into two regions Ω_i ($i = 1, 2$), $|C|$ represents the length of C, I is a piecewise smooth approximation of I_0, and I_i is the restriction of I to Ω_i. When I is a constant C_i in each Ω_i, model (1.1) reduces to the following form.

$$\min_{C_i's, C} \sum_i \int_{\Omega_i} (C_i - I_0)^2 dX + \beta |C| \tag{1.2}$$

This piecewise constant Mumford-Shah (MS) model has been well studied by Chan et al. in [2, 3]. The major advantage of this model is that it can separate two relatively homogeneous

*358 Little Hall, P. O. Box 118105, Department of Mathematics, University of Florida, Gainesville, FL, 32611, yun@math.ufl.edu

†Department of Mathematics, University of Alabama, P. O. Box 870350, Tuscaloosa, AL, 35487, wguo@as.ua.edu

regions without using edge information. However, the homogeneity assumption limits its applications.

One of the more general approaches is parametric region based active contour method. This method is based on the assumption that the image intensity $I_0(X)$ at each $X \in \Omega_i$ is an independent random variable indexed by X with p.d.f. $P(I_{0_i}(X)|\lambda_i)$ given a parameter vector λ_i, where I_{0_i} is the restriction of I_0 to Ω_i. The framework of this method minimizes the negative log-likelihood functions together with the length of the boundary, i.e.

$$\min_{C,\lambda_i's} -\sum_i \int_{\Omega_i} \log P(I_{0_i}(X)|\lambda_i)dX + \beta|C| \tag{1.3}$$

In the region competition model by Zhu et al. [4] and geodesic active region models by Rousson et al. [5] and Paragios et al. [6], $P(I_{0_i}(X)|\lambda_i)$ is chosen to be a Gaussian distribution:

$$P(I_{0_i}(X)|C_i,\sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(\frac{-(I_{0_i}(X)-C_i)^2}{2\sigma_i^2}\right) \tag{1.4}$$

When all the σ_i's are the same and pre-fixed, model (1.3-1.4) reduces to model (1.2). We call model (1.3-1.4) a global Gaussian model, since it assumes that all random variables $I_{0_i}(X)$ indexed by $X \in \Omega_i$ are normal random variables sharing the same mean C_i and variance σ_i. Parametric model (1.3) provides desirable segmentation results, when the parametric form of the intensity distribution is known. However, a specific assumption of the intensity distribution can be a significant restriction in real applications. Especially, when the image has high level of noise and/or complex multi-modal intensity distribution.

To overcome this problem nonparametric models [7] have been developed and successfully applied to image segmentation and registration. They are featured by using nonparametric density estimation to replace the parametric density estimation. For instance, the active contour for segmentation in [8] is driven by a force determined by the disparity of the p.d.f. 's of the object and background, where the p.d.f.'s are approximated by Parzen window density estimation. Similarly, in [9] the segmention of regions of interest in video sequence is obtained by minimizing the disparity of the p.d.f.'s of the current frame with the previous one, where the p.d.f. 's are also estimated using Parzen window density estimation. Also, in the variational segmentation model in [10] the energy functional incorporates edge information from the edge map image, and interior texture/intensity information captured using Parzen density estimation. In [11] Kim et al. segment images through maximizing mutual information between the region labels and image intensities, where the Parzen density estimation is applied to compute their joint p.d.f.. In all these work the p.d.f.'s are estimated using information from the whole domain Ω or segmented regions Ω_i. The implicit underneath assumption is that intensities $I_0(X)$ for all $X \in \Omega$ or $X \in \Omega_i$ are drawn as samples from the one unique distribution. This assumption may not be appropriate in some situations.

An interesting generalization is to make the density estimation local, so that it takes into account the nonstationary relation among intensities across points. In [12] Hermosillo et al. propose a local nonparametric density estimation to compute mutual information. The local density estimation is achieved by weighting the Parzen density estimate in the whole domain with a normalized Gaussian of a suitable variance. In a recent work of Chan et al. [13], a neighborhood histogram based segmentation model is proposed. This model compares the object (resp. background) histogram with all the neighborhood histograms in the object (resp. background) using the Wasserstein distance. The neighborhood size at each point is chosen to be the same. Moreover, recently, Azzabou et al. propose an adaptive denoising model [14] that estimates the intensity of smoothed image at a given pixel by maximizing its conditional probability relative to the observed noisy image. In their model the joint distribution of the

smoothed image and noisy image is estimated by applying nonparametric density estimation using multidimensional isotropic Gaussian kernel with zero mean and adaptive bandwidths. At each point the set of pairs of the smoothed image and noisy image used as "samples" to perform density estimations is extracted from the local neighborhood of this point. The neighborhood size does not change across points.

Nonparametric models do not require prior knowledge on intensity distribution, they are more applicable. However, as other data driven models, they also suffer from intensity inhomogeneity and higher level of noise. To address this problem in this paper we propose a novel local nonparametric model which is able to perform image segmentation and adaptive smoothing simultaneously. Our main idea is to treat the smoothed image intensity $I(X)$ at each point X as a random variable, whose realizations are the intensities of the observed noisy image I_0 in the neighborhood of X. Then, a nonparametric density estimation based on those realizations is applied to estimate the p.d.f. of $I(X)$, from which the likelihood for the object and background are obtained. Furthermore, by minimizing the negative log-likelihood functions together with the length of the boundary, a partition of $\Omega = \bigcup_{i=1}^{2} \Omega_i$ is obtained based on the regional statistics of the smoothed image, which is more reliable than the information from the noisy image. This is different from the nonparametric models mentioned above that are used purely for image segmentation, where the nonparametric p.d.f. estimations are applied to the observed image not the recovered image.

First variation of the energy functional with respect to the smoothed image shows that $I(X)$ at each X is a average of I_0 in the neighborhood, it is generated and updated simultaneously as the partition does. The processes of smoothing and partition assist each other to get a desired result. The neighborhood of X, where the samples of $I(X)$ is extracted, is taken as a ball center at X with radius $r(X)$ within Ω_i, i.e. $\Omega_i \cap B(X, r(X))$. By this setting the optimal $I(X)$ at each X is in fact the average of I_0 in the neighborhood $\Omega_i \cap B(X, r(X))$ (see (3.3-3.4) below). In this sense we call I a smooth version of I_0. Moreover, different from the local nonparametric models in [13] and [14] where the local neighborhood sizes at each point are the same, our local neighborhood size, i.e. the radius of the ball, varies from point to point depending on image gradients. This ensures that smoothing is much less at the locations where image gradients are larger so that features can be better preserved and segmentation is more accurate. Finally, to reduce the computation cost in nonparametric density estimation, we replace the widely used Gaussian kernel by a quadratic kernel.

The organization of the paper is as follows: in section 2 we first present the proposed model, then provide details about choice of kernel functions and adaptive neighborhood size. Section 3 discusses some numerical implementation issues. Section 4 focuses on validation and applications of the proposed model. Existence of a solution to the proposed model is presented in section 5. Conclusion and future work are drawn in the last section.

2. Proposed Model.

2.1. Energy functional. Given a noisy image I_0 our task is to find an image I smoother than I_0 and a curve C separating the object Ω_1 from background Ω_2 simultaneously. In the light of region based active contour models for segmentation one way to tackle the segmentation problem for images with non-uniform intensity or higher level noise might be using the framework (1.3) with the parametric density estimation replaced by a nonparametric density estimation (see [15, 16] for details), i.e.

$$\min_{C} -\sum_{i} \int_{\Omega_i} \log P(I_{0_i}(X)) dX + \beta |C|, \tag{2.1}$$

where I_{0_i} is the restriction of I_0 on domain Ω_i, $P(I_{0_i}(X)) = \frac{1}{|\Omega_i|}\int_{\Omega_i} K_h(I_{0_i}(X)-I_{0_i}(Y))dY$, $K_h = \frac{K(\cdot/h)}{h}$, $K(\cdot)$is a kernel function, and h is the bandwidth. The most common kernel functions are unimodal, symmetric about the origin, and fall off quickly to zero. When K is a Gaussian with zero mean and σ variance, this estimator is the popular Parzen window density estimator. In model (2.1) the segmentation is driven by the regional intensity p.d.f. estimation, which may not be trustable if the image I_0 is very noisy. To address this problem we propose to incorporate a feature preserving denosing process into segmentation process, so that the segmentation can be based on the statistical information of the smoothed image. Our approach will be using the same framework as in (2.1), but the density estimation is applied to the smoothed image rather than the observed noisy image. That is minimizing the following functional:

$$\min_C -\sum_i \int_{\Omega_i} \log P(I_i(X))dX + \beta|C|, \tag{2.2}$$

where $P(I_i(X))$ is p.d.f. of the smoothed image intensity at X. To get a simultaneous smoothing and segmentation without adding extra smoothing terms we relate the smoothed image I and observed noisy image I_0 in the following way: treat $I_i(X)$, the restriction of I on Ω_i, at each $X \in \Omega_i$ as a random variable, whose samples are extracted from I_0 at a local neighborhood of X. The local neighborhood is a ball $B(X, r(X))$ centered at X with radius $r(X)$ within Ω_i. This means that random variable $I_i(X)$ has realizations $\{I_0(Y)|Y \in B(X, r(X)) \cap \Omega_i\}$. By a Kernel based nonparametric density estimation we have

$$P(I_i(X)) = \frac{1}{|B(X, r(X)) \cap \Omega_i|}\int_{B(X,r(X))\cap\Omega_i} K_h(I_i(X) - I_0(Y))dY \tag{2.3}$$

Substitute this to (2.2), our proposed model reads

$$\min_{I_1(X),I_2(X),C} E(I_1(X), I_2(X), C),$$

where $E(I_1(X), I_2(X), C)$ is defined as follows:

$$-\sum_i \int_{\Omega_i} \log\left(\frac{1}{|B(x, r(x)) \cap \Omega_i|}\int_{B(x,r(x))\cap\Omega_i} K_h(I_i(x) - I_0(y))dy\right)dx + \beta|C| \tag{2.4}$$

The first part in the energy functional is the negative log likelihood. Through maximizing the likelihood with the kernel described in next subsection, $I_i(X)$ will be forced to be the average of I_0 in $B(X, r(X)) \cap \Omega_i$. Since every neighborhood of $X \in \Omega$ is inside Ω_i, the averaging never crosses boundary. Moreover, the radius of the ball $r(X)$ varies from point to point depending on image gradients (see subsection below). The smoothing, which means averaging in this model, is adaptive and preserves features. Moreover, by maximizing the likelihood term we also get a segmentation based on the statistical information of I. The second part represents the length of the contour, and contributes to the smoothness of the boundary.

