

# Level Set Segmentation with Both Shape and Intensity Priors

Siqi Chen and Richard J. Radke\*

Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY

chens@rpi.edu, rjradke@ecse.rpi.edu

## Abstract

We present a new variational level-set-based segmentation formulation that uses both shape and intensity prior information learned from a training set. By applying Bayes' rule to the segmentation problem, the cost function decomposes into shape and image energy parts. The shape energy is based on recently proposed nonparametric shape distributions, and we propose a new image energy model that incorporates learned intensity information from both foreground and background objects. The proposed variational level set segmentation framework has two main advantages. First, by characterizing image information with regional intensity distributions, there is no need to balance image energy and shape energy using a heuristic weighting factor. Second, by incorporating learned intensity information into the image model using a nonparametric density estimation method and an appropriate distance measure, our segmentation framework can handle problems where the interior/exterior of the shape has a highly inhomogeneous intensity distribution. We demonstrate our segmentation algorithm using challenging pelvis CT scans.

## 1. Introduction

In recent years, much work on geometric active contours, i.e., active contour models implemented via level set methods, has been proposed to solve a large number of image segmentation problems. The basic idea is that a contour  $C$  in a domain  $\Omega$  can be represented by the zero level set of a higher level embedding function  $\phi : \Omega \rightarrow \mathbb{R}$ . Rather than directly evolving the contour  $C$ , one evolves the level set function  $\phi$ . The embedding function  $\phi$  is defined as the signed distance function (SDF) with  $\phi > 0$  inside the shape,  $\phi < 0$  outside the shape and  $|\nabla\phi| = 1$  almost everywhere. Compared with traditional parametric active contours, also known as “snakes” [15], geometric active contour models have several advantages. For example, re-parametrization

is not required and the curve evolution process can easily handle topology changes like merging and splitting.

The evolution of the level set function  $\phi$  is governed by a partial differential equation (PDE). One way of obtaining this PDE is to convert from an evolution PDE of a parametric active contour [27]. Alternately, one can directly derive the PDE from a certain energy functional  $E(\phi)$  on the space of level set functions and subsequently minimize this functional by calculus of variations methods:

$$\frac{\partial\phi}{\partial t} = -\frac{\partial E(\phi)}{\partial\phi}. \quad (1)$$

These methods are known as variational level set methods [32]. One advantage of this formalism is that we can directly incorporate statistical prior information into the segmentation framework by the careful design of  $E(\phi)$ .

The introduction of statistical prior information into the level set framework has greatly enhanced the performance of many image segmentation algorithms. The basic idea is to obtain the best  $\phi$ , and thus the segmentation boundary  $C$ , by maximizing the conditional probability:

$$\begin{aligned} \hat{\phi} &= \arg \max p(\phi|I) \\ &\propto \arg \max (p(I|\phi) \cdot p(\phi)) \\ &= \arg \min (-\log(p(I|\phi)) - \log(p(\phi))) \\ &= \arg \min (E_{image} + E_{shape}) \end{aligned} \quad (2)$$

where  $I$  is the input image, and  $E_{image}$  and  $E_{shape}$  denote image and shape energy functions.

In this paper, we propose a new variational model that has several advantages over previous methods:

1. Although numerous image energies and shape energies have been proposed in recent years, most of them need a weighting factor to balance the two energy functionals in order to make both of them equally important to the segmentation, i.e.,

$$E = \frac{1}{\beta} E_{image} + E_{shape}. \quad (3)$$

Most papers treat this problem as minor and no further discussion on how to choose this weighting factor is

\*This work is supported in part by CenSSIS, the NSF Center for Subsurface Sensing and Imaging Systems, under the award EEC-9986821. We thank D. Michael Lovelock and Steve Jiang for providing the datasets.

made. However, as Dambreville et al. [11] pointed out, careful analysis is often required to tune this factor to get good segmentations. In this paper, we base the image energy on regional intensity distributions (i.e., normalized histograms) and show that the image energy and shape energy are approximately the same magnitude; hence no heuristic weighting factor is required.

2. When training images are available and faithfully represent the expected intensity variation of regions inside and outside the desired contour, we present a method that learns these intensity distributions by using non-parametric density estimation techniques and drives the segmentation accordingly. We demonstrate that the added information is important in a discussion of challenging medical image segmentation problems that exhibit highly inhomogeneous intensities, or no visible intensity changes around the true object boundaries.

## 2. Related Work

In the following, we briefly review some of the important related work on statistical approaches to level set segmentation. For a complete review and discussion, we refer the reader to a recent review paper [9].

### 2.1. Shape Energy

An early approach to modeling shape energy was based on the assumption that the segmentation should prefer a smooth partitioning boundary. This is usually referred to as a shape regularization term, where the curve's length, curvature or interior area is typically incorporated into a penalty term [15, 22]. Such geometric shape priors are still widely used in general image segmentation when further shape prior information is not available.

While this geometric shape constraint term helps to overcome a certain amount of noise, it gives no further information about which contours are more likely to appear than others beyond being smooth. Furthermore, penalizing the curve to have smaller length or curvature may cause the contour to shrink too much or miss some fine details. Because the segmentation problem can be formulated as a Bayesian inference, we can introduce higher-level prior knowledge into the shape model.

Leventon et al. [19] proposed to represent shapes by signed distance functions (SDFs) and perform Principal Component Analysis (PCA) on the set of training SDFs. They incorporated the learned shape information into a geodesic active contour model (GAC) [4] as an additional global shape force to drive the embedding function to the most likely shape. Tsai et al. [29] also performed PCA over the space of SDFs, proposing an efficient method of optimizing the PCA coefficients of the first few eigenmodes directly. However, direct PCA on the space of SDFs has the

limitation that since the space of SDFs is not linear, neither the mean level set nor a linear combination of eigenmodes will correspond to a valid SDF. Nevertheless, Kim et al. [16] pointed out that these methods work well in practice since the shape manifold that holds the training shapes is approximately flat and the space of SDFs can be approximated by a linear space when the shape variation is relatively small. Instead of linear PCA, Dambreville et al. [11] proposed to use kernel PCA (KPCA) to model the shape variation and incorporate the distance between shapes in feature space into a GAC framework. The use of KPCA allows the extraction of nonlinear structure in the space of SDFs, and therefore performs better than linear PCA.

One limitation associated with the PCA-based methods is that they typically use the first few principal components in the feature space to define the shape distance and shape prior. As a consequence, they may need a large number of eigenmodes to capture fine details. Defining a shape prior on the space of SDFs directly seems to be a more natural approach. Various groups have proposed to use a distance measure between the evolving level set and a reference level set as the shape prior term [2, 10, 26]. For a detailed analysis of different shape distances and statistical shape analysis, we refer the reader to the work by Charpiat et al. [7].

Learning a reference level set from training shapes often makes some statistical assumptions, e.g., the training shapes are distributed according to a Gaussian distribution. This is not always realistic since many real-world objects undergo complex shape variations in different scenarios. Recently, Cremers et al. [8] and Kim et al. [16] proposed to use the nonparametric technique of Kernel Density Estimation (KDE) to model the shape variation. This allows the shape prior to approximate arbitrary distributions. We will make use of this powerful technique in our model.

Often, rigid pose transformations are treated independently of shape. Frequently, another set of explicit pose parameters is added into the shape model and both types of parameters are iteratively optimized. As Cremers et al. [8] pointed out, such iterative optimization schemes often need careful tuning (e.g., gradient descent step size, order of parameter iterations). Instead, they proposed a shape dissimilarity measure invariant to certain pose transformations, using an intrinsic registration of the evolving level set function. We also use this approach in our work.