2.2. Kernel. In model (2.4) the kernel function $K_h(\cdot)$ is chosen as a quadratic function below.

$$Q_h(u) = \begin{cases} \frac{3(1-(u/h)^2)}{4h} & |u| \le h \\ 0 & |u| \ge h \end{cases} \tag{2.5}$$

with fixed bandwidth $h = 1$. FIG.2.1 illustrates the graphs of a gaussian kernel with zero mean and 0.6 variance restricted to interval $[-1.5, 1.5]$ and Q_1 defined above. The purpose of using a quadratic kernel instead of a Gaussian kernel is to reduce computational cost. Computation involving exponential function is much more than a quadratic function.

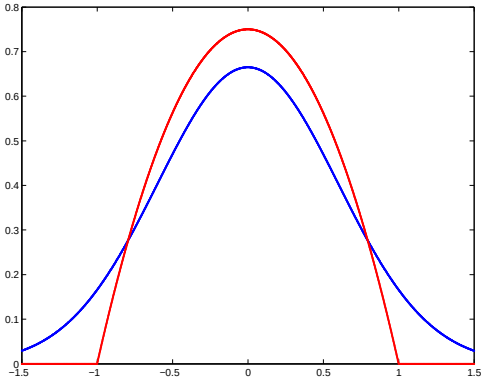


FIG. 2.1. *Graphs of a gaussian kernel with 0 mean , .4 variance (blue) and a quadratic kernel (2.5) with* $h = 1$ *(red).*

In general, the bandwidth h is an important parameter in nonparametric p.d.f. estimation, it determines how we trust each sample, also affects how many samples are incorporated into estimation. We choose $h = 1$ because of the following consideration. Practically when we apply the proposed model, we first scale the intensity of the observed image I_0 to be between 0 and .9. By the Euler Lagrangian equations (EL) of (2.4) (see (3.3-3.4) below), $I_i(X)$ must be the mean of I_0 in the neighborhood $\Omega_i \cap B(X, r(X))$, hence $|I_i(X) - I_0(Y|$ can't be beyond one for all $Y \in \Omega_i \cap B(X, r(X))$. This means that if the bandwidth $h = 1$, I_0 at each point of this neighborhood is picked as a sample contributing to the estimate of the p.d.f. of $I_i(X)$. Then, the proposed local density estimator is built by varying the neighborhood size. In this way an averaging neighborhood intensity process, regarding as smoothing, is performed simultaneously with segmentation.

2.3. Adaptive radius. To have an adaptive smoothing for feature preserving, the radius $r(X)$ of the ball is chosen depending on image gradients. More precisely, one of the choices is the following.

$$r(X) = \frac{a}{1 + b|\nabla \widetilde{I}_0(X)|} \tag{2.6}$$

where a and b are positive numbers, $\widetilde{I}_0$ is a smoothed version of the given image I_0 that could be obtained through convolving I_0 with a smooth kernel. By this choice, if X is near the edge of I_0, where the gradient is larger, the averaging to generate I is restricted in a smaller neighborhood of this point, hence, feature can be preserved. For those points, where $|\nabla \widetilde{I}_0(X)|$ is smaller, the averaging is carried out in a larger neighborhood to speed up the smoothing. Our experimental results below will verify this advantage.

If the noise distributions of the observed image in the object and background are not too inhomogeneous, we may choose $r(X)$ as below.

$$r(x) = \begin{cases} M, & |\nabla \widetilde{I}_0(x)| \le s \\ N, & |\nabla \widetilde{I}_0(x)| > s \end{cases} \tag{2.7}$$

where M is an integer depending on noise level of the input image I_0. It is chosen to be relatively larger, if the noise level is higher. N should be smaller than M. For most of the images we have tried, s is set to be the 85th percentile of $|\nabla \widetilde{I}_0(x)|$. Experimental results are not very sensitive to these parameters; the general rule to adjust N and s is to set them to be smaller if fine structures in the noisy image need to be persevered.

One more option is to replace I_0 in the above two choices by the smoothed image I, i.e.

$$r(x) = \frac{a}{1 + b|\nabla I(x)|} \tag{2.8}$$

or

$$r(x) = \begin{cases} M & |\nabla \widetilde{I}(x)| \le s \\ N & |\nabla \widetilde{I}(x)| > s \end{cases} \tag{2.9}$$

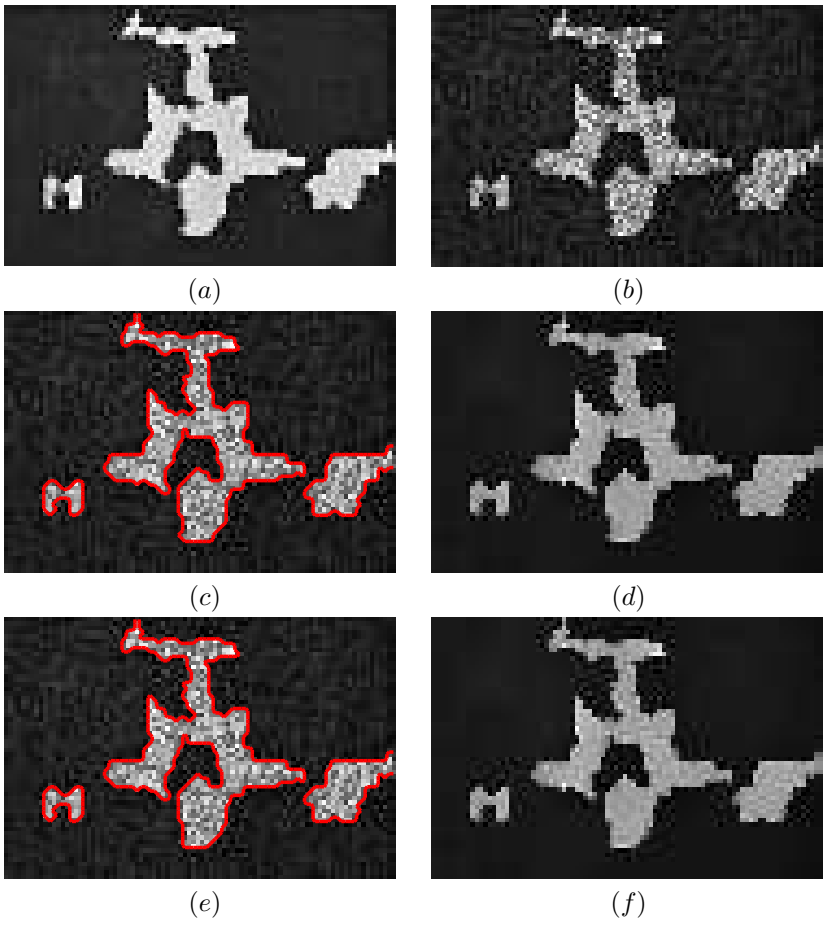


FIG. 2.2. *Compare segmentation and smoothing results of model(2.10) with $r(x)$ as defined in (2.7) and (2.9). (a) A clean plane image (60×90); (b) a noisy plane image obtained by adding speckle noise with parameter .05 to (a); (c)-(d): segmentation and smoothing results of (b) using model(2.10) with radius defined in (2.7) respectively; (e)-(f) segmentation and smoothing results of (b) using model(2.10) with radius defined in (2.9) respectively. For all the results, time step size = .1, $\beta = .5$, M = 5, N=1, converge in 5 iterations.*

To see if $r(X)$ depending on I rather than I_0 can significantly improve the segmentation and smoothing results we did an experiment shown in FIG.2.2. FIG.2.2(a) shows a clean plane image. A speckle noise with parameter .05 is added to generate a noisy image shown in (b). The second and the third row demonstrates the segmentation and smoothing results from model (2.4) with radius defined in (2.7) and (2.9) respectively. There is no significant difference observed in both segmentation and smoothing results. Therefore, we take $r(x)$ depending only on I_0 in our experiments. Also, if $r(X)$ depends on I, the model will be more nonlinear and the discussion on existence of a model solution will be more difficult.

2.4. Level set formulation. Level set method was invented and continuously advanced by S. Osher and J. Sethian in 1988 [17]. The main advantage of level set method is that topological changes, such as merging and pinching off of contours can be captured naturally, and all the numerical computation is on fixed grid.