### 2.2. Image Energy

Unlike traditional edge-based active contour models [4, 15] that depend on large image gradients to stop the curve evolution, region based active contour models are more robust to noise and can detect objects with very diffuse boundaries. In this section, we focus on formulations where the image domain can be partitioned into a foreground object  $\Omega_{in}$  and background  $\Omega_{out}$ , and the intensity distributions in

different partitions are assumed to be independent. This lets us rewrite the image energy:

$$-\log(p(I|\phi)) = -\log(p(I|\Omega_{in})) - \log(p(I|\Omega_{out})). \quad (4)$$

One way of modeling the image energy is to assume that the image  $I$  in each region is characterized by the individual intensity values at different locations  $x$  and the intensity values are i.i.d. The above functional reduces to:

$$-\log(p(I|\phi)) = -\int_{\Omega_{in}} \log(p_{in}(I(x)))dx - \int_{\Omega_{out}} \log(p_{out}(I(x)))dx. \quad (5)$$

Equation (5) is the image energy model for many papers [33]. Chan and Vese [6] pioneered the idea of “active contours without edges” by assuming that both  $p_{in}$  and  $p_{out}$  are Gaussian distributions with the same variance. Their method can be also viewed as an extension of the classical Mumford-Shah model [22]. Rousson and Deriche [25] also proposed to model  $p_{in}$  and  $p_{out}$  as Gaussian distributions with different variances. Heiler et al. [12] used Laplace distributions to segment natural images. These parametric methods can be severely affected when the parametric model is not correct. Later, several authors proposed to use nonparametric methods to model those distributions [13, 17]. These methods are all based on the assumption that pixels belonging to one region all share the same probability distribution; thus, they can not handle the image inhomogeneity problem. In [28] and [30], the authors proposed to approximate the image  $I$  by two smooth functions  $u^-(x)$  and  $u^+(x)$  defined in the subregions  $\Omega_{in}$  and  $\Omega_{out}$ . Such methods have the ability to overcome the image inhomogeneity problem, but they are computationally expensive since they need to update  $u^-$  and  $u^+$  constantly by solving extra PDEs. Recently, several authors proposed to incorporate local intensity statistics into a variational framework [3, 18, 20]. Instead of modeling the intensity distribution globally for each region, local region models take the pixels’ spatial position into account. For example, in (5),  $p_{in}$  and  $p_{out}$  would also be functions of  $x$ .

Most of the above methods make statistical assumptions on the intensity values, and solve the segmentation in terms of a Maximum Likelihood approach. Another type of image energy model tries to partition the image so that the discrepancy between the intensity distributions of the inside and the outside is maximized. Yezzi et al. [31] proposed to use empirical moments such as the regional mean and/or variance to drive the segmentation. Although virtually all probability distributions can be uniquely represented by (either finite or infinite) sets of their moments, using moments to compare two distributions raises the question of the numbers and orders needed. Instead, one can use the entire shape of

the distributions, therefore taking into account the infinite number of empirical moments. Recently, Michailovich et al. [21] proposed to use the Bhattacharyya distance between the inside and outside distributions to drive the segmentation. However, since these methods have no probabilistic meaning, statistical information is difficult to incorporate.

In real world applications, especially in medical images, objects of interest often exhibit highly inhomogeneous intensity distributions that are problematic for existing methods. For example, Fig. 1 illustrates the segmentation of the rectum in Computed Tomography (CT) imagery using the Chan-Vese method [6] and a recent localized active contour method [18]. Due to the highly inhomogeneous intensity distribution of both the foreground and the background, neither of the two methods gives satisfactory results.

Incorporating learned information is one solution. When a reference contour or contours are available, many papers propose to combine a shape energy with an image energy discussed above. Since the shape energy is usually defined using the contour itself or the corresponding SDF, it is a global energy. However, most work assumes that intensities at different locations are independent of each other, and the image energy is based on local pixels and integrated over the domain. Therefore, the image energy is typically very large compared to the shape energy, and a weighting factor is introduced to balance them. In Fig. 2, we show the segmentation results using the nonparametric shape distribution [8] as the shape energy and the Chan-Vese model as the image energy using different values for the weighting parameter in (3). We can observe that in this case, the value of the weighting parameter plays an important role. If it is too small, then the image energy will dominate the energy functional and the segmentation result is similar to Fig. 1(c). If the weighting parameter is too large, the shape energy will dominate the energy functional and the result will not take the intensity information into account. In practice, tuning this parameter is not easy and may not be done automatically. Furthermore, too many local minima will make choosing a stopping criterion difficult.