For notation simplicity, we denote the object Ω_1 by A while the background Ω_2 by $\Omega\backslash A$. We represent the curve C by the zero level set of a Lipschitz function ϕ, i.e. $C = \{x|\phi(x) = 0\}$. The object A and background $\Omega\backslash A$ are represented by $\{x|\phi(x) > 0\}$ and $\{x|\phi(x) < 0\}$, respectively. A Heaviside function defined as $H(\phi) = 1$ when $\phi > 0$ and 0 elsewhere is used to distinguish A and $\Omega\backslash A$. Thus, set A corresponds to the region where $H(\phi) = 1$, while $\Omega\backslash A$ corresponds to the region where $H(\phi) = 0$. Then, the level set formulation of the proposed model (2.4) is as follows:

$$\begin{aligned}\min_{I_1(X),I_2(X),\phi(X)} E(I_1(X), I_2(X), \phi(X)) &= \beta \int_\Omega |\nabla H(\phi)| \\ - \int_\Omega H(\phi(X)) \log \left(\frac{\int_\Omega \chi_{B(X,r(X))}(Y) H(\phi(Y)) Q_1(I_1(X) - I_0(Y)) dY}{\int_\Omega \chi_{B(X,r(X))}(y) H(\phi(Y)) dY} \right) dx \\ - \int_\Omega (1 - H(\phi(X))) \log \left(\frac{\int_\Omega \chi_{B(X,r(X))}(Y)(1 - H(\phi(Y))) Q_1(I_1(X) - I_0(Y)) dy}{\int_\Omega \chi_{B(X,r(X))}(Y)(1 - H(\phi(Y))) dY} \right) dX\end{aligned} \tag{2.10}$$

where $\chi_{B(X,r(X))}(Y)$ is the characteristic function of $B(X, r(X))$.

3. Numerical Implementation . The level set method is used to represent C implicitly and to express each subregion.

We implement the proposed model using its level set formulation (2.10). We first derive the gradient flows then apply evolution techniques.

3.1. Certain calculation details. In our implementation we use C^∞ regularized version of Heaviside function as defined in [2] to make the functional differentiable with respect to ϕ. The regularized Heaviside function is denoted by $\widetilde{H}$, the derivative of which is called regularized Delta function and is denoted by $\widetilde{\delta}$. Moreover, a linear transformation, that does not affect segmentation results, is applied to make the observed image I_0 having range $[0, .9]$. Because $0 \leq I_0 \leq .9$ and the bandwidth of the kernel is one (see definition of Q_1) all $I_0(Y)$'s for $Y \in \Omega_i \cap B(X, r(X))$ contributes to the p.d.f. estimation of $I_i(X)$.

To get an optimal ϕ the functional in (2.10) is discretized first in order to avoid the difficulty in getting the EL equation for ϕ. Then we take derivative with respect to ϕ at each grid point. EL equations with respect to I_1 and I_2 can be derived directly from the first variation. For notation simplicity, we present our equations in two dimensional cases. Let $(x_i, y_j) = (ih, jh)$, where h is spacial step size, be a grid point ($1 \leq i \leq M, 1 \leq j \leq N$). The values of radius r and ϕ at point (x_i, y_j) are denoted by $r_{i,j}$ and $\phi_{i,j}$ respectively. The length term is discretized using backward finite difference for gradient in the x-direction,

and forward difference for y-direction. The corresponding discretized version of the energy functional in (2.10) is shown below.

$$
\begin{aligned}
&-\sum_{ij}\widetilde{H}(\phi_{i,j})\log\left(\frac{\sum\limits_{k,l\in B((x_i,y_j),r_{ij})}\widetilde{H}(\phi_{k,l})Q_1(I_1(x_i,y_j)-I_0(x_k,y_l))}{\sum\limits_{k,l\in B((x_i,y_j),r_{ij})}\widetilde{H}(\phi_{k,l})}\right)\\
&+\sum_{i=1}^{M}\sum_{j=1}^{N}\widetilde{H}(\phi_{ij})\log\sum_{k=max(i-r_{ij},1)}^{min(i+r_{ij},M)}\sum_{l=max(j-r_{ij},1)}^{min(j+r_{ij},N)}H(\phi(x_k,y_l))\\
&-\sum_{ij}(1-\widetilde{H}(\phi_{i,j}))\log\left(\frac{\sum\limits_{k,l\in B((x_i,y_j),r_{ij})}(1-\widetilde{H}(\phi_{i,j}))Q_1(I_2(x_i,y_j)-I_0(x_k,y_l))}{\sum\limits_{k,l\in B((x_i,y_j),r_{ij})}(1-\widetilde{H}(\phi_{i,j}))}\right)\\
&+\beta\sum_{ij}\widetilde{\delta}(\phi_{i,j})\frac{\sqrt{(\phi_{i,j}-\phi_{i-1,j})^2+(\phi_{i,j+1}-\phi_{i,j})^2}}{h}
\end{aligned}
\tag{3.1}
$$

Then, the corresponding evolution equation for ϕ_{ij} is the following:

$$
\begin{aligned}
\tfrac{d\phi_{ij}}{dt}=&\sum_{ij}\widetilde{\delta}(\phi_{i,j})\log\left(\frac{\sum\limits_{(k,l)\in B((x_i,y_j),r_{ij})}\widetilde{H}(\phi_{k,l})Q_1(I_1(x_i,y_j)-I_0(x_k,y_l))}{\sum\limits_{(k,l)\in B((x_i,y_j),r_{ij})}\widetilde{H}(\phi_{k,l})}\right)\\
&-\sum_{ij}\widetilde{\delta}(\phi_{i,j})\log\left(\frac{\sum\limits_{(k,l)\in B((x_i,y_j),r_{ij})}(1-\widetilde{H}(\phi_{i,j}))Q_1(I_2(x_i,y_j)-I_0(x_k,y_l))}{\sum\limits_{(k,l)\in B((x_i,y_j),r_{ij})}(1-\widetilde{H}(\phi_{i,j}))}\right)\\
&-\delta(\phi(X))\log\textstyle\int_{\Omega\bigcap B(X,r(X))}(1-H(\phi(Y)))dY\\
&+\delta(\phi(X))\log\textstyle\int_{\Omega\bigcap B(X,r(X))}(1-H(\phi(Y)))Q_1(I_2(X)-I_0(Y))dY\\
&+\sum_{m,n}H(\phi_{m,n})\frac{\delta(\phi_{i,j})Q_1(I_1(x_m,y_n)-I_0(x_i,y_j))}{\sum\limits_{(k,l)\in B((x_i,y_j),r_{ij})}H(\phi_{k,l})Q_1(I_1(x_m,y_n)-I_0(x_k,y_l))}\\
&-\sum_{m,n}(1-H(\phi_{m,n}))\frac{\delta(\phi_{i,j})Q_1(I_2(x_m,y_n)-I_0(x_i,y_j))}{\sum_{(k,l)\in B((x_i,y_j),r_{ij})}(1-H(\phi_{k,l}))Q_1(I_2(x_m,y_n)-I_0(x_k,y_l))}\\
&-\beta\widetilde{\delta}'(\phi_{i,j})\tfrac{\sqrt{(\phi_{i,j}-\phi_{i-1,j})^2+(\phi_{i,j+1}-\phi_{i,j})^2}}{h}\\
&-\beta\widetilde{\delta}(\phi_{i,j})\tfrac{2\phi_{i,j}-\phi_{i-1,j}-\phi_{i,j+1}}{h\sqrt{(\phi_{i,j}-\phi_{i-1,j})^2+(\phi_{i,j+1}-\phi_{i,j})^2}}\\
&-\beta\widetilde{\delta}(\phi_{i+1,j})\tfrac{\phi_{i,j}-\phi_{i+1,j}}{h\sqrt{(\phi_{i+1,j}-\phi_{i,j})^2+(\phi_{i+1,j+1}-\phi_{i+1,j})^2}}\\
&-\beta\widetilde{\delta}(\phi_{i,j-1})\tfrac{\phi_{i,j}-\phi_{i,j-1}}{h\sqrt{(\phi_{i,j-1}-\phi_{i-1,j-1})^2+(\phi_{i,j}-\phi_{i,j-1})^2}}
\end{aligned}
\tag{3.2}
$$

The first variation of (2.10) with respect to I_1 and I_2 leads to the following equations:

$$I_1(x) = \frac{\int_{B(x,r(x))} H(\phi)(y) I_0(y) dy}{\int_{B(x,r(x))} H(\phi)(y) dy} \tag{3.3}$$

$$I_2(x) = \frac{\int_{B(x,r(x))} (1 - H(\phi)(y)) I_0(y) dy}{\int_{B(x,r(x))} (1 - H(\phi)(y)) dy} \tag{3.4}$$

For ϕ, we impose boundary condition $\frac{\partial \phi}{\partial n} = 0, for\ X \in \partial\Omega, t > 0$ and initial condition $\phi(X,0) = \phi_0(X) for X \in \Omega$. I_1 and I_2 are always updated using (3.3) and (3.4) respectively. The choice of the initial ϕ is extremely flexible, even simple constant initial work for most of the images we have tried. This eliminates the work of creating the distance function of an initial curve that is required by regular level set methods.

The evolution equation (3.2) is solved using a semi-implicit scheme invoking the additive operator splitting (AOS) [18, 19, 20]. The AOS scheme splits the diffusion operator into a coordinate-by-coordinate fashion, and replaces the inverse of the sum by the sum of the inverse. It guarantees equal treatment of all coordinate axes, and can be implemented easily in arbitrary dimensions. Its computational complexity and memory requirement are linear to the number of pixels. It is stable for large step sizes.