The fundamental reason for the above segmentation failures or difficulties is that the image descriptor (e.g., pixel) is local while the shape descriptor (e.g., SDF) is global. Incorporating learned global intensity information (e.g., histograms) along with shape information into the segmentation framework is one possible solution. Existing methods typically assume that one reference distribution  $h_{in}$  and  $h_{out}$  for each region is available. Rousson and Cremers [24] incorporated such distributions into (5) and defined the image energy accordingly; however, the pixel independence assumption is still made. In [1, 5, 14], the authors proposed to minimize a distance measure between the reference and the current distributions  $h_{in}$ ,  $h_{out}$ ; however, since no training shapes are used, the competition between the shape en-

ergy and the image energy is not a issue in their work. In this paper, we extend the second approach to the situation when multiple training images and shapes are available, and combine the learned intensity and shape information into a systematic method where no weighting factor is required.

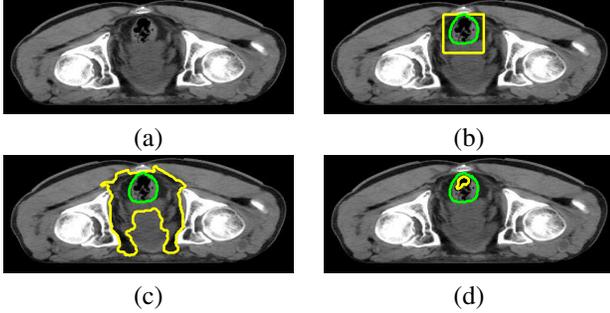


Figure 1. Rectum segmentation in CT imagery. The green curve is the ground truth. (a) The original image. (b) The initialization. (c) The segmentation results of the Chan-Vese method [6] with 500 iterations. (d) The segmentation results of the localized active contour method [18] with 400 iterations.

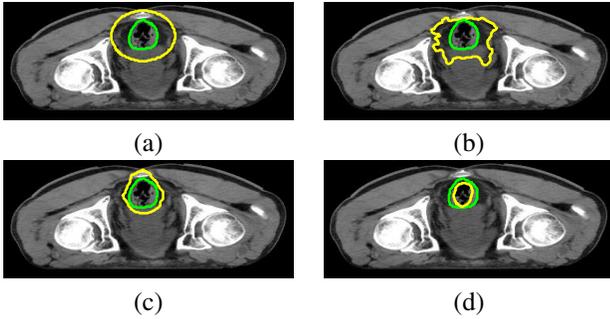


Figure 2. Rectum segmentation in CT imagery using KDE shape energy and Chan-Vese image energy in (3). The green curve is the ground truth. (a) The initialization. (b) The segmentation results with  $\beta = 10^3$ . (c) The segmentation results with  $\beta = 10^5$ . (d) The segmentation results with  $\beta = 10^7$ .

### 3. Proposed Segmentation Framework

#### 3.1. Bayesian Formulation

In this section, we present our Bayesian formulation of the level-set-based segmentation problem. Recall that the level set function is denoted  $\phi$  and the image information is denoted  $I$ . We revisit (2) by explicitly incorporating training shapes (which we denote  $\tilde{\phi}$ ) and corresponding images (which we denote  $\tilde{I}$ ) into the Bayesian analysis:

$$\begin{aligned}
 \hat{\phi} &= \arg \max p(\phi|I, \tilde{\phi}, \tilde{I}) \\
 &\propto \arg \max p(I|\phi, \tilde{\phi}, \tilde{I}) \cdot p(\phi|\tilde{\phi}, \tilde{I}) \\
 &= \arg \min \left( -\log(p(I|\phi, \tilde{\phi}, \tilde{I})) - \log(p(\phi|\tilde{\phi})) \right) \\
 &= \arg \min (E_{image} + E_{shape}) \tag{6}
 \end{aligned}$$

In our approach, we make no i.i.d. assumptions about the intensity values at different locations  $x$  in each region. Instead, we characterize the image information  $I$  using the intensity distributions inside the contour, outside the contour and over the entire region, denoted respectively as  $h_{in}$ ,  $h_{out}$  and  $h_{en}$ . Three key things to notice are that: 1)  $h_{in}$  and  $h_{out}$  are functions of  $\phi$ , 2)  $h_{en}$  is independent of  $\phi$  and 3) there is a simple relation between  $h_{in}$ ,  $h_{out}$ , and  $h_{en}$ :

$$h_{en} = \alpha h_{in} + (1 - \alpha) h_{out} \tag{7}$$

where  $\alpha$  is the ratio of the area inside the evolving contour to the entire area.