The algorithm usually converges in fewer than 10 iterations, this makes reinitialization of level set function ϕ unnecessary. Parameter β affects experimental results and needs to be tuned for each image, this is inevitable for all the models that involve length term. All the results shown are based on the best choice of β that we have tried.

Table 3.1: Segmentation Accuracy

Gaussian Noise	0.01	0.05	0.1	0.25
SA	0.996	0.979	0.957	0.868
Salt Pepper Noise	0.05	0.15	0.25	0.50
SA	0.996	0.988	0.970	0.885

4. Validation and Application. We first validate the ability of model (2.10) (or equivalently (2.4)) in simultaneously segmenting and smoothing noisy images. Then, to demonstrate the advantage of the proposed model in dealing with segmentation and smoothing for images with higher level noise, we compare the proposed model with piecewise constant MS model (1.2), global Gaussian model (1.3-1.4) and the model proposed in [21] which is basically model (2.10) with fixed radius. We also compare our local nonparametric model (2.10) with the following global nonparametric model which utilizes intensities of I_0 in the whole Ω_i to estimate p.d.f. of $I_i(X)$ for $X \in \Omega_i$.

$$\begin{aligned} \min_{I_1, I_2, \phi} & -\int_\Omega H(\phi(x)) \log \left(\frac{\int_\Omega H(\phi(y)) Q_1(I_1(x) - I_0(y)) dy}{\int_\Omega H(\phi(y)) dy} \right) dx \\ & -\int_\Omega (1 - H(\phi(x))) \log \left(\frac{\int_\Omega (1 - H(\phi(y))) Q_1(I_1(x) - I_0(y)) dy}{\int_\Omega (1 - H(\phi(y))) dy} \right) dx \\ & +\beta \int_\Omega \delta(\phi(x)) |\nabla \phi(x)| dx \end{aligned} \tag{4.1}$$

Finally, we apply the proposed model to human brain T1 weighted magnetic resonance (MRI) images, FLAIR human brain MRI images and short axis human echocardiographic images to show the effectiveness of the model in medical image analysis .

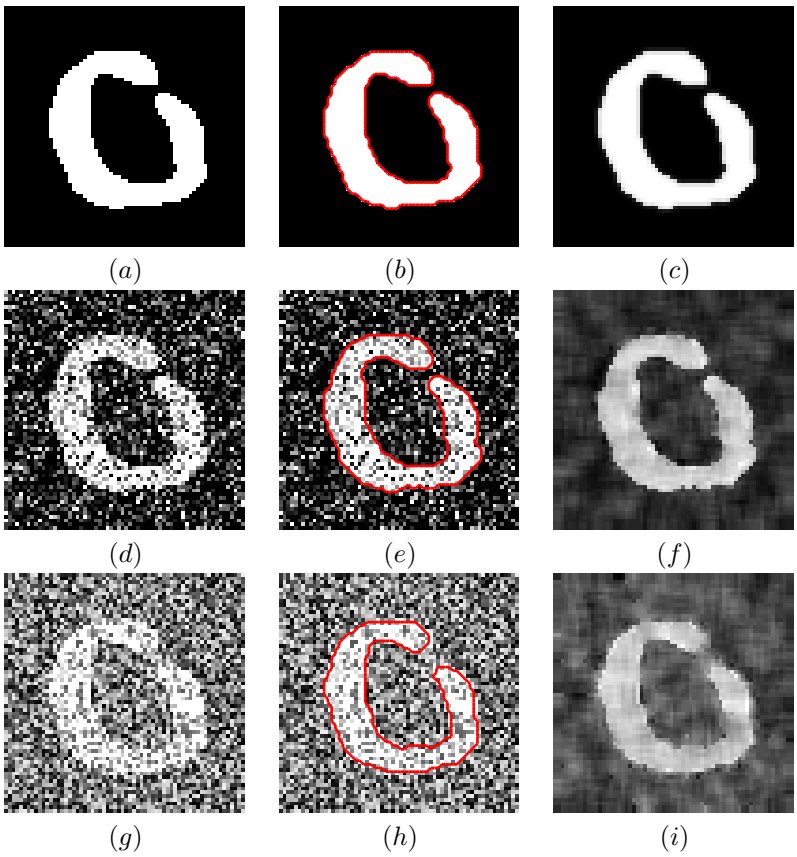


FIG. 4.1. *Segmentation results of a clean image and its noisy versions. From top to bottom, the left column: a clean cartoon C-shaped image, its noisy versions generated by adding white Gaussian noise with variance 0.05, and speckle noise with parameter 0.05; the middle column: segmentation results (red contours) obtained by applying the proposed model to the images next to them; the right column: smoothed versions of the images on the first column respectively. For all the results, time step size = .1,* $\beta = .01$*, M = 2, N=1, converge in 20 iterations.*

The first set of experiments aims to examine the accuracy of the proposed model (2.10) in segmenting a noisy image. We start from a clean image C-shaped cartoon image as shown in FIG. 4.1(a). By adding different levels of Gaussian noise and salt & pepper noise to it we generate various noisy images. Then we apply model (2.10) to each of the noisy image and the clean cartoon image to get segmentation of these images. To quantitatively observe how accurate the model (2.10) is in segmenting noisy images, we treat segmentation result of the clean cartoon image as the ground truth, then define a segmentation accuracy (SA) measure to be the ratio of number of pixels sharing the same segmentation with the ground truth over total pixel number. SA is between 0 and 1, the closer it is to 1, the better the model (2.10) is in segmenting this noisy image. Table 3.1 lists SA's corresponding to images obtained by adding random Gaussian noise with mean 0 and four different variances, and salt & pepper noise with four different parameters, respectively. Each of the SA shown in the table is the average value of two SA's corresponding to two different images obtained through adding the same type and the same strength of random noise at two different times. It is observed from Table 3.1 that when variance of Gaussian noise is lower than .1 or parameter of salt &

pepper noise is less than 0.25, the segmentation accuracy could be higher than 95%. And an obvious trend is that as noise level increases, SA decreases as expected.

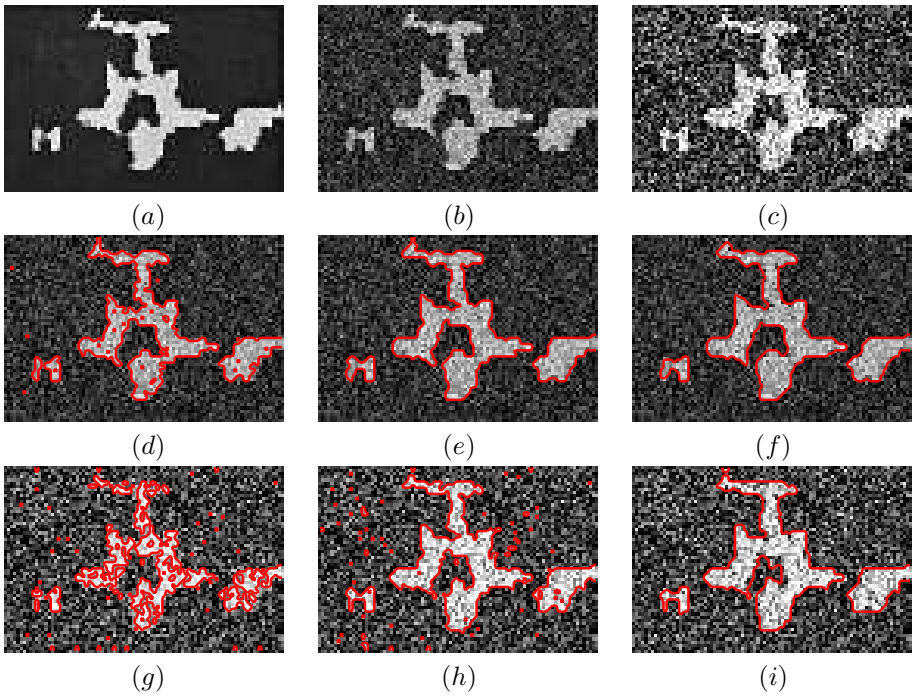


FIG. 4.2. *Compare three models in segmenting images with lower and higher level of noises. (a) A clean plane image; (b)-(c) Noisy plane images obtained by adding a random Gaussian noise with mean* 0*,variance* 0.01 *and* 0.05 *to (a), respectively; (d)-(f) Segmentation results of (b) obtained from piecewise constant MS with* $\beta = .05$*, global Gaussian model with* $\beta = .8$ *and the proposed model (2.10) with* $\beta = 10$*, respectively; (g)-(i) Segmentation results of (c) from piecewise constant MS model with* $\beta = .1$*, global Gaussian model with* $\beta = 1$ *and (2.10) with* $\beta = 10$*, respectively.*

FIG.4.1 shows segmentation and smooth results obtained by applying model (2.10) to the clean C-shaped cartoon image (FIG.4.1(a)) and its two noisy versions shown in FIG.4.1(d) and FIG.4.1(g). The image in FIG.4.1(d) is generated by adding a random Gaussian noise with zero mean and variance 0.05 to the image in (a), while the one in FIG.4.1(g) is obtained by adding speckle noise with parameter 0.05. Their boundaries depicted in red curves, which are found by applying model (2.10), are superimposed on the images in (a), (d), and (g), and shown in FIG.4.1(b), (e), and (h) respectively. It is observed that they are almost the same, actually, they are consistent up to 99.6% (SA = .996). FIG.4.1(c), (f), and (i) demonstrate the reconstructed smooth images of FIG.4.1(a), (d), and (g), respectively. The recovered images are much smoother and sharper than the corresponding original noisy images.

FIG.4.2 shows the comparison in effectiveness of segmentation among piecewise constant MS (1.2) model, global Gaussian model (1.3-1.4) and the proposed model (2.10). We add random Gaussian noise with mean 0 and variance 0.01 and 0.05 to a plan image as shown in FIG.4.2(a) to create two noisy images as shown in FIG.4.2(b) and FIG.4.2(c), respectively. Then all these three models are applied to the image in FIG.4.2(b) to get the segmentations (superimposed red contours) as shown in FIG.4.2 (d)-(f) respectively. Visually, not much difference among them is detected. This indicates all the three models are able to segment images with low level noise. But after applying these three models to the noisier image in

FIG.4.2(c), the segmentation results shown in FIG.4.2(g)-(i) respectively are very different. Moreover, it is visible that the result from model (2.10) is the best as it simultaneously segments and smoothes the image so that the noise are removed while the boundary of the plane is well kept. Since the image intensities of the object and background are no longer homogeneous, the piecewise constant MS is not suitable in this case. The result from global Gaussian model keeps plane boundary well but gains more background fault boundaries, as it allows variation of intensity in Ω_i's.

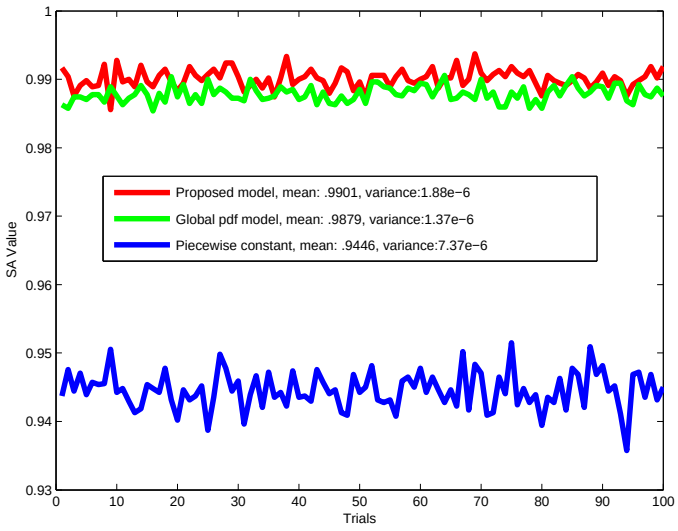


FIG. 4.3. *SA Graphs of images with Gaussian noise, obtained from three models: red: proposed model, green: global Gaussian model, blue: piecewise constant MS model.*

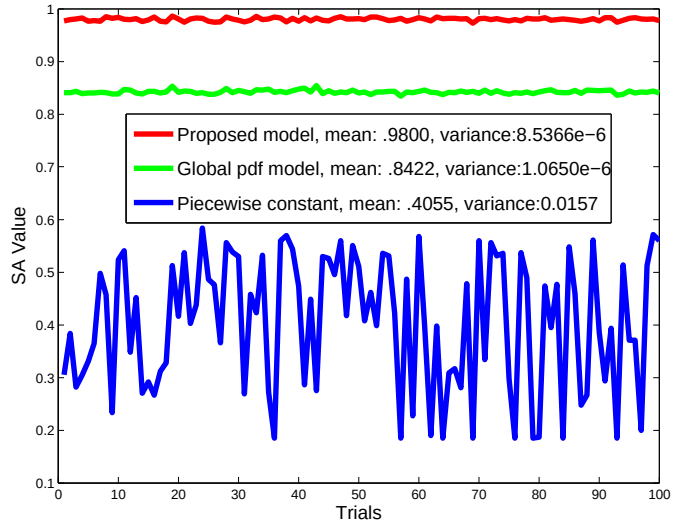


FIG. 4.4. *SA Graphs of images with speckle noise, obtained from three models: red: proposed model, green: global Gaussian model, blue: piecewise constant MS model.*

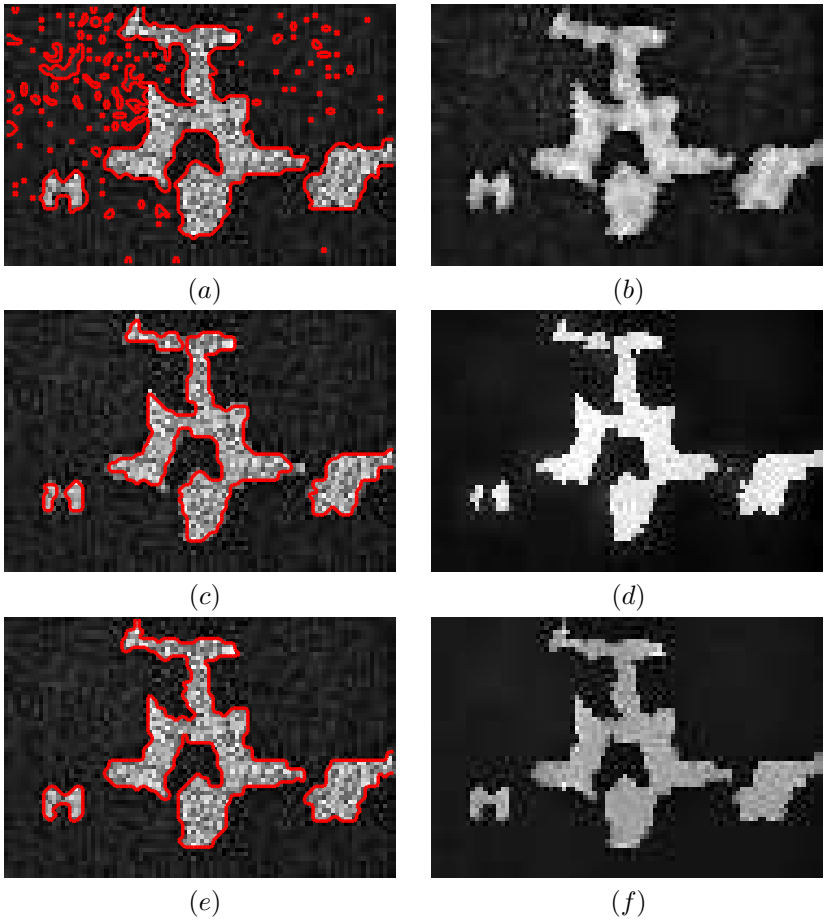


FIG. 4.5. *Comparison of segmentation and recovered image from the image in FIG.2.2(b) by the proposed model with fixed and adaptive radii. (a)-(b) Results from the proposed model with fixed radius 1; (c)-(d) Results from the proposed model with fixed radius 5;(e)-(f) Results from proposed model with adaptive radii defined in (2.7) where M = 5, and N=1. For all the results, time step size = .1,* $\beta = .5$*, converge in 10 iterations.*

A quantitative comparison of the three models is provided using SA. We add random Gaussian noise with mean 0 and variance 0.05 for 100 times to the clean plane image FIG.4.2(b) to get 100 noisy images. The three models are then applied to these 100 noisy images to obtain 100 SA's for each model, FIG.4.3 shows the graphs of SA for the three models. Mean value and variance of the 100 SA's obtained from the proposed model (shown in red contour), global Gaussian model (shown in green contour) and piecewise constant model (shown in blue contour) are $(.9901, 1.88 \times 10^{-6})$, $(.9879, 1.37 \times 10^{-6})$, $(.9446, 7.37 \times 10^{-6})$ respectively. The proposed model has the highest mean and smaller variance, while the piecewise constant has the lowest mean and smallest variance. It is obvious that the proposed model and the global Gaussian model are significantly better than piecewise constant MS, and the proposed model is slightly better than the global Gaussian model. The reason for global Gaussian model (1.3-1.4) being comparable with the proposed model (2.10) in this experiment is that the noisy images are generated by adding Gaussian noise, while model (1.3-1.4) is under Gaussian intensity assumption. For a better comparison, we repeat the

above work for 100 noisy images obtained by adding random speckle noise with parameter .04 for 100 times. Then the significant difference is observed and shown in FIG. 4.4: (2.10) maintains high mean (.9800) and low variance(8.5366×10^{-6}), but model (1.3) gets lower mean (.8422). Student t-test shows that model (2.10) is significantly better than (1.3-1.4) in processing these images with non-gaussian noise.