These allow us to decompose the image energy term as follows:

$$\begin{aligned}
 &p(I|\phi, \tilde{\phi}, \tilde{I}) \\
 &= p(h_{in}(\phi), h_{out}(\phi), h_{en}|\tilde{h}_{in}, \tilde{h}_{out}, \tilde{h}_{en}) \\
 &\propto p(h_{en}|h_{in}(\phi), h_{out}(\phi), \tilde{h}_{in}, \tilde{h}_{out}, \tilde{h}_{en}) \times \\
 &\quad p(h_{in}(\phi), h_{out}(\phi)|\tilde{h}_{in}, \tilde{h}_{out}, \tilde{h}_{en}) \\
 &\stackrel{=1}{=} \int p(h_{en}, \alpha|h_{in}(\phi), h_{out}(\phi), \tilde{h}_{in}, \tilde{h}_{out}, \tilde{h}_{en}) d\alpha \times \\
 &\quad p(h_{in}(\phi)|\tilde{h}_{in}) \times p(h_{out}(\phi)|\tilde{h}_{out}) \\
 &\stackrel{=2}{=} \int p(h_{en}|h_{in}(\phi), h_{out}(\phi), \alpha) p(\alpha|\tilde{h}_{in}, \tilde{h}_{out}, \tilde{h}_{en}) d\alpha \times \\
 &\quad p(h_{in}(\phi)|\tilde{h}_{in}) \times p(h_{out}(\phi)|\tilde{h}_{out}) \\
 &\stackrel{=3}{=} \int \delta(\alpha - \alpha(\phi)) p(\alpha|\tilde{\alpha}) d\alpha \times \\
 &\quad p(h_{in}(\phi)|\tilde{h}_{in}) \times p(h_{out}(\phi)|\tilde{h}_{out}) \\
 &= p(\alpha(\phi)|\tilde{\alpha}) \times p(h_{in}(\phi)|\tilde{h}_{in}) \times p(h_{out}(\phi)|\tilde{h}_{out}) \tag{8}
 \end{aligned}$$

Equality 1 holds because we assume that  $h_{in}$  and  $h_{out}$  are independent. Equality 2 holds because  $h_{en}$  is conditionally independent with the training images given  $h_{in}$ ,  $h_{out}$  and  $\alpha$ . Similarly,  $\alpha$  is independent of  $h_{in}$  and  $h_{out}$  when  $h_{en}$  is unknown. Equality 3 holds because of (7), where  $\delta$  is the Kronecker delta function.

Now (6) can be written as:

$$\begin{aligned}
 \hat{\phi} &= \arg \min \left( -\log(p(\alpha(\phi)|\tilde{\alpha})) - \log(p(h_{in}(\phi)|\tilde{h}_{in})) - \right. \\
 &\quad \left. \log(p(h_{out}(\phi)|\tilde{h}_{out})) - \log(p(\phi|\tilde{\phi})) \right) \\
 &= \arg \min (E_{\alpha} + E_{h_{in}} + E_{h_{out}} + E_{shape}) \tag{9}
 \end{aligned}$$