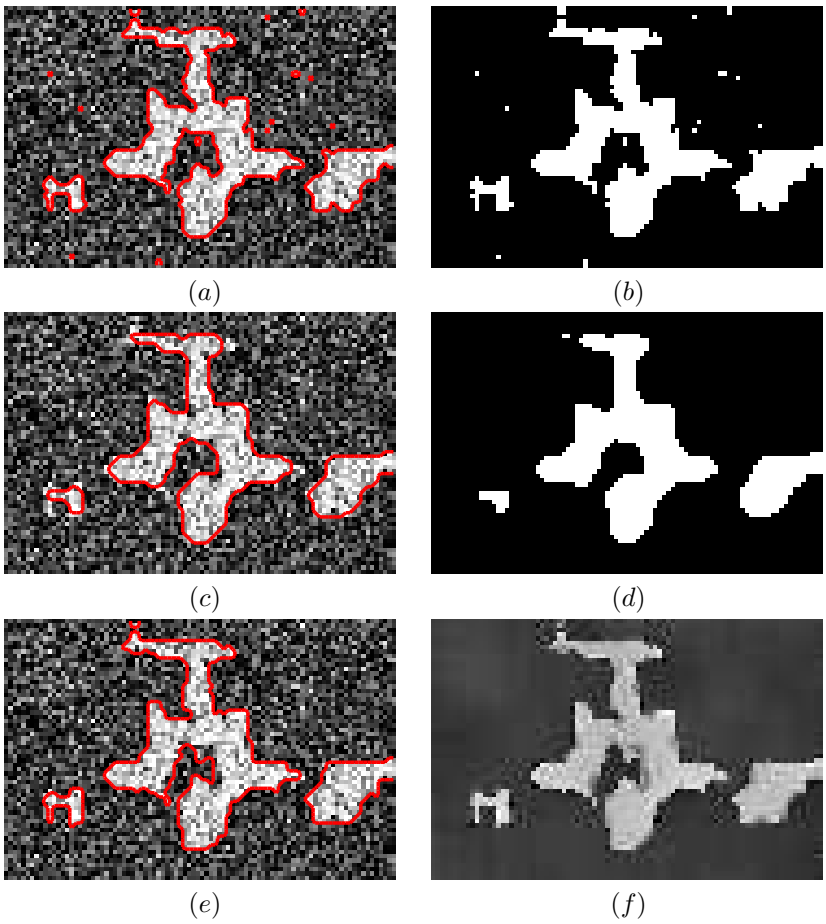


FIG. 4.6. *Comparison of global nonparametric model (4.1) with local nonparametric model (2.10). First row: intermediate segmentation and smoothing result from (4.1); Second row: final result from (4.1); Third row: final result from model(2.10).*

The second comparison is aimed to show the advantage of taking adaptive radius in simultaneous segmentation and smoothing. For this purpose we compare model (2.10) that uses adaptive radius with model in [21] that utilizes fixed radius. FIG.4.5 shows the segmentation and smoothing results of FIG.2.2(b) obtained from them. The results from model in [21] with fixed radius 1, and 5 are shown in the first and second rows, respectively. The third row depicts the results from model (2.10) with radius as defined in (2.7), $M = 5$ and $N = 1$. It is evident that the third row gives better results in both segmentation and smoothing. When the ball sizes are fixed for all X as in [21], it is very harder to choose one uniform ball size that works for all the locations. If it is too small, noise may not be removed sufficiently, while should a too big ball size is selected, fine structures may be lost.

The third comparison is to demonstrate the advantage of model (2.10) using local non-

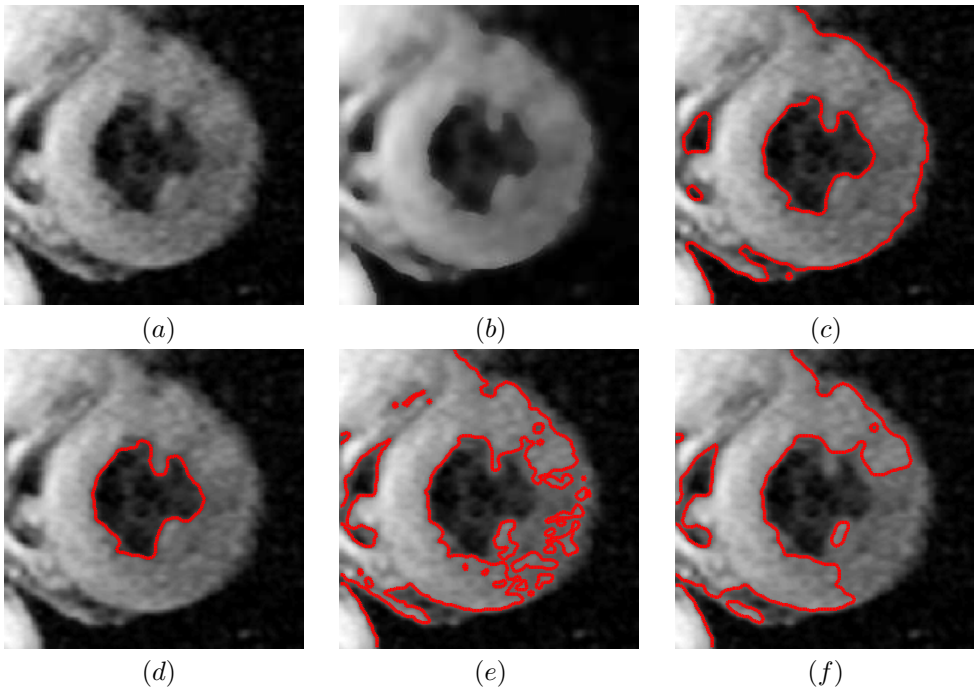
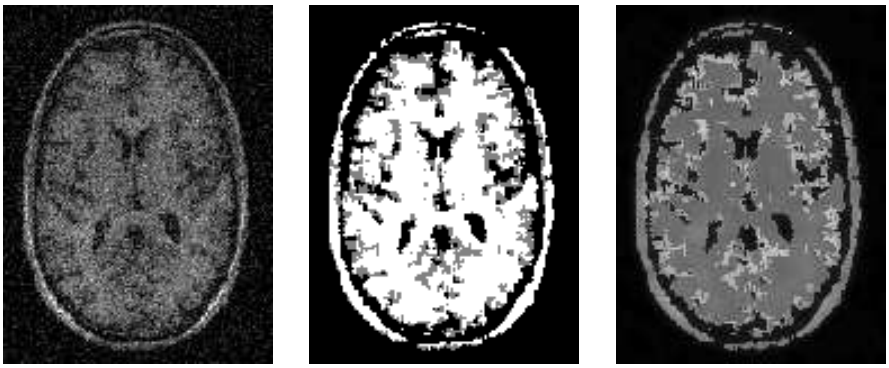


FIG. 4.7. *Segmentation and smoothing result of a human echocardiographic image. (a) A noisy human echocardiographic image; (b)-(c) Recovered image and segmentation result obtained by using model (2.10) respectively, (time step size is .1, $\beta = 70$, a = 4, b = 5, converge in 10 iterations); (d) Segmentation result (c) after applying mask of region of interest; (e)-(f) Segmentation result obtained by using global Gaussian model with $\beta = 0$ and 1 respectively.*

parametric density estimation (2.3) over the global nonparametric model (4.1) where the p.d.f. of I_i is estimated using all the $I_0(Y)$ for $Y \in \Omega_i$ as samples.

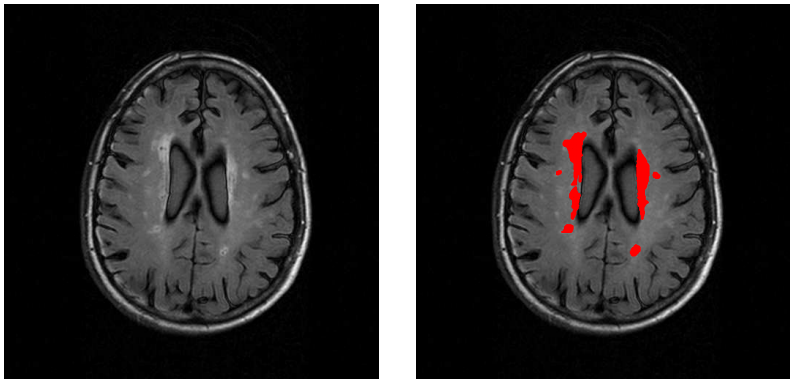
In FIG.4.6, the first row corresponds to the segmentation and smoothing results of (4.1) at the third iteration step. At this stage boundary of the object has been well caught, but the background noise has not been completely removed yet. While at the 10th iteration step (shown in the second row), when all the background noise has been removed for the first time, some fine structures of the object have disappeared. However, model (2.10) gives both good segmentation and smooth result at its convergent state as shown in the third row.

The last comparison is made on the segmentation of the left ventricle, whose volume is a measure to evaluate heart function, in a short axis echocardiographic image. It has been known that intensity distribution of Echocardiographic images is non-gaussian. In FIG. 4.7, (a) shows the observed image, the mushroom shaped object in the middle is the heart with papillary muscles excluded. (b) and (c) represent smooth and segmentation result from model (2.10) respectively. This time, we use radius $r(X)$ as defined in (2.6) with $a = 4, b = 5$. We do not use any restriction to the initial and evolving contours, so we need a mask to exclude background from result (b). (d) shows segmentation result after applying a mask of region of interest on (c), the red contour represents the boundary of the left ventricle with papillary muscles excluded. For comparison, we apply global Gaussian model to the same image with many different choices of β, all the segmentation results are not good, (e) and (f) correspond to those with $\beta = 0$ and 1. It is observed that model (2.10) successfully handles images whose intensity is of non-gaussian distribution.



(a) (b) (c)

FIG. 4.8. *Segmentation and smoothing result of a noisy T_1 human brain image. (a) A noisy T1 image;(b) Segmentation result obtained by using model (2.10): black: background and CSF; gray:gray matter; white:white matter; (c) Recovered smooth T1 image. Time step size is .1, $\beta = 80$, M = 3, N=2, convergence is in 10 iterations.*



(a) (b)

FIG. 4.9. *Segmentation result of a human brain FLAIR MR image. (a) A noisy human brain FLAIR MR image;(b) Segmentation obtained by applying model (2.10). Time step size is .1, $\beta = 60$, M = 0, N=0, convergence is in 10 iterations.*

Finally, we apply model (2.10) with radius as defined in (2.7), $M = 3, N = 2$ to segment several real medical images. FIG.4.8 demonstrates segmentation and smoothing results of a T1 weighted MRI human brain image as shown in FIG.4.8(a). This T1 image involves four phases: background, cerebral spinal fluid(CSF), gray matter, and white matter. Noise with unknown type is involved, intensity distribution is of unknown type also, model (2.10) is applied to this image, segmentation and smoothing results are shown in FIG.4.8(b) and (c) respectively. The segmentation result is quite comparable to known anatomy, and it is clearly visible that the recovered image FIG.4.8(c) is significantly sharper than the original one in FIG.4.8(a). In this experiment we use one level set in a hierarchical way [22]. First, we segment the image into two regions: white matter together with gray matter, and CSF together with background. Then we segment white matter from gray matter in the first subregion.

We have also applied model (2.10) to a slice of human brain MR fluid-attenuated inversion recovery (FLAIR) image shown in FIG. 4.9(a). To assist an automated measuring of white matter hyperintensity (leukoaralosis) volume we apply the proposed model to find its

boundary. Since this image is clean enough, smoothing is not necessary. Hence, we choose M and N to be 0 in our model. The segmentation result is given in (b), which is close to the one drawn by an expert.

5. An Existence Theorem for the Proposed Model. We first recall the definition and a few properties for the space of functions of bounded variation, that are needed in our proof [23, 24, 25]. Let Ω be an open subset of $\mathbb{R}^N$. We say $u \in L^1(\Omega)$ is a function of bounded variation, denoted by $u \in BV(\Omega)$, if the total variation of u

$$|Du|(\Omega) =: \sup_f \{ \int_\Omega u div f dx | f \in C_c^1(\Omega; \mathbb{R}^N), |f| \leq 1 \} < \infty.$$

$|Du|(\Omega)$ is also denoted by $\int_\Omega |Du|$. The space of functions of bounded variation, $BV(\Omega)$, is a Banach space endowed with the norm:

$$||u||_{BV(\Omega)} = ||u||_{L^1(\Omega)} + |Du|(\Omega).$$

Lower Semi-Continuity of variational measure
If $u, \{u_j\} \subset BV(\Omega)$ and $u_j \to u$ in $L^1(\Omega)$, then

$$|Du|(\Omega) \leq \liminf_{j \to \infty} |Du_j|(\Omega).$$

Compactness theorem in BV
Assume $\Omega \subset \mathbb{R}^N$ is open and bounded with Lipschitz boundary. If $\{u_n\}_{n \geq 1}$ is a bounded sequence in $BV(\Omega)$, then there exists a subsequence $\{u_{n_j}\}$ of $\{u_n\}$, and a function $u \in BV(\Omega)$, such that $u_{n_j} \to u$ strongly in $L^1(\Omega)$ as $j \to \infty$.

Discussions on existence
Denote the set $\{x \in \Omega | \phi(x) > 0\}$ by A, and rewrite (2.10) by using the characteristic function χ_A of A as follows:

$$E(\chi_A, I_1, I_2) = \beta \int_\Omega |D\chi_A| - \int_\Omega \chi_A(x) \log \left(\frac{\int_{B(x,r(x))} \chi_A(y) Q_1(I_1(x) - I_0(y)) dy}{\int_{B(x,r(x))} \chi_A(y) dy} \right) dx$$
$$- \int_\Omega (1 - \chi_A(x)) \log \left(\frac{\int_{B(x,r(x))} (1 - \chi_A(y)) Q_1(I_2(x) - I_0(y)) dy}{\int_{B(x,r(x))} (1 - \chi_A(y)) dy} \right) dx \qquad (5.1)$$

where the first term is the total variation of χ_A, which computes the length of the boundary of A. Model (2.10) aims to find a nonempty set A having finite perimeter in Ω, i.e. $\chi_A \in BV(\Omega)$, and functions $I_1(x),\ I_2(x) \in L^\infty(\Omega)$ that minimizes the energy functional (5.1).

Observe that if $I_1(x)$ is a minimizer of (5.1), $Q_1(I_1(x) - I_0(y)) = \frac{3}{4}(1 - (I_1(x) - I_0(y))^2)$ for a.e. $x \in A \cap \Omega$ and a.e. $y \in A \cap B(x, r(x))$ (otherwise the energy functional (5.1) at this minimizer would be infinity). Then, from Euler-Lagrange equations, the optimal I_1 and I_2 must be related to the optimal A (or χ_A) by

$$I_1(x) \int_{B(x,r(x))} \chi_A(y) dy - \int_{B(x,r(x))} \chi_A(y) I_0(y) dy = 0, x \in A \qquad (5.2)$$

$$I_2(x) \int_{B(x,r(x))} (1 - \chi_A(y)) dy - \int_{B(x,r(x))} (1 - \chi_A(y)) I_0(y) dy = 0, x \in \Omega \setminus A \qquad (5.3)$$

Since we are only interested in finding a nonempty set A with $\chi_A \in BV(\Omega)$, we may assume $\int_{B(x,r(x))} \chi_A(y) dy > 0$ for all $x \in A$. Then, equations (5.2) yields

$$I_1(x) = \frac{\int_{B(x,r(x))} \chi_A(y) I_0(y) dy}{\int_{B(x,r(x))} \chi_A(y) dy}, \quad x \in A \qquad (5.4)$$

Similarly, we have

$$I_2(x) = \frac{\int_{B(x,r(x))}(1-\chi_A(y))I_0(y)dy}{\int_{B(x,r(x))}(1-\chi_A(y))dy}, x \in \Omega \setminus A \tag{5.5}$$

By substituting (5.4) and (5.5) to (5.1) and denoting the resulting energy functional by E_1, we have

$$E_1(\chi_A) := E(\chi_A, I_1(\chi_A), I_2(\chi_A)),$$

where $I_1(\chi_A)$ and $I_2(\chi_A))$ are given by the right hand side of (5.4)and (5.5), respectively. Now we reduce the problem (5.1) to

$$\min_{\chi_A \in BV(\Omega)} E_1(\chi_A). \tag{5.6}$$

Next we will prove the following existence theorem for this problem.

Theorem

Let $\Omega \subset \mathbb{R}^N$ be open and bounded with Lipschitz boundary. Assume $I_0 \in L^\infty$ and $||I_0||^2_{L^\infty} \leq 1-\delta$ for a small constant $\delta > 0$. Then, there exists a set D with finite perimeter in Ω and

$$\chi_D = \min_{\chi_A \in BV(\Omega)} E(\chi_A).$$

Proof: Observe that $0 \leq Q_1(\cdot) \leq 1$, so $0 \leq \frac{\int_{B(x,r(x))}\chi_A(y)Q_1(I_1(x)-I_0(y))dy}{\int_{B(x,r(x))}\chi_A(y)dy} \leq 1$, from which we have $\log(\frac{\int_{B(x,r(x))}\chi_A(y)Q_1(I_1(x)-I_0(y))dy}{\int_{B(x,r(x))}\chi_A(y)dy}) \leq 0$, hence the second term in (5.6) is nonnegative. Similarly, the third term is also nonnegative. Therefore, the energy functional $E(\chi_A)$ is bounded below by 0 for all $A \subset \Omega$ with $\chi_A \in BV(\Omega)$. So there exists a minimizing sequence $\{\chi_{A_n}\}$ such that

$$\lim E(\chi_{A_n}) = \inf_{\chi_A \in BV(\Omega)} E(\chi_A). \tag{5.7}$$

Since the second and third terms in $E(\chi_{A_n})$ are nonnegative, from (5.7) we have that $\int_\Omega |D\chi_{A_n}|$ are uniformly bounded. Moreover, it is easy to see that $\{\chi_{A_n}\}$ is a bounded sequence in $L^1(\Omega)$. Hence, $\{\chi_{A_n}\}$ is a bounded sequence in $BV(\Omega)$. By compactness theorem in BV, there exists a subsequence of $\{\chi_{A_n}\}$, still denoted by $\{\chi_{A_n}\}$, and a function $u \in BV(\Omega)$, such that as $n \to \infty$

$$\chi_{A_n} \to u \ \textit{strongly in} \ L^1(\Omega), \ \textit{and a.e. on} \ \Omega.$$

Since χ_{A_n} takes values of either 0 or 1, u must be a characteristic function χ_D of some D with finite perimeter in Ω. So we have

$$\chi_{A_n} \to \chi_D \ \textit{strongly in} \ L^1(\Omega), \ \textit{and a.e. on} \ \Omega. \tag{5.8}$$

Moreover, by lower semi-continuity of BV-norm in L^1 topology, we have

$$\int_\Omega |D\chi_D| \leq \liminf_{j\to\infty} \int_\Omega |D\chi_{A_n}|. \tag{5.9}$$

From (5.8) it is not difficult to see that as $n \to \infty$,

$$\int_{B(x,r(x))} \chi_{A_n}(y)dy \to \int_{B(x,r(x))} \chi_D(y)dy, \tag{5.10}$$

and at each fixed $x \in \Omega$,

$$Q_1(I_1(\chi_{A_n})(x) - I_0(y)) \to Q_1(I_1(\chi_D)(x) - I_0(y)), \ \ for\ a.e.\ y \in \Omega,$$

as well as

$$\chi_{A_n}(y)Q_1(I_1(\chi_{A_n})(x) - I_0(y)) \to \chi_D(y)Q_1(I_1(\chi_D)(x) - I_0(y)), \ \ for\ a.e.\ y \in \Omega.$$