We use kernel density estimation [8, 16] based on  $N$

training samples to model each of the above probabilities:

$$p(\alpha|\tilde{\alpha}) \propto \frac{1}{N} \sum_{i=1}^N \exp\left(\frac{-(\alpha - \tilde{\alpha}_i)^2}{2\sigma_\alpha^2}\right) \quad (10)$$

$$p(h_{in}|\tilde{h}_{in}) \propto \frac{1}{N} \sum_{i=1}^N \exp\left(\frac{-D_h(h_{in}, \tilde{h}_{in}^i)^2}{2\sigma_{h_{in}}^2}\right) \quad (11)$$

$$p(h_{out}|\tilde{h}_{out}) \propto \frac{1}{N} \sum_{i=1}^N \exp\left(\frac{-D_h(h_{out}, \tilde{h}_{out}^i)^2}{2\sigma_{h_{out}}^2}\right) \quad (12)$$

$$p(\phi|\tilde{\phi}) \propto \frac{1}{N} \sum_{i=1}^N \exp\left(\frac{-D_\phi(\phi, \tilde{\phi}_i)^2}{2\sigma_\phi^2}\right) \quad (13)$$

Here,  $D_h$  and  $D_\phi$  are suitable distance measures defined in the next section.

### 3.2. Variational Level Set Formulation

With the level set representation, the above probability model (9) can be rewritten explicitly in terms of  $\phi$  where we make use of the following relations:

$$\alpha = \frac{\int_\Omega H(\phi(x))dx}{\int_\Omega dx} \quad H(x) = \begin{cases} 1, & \text{if } x \geq 0; \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

$$h_{in}(y) = \frac{\int_\Omega H(\phi(x))\delta(y - I(x))dx}{\int_\Omega H(\phi(x))dx} \quad (15)$$

$$F_{in}(y) = \frac{\int_\Omega H(\phi(x))H(y - I(x))dx}{\int_\Omega H(\phi(x))dx} \quad (16)$$

where  $F_{in}$  is the Cumulative Density Function (CDF) of  $h_{in}$ .  $F_{out}$  and  $h_{out}$  can be defined in a similar way.

The area of the set symmetric difference is commonly used as the distance measure between shapes:

$$D_\phi(\phi, \tilde{\phi}_i)^2 = \int_\Omega \left(H(\phi(x)) - H(\tilde{\phi}_i)\right)^2 dx \quad (17)$$

In our experiment, we use the extension proposed by Cremers et al. [8],

$$D_\phi(\phi, \tilde{\phi}_i)^2 = \int_\Omega \left(H(\phi(x + \mu_\phi)) - H(\tilde{\phi}_i)\right)^2 dx \quad (18)$$

where  $\mu_\phi$  is the center of gravity of the shape  $\phi$  and it is assumed that all training shapes have center of gravity at the origin. This distance measure has the advantage of being translation invariant.

The distance between two normalized histograms is defined by the CDF distance, also known as the linear Wasserstein distance ( $W_1$  distance) [5]:

$$D_h(h, \tilde{h}_i) = \int |F(y) - F_i(y)|dy \quad (19)$$

### 3.3. Gradient Descent Flow

By defining the total energy functional

$$E(\phi) = E_\alpha(\phi) + E_{h_{in}}(\phi) + E_{h_{out}}(\phi) + E_{shape}(\phi) \quad (20)$$

The gradient descent equation (1) can be written as:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E_\alpha(\phi)}{\partial \phi} - \frac{\partial E_{h_{in}}(\phi)}{\partial \phi} - \frac{\partial E_{h_{out}}(\phi)}{\partial \phi} - \frac{\partial E_{shape}(\phi)}{\partial \phi} \quad (21)$$

Since all the energy functionals are defined in terms of KDEs, we have the following common gradient flow:

$$\frac{\partial E_k(\phi)}{\partial \phi} = \frac{\sum_i w_i D(f(\phi), f_i(\tilde{\phi})) \frac{\partial D(f(\phi), f_i(\tilde{\phi}))}{\partial \phi}}{\sigma^2 \sum_i w_i} \quad (22)$$

where  $E_k$  can be any of the four functionals in (20) and  $f$  is correspondingly  $\alpha$ ,  $h_{in}$ ,  $h_{out}$  or the identity. We denote  $w_i = \exp\left(\frac{-D^2(f(\phi), f_i(\tilde{\phi}))}{2\sigma^2}\right)$ . The only thing we need to calculate is  $\frac{\partial D(f(\phi), f_i(\tilde{\phi}))}{\partial \phi}$  where  $D$  is taken as the four different distance measures.