Since $\chi_{A_n}(y)Q_1(I_1(\chi_{A_n})(x) - I_0(y))$ is bounded by one, by the dominated convergence theorem, for each $x \in \Omega$, as $n \to \infty$

$$\int_{B(x,r(x))} \chi_{A_n}(y)Q_1(I_1(\chi_{A_n})(x) - I_0(y))dy \to \int_{B(x,r(x))} \chi_D(y)Q_1(I_1(\chi_D)(x) - I_0(y))dy. \tag{5.11}$$

Furthermore, from (5.8), (5.10) and (5.11) we get

$$\chi_{A_n}(x) \log\left(\frac{\int_{B(x,r(x))} \chi_{A_n}(y)Q_1(I_1(\chi_{A_n})(x) - I_0(y))dy}{\int_{B(x,r(x))} \chi_{A_n}(y)dy}\right) \tag{5.12}$$

$$\to \chi_D(x) \log\left(\frac{\int_{B(x,r(x))} \chi_D(y)Q_1(I_1(\chi_D)(x) - I_0(y))dy}{\int_{B(x,r(x))} \chi_D(y)dy}\right), \ \ pointwisely\ on\ \Omega. \tag{5.13}$$

On the other hand, note that from (5.4) we have $I_1(\chi_{A_n})(x)$ is the mean of I_0 on $B(x, r(x)) \cap A_n$. Hence, for each $x \in \Omega$

$$\sup_{y \in B(x,r(x)) \cap A_n} |I_1(\chi_{A_n})(x) - I_0(y)|^2 \leq \sup_{y \in B(x,r(x)) \cap A_n} |I_0(y)|^2 \leq 1 - \delta.$$

This implies that for all $x \in \Omega$ and $y \in B(x, r(x)) \cap A_n$

$$3\delta/4 \leq \chi_{A_n}(y)Q_1(I_1(\chi_{A_n})(x) - I_0(y)) \leq 1,$$

and hence, for all $x \in \Omega$

$$|\chi_{A_n}(x) \log\left(\frac{\int_{B(x,r(x))} \chi_{A_n}(y)Q_1(I_1(\chi_{A_n})(x) - I_0(y))dy}{\int_{B(x,r(x))} \chi_{A_n}(y)dy}\right)| \leq \log(3\delta/4). \tag{5.14}$$

Applying dominant convergence theorem to

$$-\int_\Omega \chi_{A_n}(x) \log\left(\frac{\int_{B(x,r(x))} \chi_{A_n}(y)Q_1(I_1(\chi_{A_n})(x) - I_0(y))dy}{\int_{B(x,r(x))} \chi_{A_n}(y)dy}\right) dx,$$

by using (5.13)and (5.14), we get that as $n \to \infty$ this term converges to

$$-\int_\Omega \chi_D(x) \log\left(\frac{\int_{B(x,r(x))} \chi_D(y)Q_1(I_1(x) - I_0(y))dy}{\int_{B(x,r(x))} \chi_D(y)dy}\right) dx.$$

By using the same argument we can also have $n \to \infty$

$$-\int_{\Omega}(1-\chi_{A_n}(x))\log\left(\frac{\int_{B(x,r(x))}(1-\chi_{A_n}(y))Q_1(I_1(\chi_{A_n})(x)-I_0(y))dy}{\int_{B(x,r(x))}(1-\chi_{A_n}(y))dy}\right)dx$$

$$\to -\int_{\Omega}(1-\chi_D(x))\log\left(\frac{\int_{B(x,r(x))}(1-\chi_D(y))Q_1(I_1(x)-I_0(y))dy}{\int_{B(x,r(x))}(1-\chi_D(y))dy}\right)dx.$$

Combining these with (5.8), (5.9) it yields that

$$E(\chi_D) \leq \liminf_{n\to\infty} E(\chi_{A_n}).$$

Hence χ_D is a minimizer.

6. Conclusion. A new nonparametric region based active contour model that utilizes both intensity and edge information was introduced to simultaneously segment and smooth images. To handle images with intensity inhomogeneity and higher level of noise, we apply a local nonparametric method to estimate intensity distribution of the smoothed image, using local intensity information of the observed image. Segmentation and smoothing results were obtained through minimizing the negative log likelihood the smoothed image and the length of objects boundaries. The minimization problem was implemented using semi-implicit iterative scheme. During the iteration, segmentation results were updated based on better smoothing images, smoothing results were then updated by using the better edge information derived from the better segmentation results. The local neighborhood sizes in density estimation were chosen adaptively based on image gradients, so that the model carries variate levels of smoothing at different locations, as a result, image features were well kept and edges were enhanced. The experimental results and comparison with several existing models on synthetic and real medical images with different types and levels of noise indicate the potential and advantage of the proposed model in simultaneously segmenting and smoothing images with higher level of noise.

In the proposed model the band width of the quadratic kernel is set to be one. We use variate local neighborhood sizes to control the use of samples to achieve adaptive smoothing. However, in the work of Azzabou et al. [14] the adaptive image smoothing is obtained by varying bandwidth in density estimation. As our future work, the relationship between variate bandwidth and adaptive neighborhood size on controlling image smoothing will be studied.

7. Acknowledgement. We thank Cate Price from college of public health and health professions at University of Florida for providing the FLAIR MRI image. The first author is supported by NIH R01 NS052831-01 A1 and NIH 5 R01 HL077810-02.

REFERENCES

[1] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, vol. 42, pp. 557–685, 1989.

[2] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[3] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shand model," *Int. Journal of Computer Vision*, vol. 50(3), pp. 271–293, 2002.

[4] S. C. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE PAMI*, vol. 18, pp. 884–90, 1996.

[5] M. Rousson and R. Deriche, "A variational framework for active and adaptive segmentation of vector values images," *Proceedings of workshop on motion and video computing*, pp. 56–61, Dec. 2002.

[6] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *Int. J. Comput. Vision*, vol. 46, no. 3, pp. 223–247, 2002.

[7] R. Duda and P. Hart, *Pattern classification and scene analysi*. John Wiley and Sons, Inc., 1973.

[8] W. Abd-Almageed, C. E. Smith, and S. Ramadan, "Kernel snakes: non-parametric active contour models," *IEEE Int. Conf. on Systems, Man and Cybernetics*, vol. 1, pp. 240–244, 2003.

[9] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, "Image segmentation using active contours: Calculus of variations or shape gradients," *SIAM Applied Mathematics*, vol. 63, no. 6, pp. 2128–2154, 2003.

[10] X. Huang, D. Metaxas, and T. Chen, "Metamorphs: Deformable shape and texture models," *Proceedings of IEEE Computer Soceity Conf. on Computer Vision and Pattern Recognition*, vol. I, pp. 496–503, 2004.

[11] J. Kim, J. W. F. III, A. Yezzi, M. Cetin, and A. S. Willsky, "A nonparamteric statistical method for image segmentation using information theory and curve evolution," *IEEE Transactions on Image Processing*, vol. 14:10, pp. 1482–1502, Oct. 2005.

[12] G. Hermosillo, C. Chefd'hotel, and O. Faugeras, "Variational methods for multimodal image matching," *Intel. Journal of Computer Vision*, vol. 50, no. 3, pp. 329–343, 2002.

[13] T. F. Chan, S. Esedoglu, and K. Ni, "Histogram based segmentation using wasserstein distances," in *First Int'l Conf. Scale Space and Variational Methods in Computer Vision, Springer*, Ischia, Italy, May/June 2007, pp. 697–708.

[14] N. Azzabou, N. Paragios, and F. Guichard, "Uniform and texture regions seperation in natural images towards mpm adaptive denoising," in *First Int'l Conf. Scale Space and Variational Methods in Computer Vision, Springer*, Ischia, Italy, May/June 2007, pp. 418–429.

[15] R. L. Rubank, *Nonparametric Regression and Spline Smoothing, 2nd Edition*. Marcel Dekker,Inc., 1999.

[16] P. Green and B. W. Silverman, *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall/CRC, 1994.

[17] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithm based on Hamilton-Jacobi formulation," *Journal of Computational Physics*, vol. 70, pp. 12–49, 1988.

[18] X. Tai, "Global extrapolation with a parallel splitting method," *Numerical Algorithm*, vol. 3, pp. 527–440, 1991.

[19] J. Weickert, B. Romeny, and M. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Trans. on Img. Proc.*, vol. 7, no. 3, pp. 398–410, March 1998.

[20] G. Kühne, J. Weickert, M. Beier, and W. Effelsberg, "Fast implicit active contour models," *Lecture Notes on Computer Science*, vol. 2449, pp. 133–140, 2002.

[21] W. Guo and Y. Chen, "Using non-parametric kernel to segment and smooth images simultaneously," *Proceedings of International Confernce in Image Processing*, pp. 217–220, 2006.

[22] S. Gao and T. Bui, "A new image segmentation and smoothing model," *Proc. of Int. Sym. on Biomed. Img.*, pp. 137–140, 2004.

[23] H. Federer, "Geometric measure theory," *Berlin, Heidelberg, New-York: Springer-Verlag*, 1969.

[24] L. C. Evans and R. R. Gariepy, "Measure theory and fine properties of functions," *CRC Press*, 1992.

[25] E. Guisti, "Minimal surfaces and functions of bounded variation," *Basel, Boston, Stuttgart: Birkhäuser Verlag*, 1984.