By calculus of variations, it can be shown that:

$$\frac{\partial(\alpha(\phi) - \tilde{\alpha})}{\partial \phi} = \frac{\delta(\phi(x))}{\int_\Omega dx} \quad (23)$$

$$\frac{\partial D_h(h(\phi), h_i(\tilde{\phi}))}{\partial \phi} = \frac{\delta(\phi(x))}{Area} \times$$

$$\int \text{sign}(F(y) - F_i(y))(H(y - I(x)) - F(y))dy \quad (24)$$

where  $Area$  in (24) denotes the area of region inside or outside the contour respectively.  $F(y)$  denotes either  $F_{in}$  or  $F_{out}$  and  $F_i(y)$  is the training CDF defined accordingly.

The gradient flow of (18) is given in [8]:

$$\begin{aligned} \frac{\partial D_\phi(\phi, \tilde{\phi}_i)}{\partial \phi} &= 2\delta(\phi(x)) \left( (H(\phi(x)) - H(\tilde{\phi}_i(x - \mu_\phi))) + \right. \\ &\int H(\phi(x')) - H(\tilde{\phi}_i(x' - \mu_\phi))\delta(\phi(x'))\nabla\phi(x')dx' \\ &\left. \times \frac{x - \mu_\phi}{\int_\Omega H(\phi(x))dx} \right) \quad (25) \end{aligned}$$

### 3.4. Implementation

The numerical implementation of the gradient descent is done using a narrowband approach with the standard re-initialization technique [23]. In practice, the Heaviside function  $H$  is approximated by a smooth function  $H_\epsilon$  [6]:

$$H_\epsilon = \frac{1}{2} \left[ 1 + \frac{2}{\pi} \arctan\left(\frac{x}{\epsilon}\right) \right] \quad (26)$$

and  $\delta$  is approximated by the derivative of  $H_\epsilon$ :

$$\delta_\epsilon = \frac{\epsilon}{\pi(\epsilon^2 + x^2)} \quad (27)$$

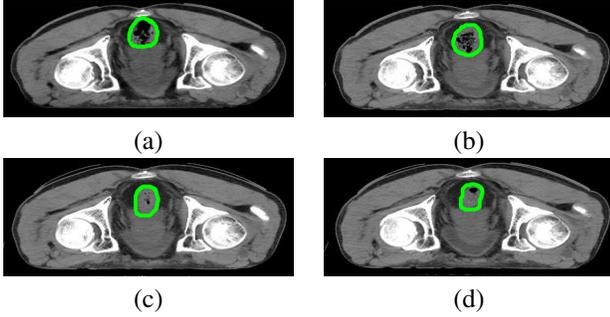


Figure 3. Four out of ten CT images of the rectum.

In this work, we choose  $\varepsilon = 1.0$ .

We defined the  $\sigma$  in each kernel density estimator as the mean inter-training-sample distance. This will give neither a too narrow nor too broad kernel width.

#### 4. Experimental Results

In this section, we present the results of our proposed method on challenging medical image processing problems. In our experiments, the image size is  $512 \times 512$ , but may be cropped for better visualization. The “outside” is defined as the region within the body but outside the object of interest. In our numerical experiments, we generally choose the time step  $\Delta t = 1$ . The evolution stops when  $\Delta E < 10^{-3}$ . Our algorithm runs at about 1 second for each iteration using unoptimized Matlab code on an Intel Core 2.66GHz CPU, 3.5GB RAM computer.

As a first experiment, we investigate the automatic segmentation of the rectum in CT imagery. Due to gas and filling, the rectum has unpredictable shape and intensity variation; furthermore, the intensity is usually inhomogeneous. Fig. 3 illustrates four out of the ten images along with the contours we used to build the model and test the method. The images come from several daily scans of a single patient during fractionated prostate cancer radiotherapy, where physicians outlined the ground truth by hand. All the experiments are done with the leave-one-out method. That is, each image was segmented using the other nine images for training.

In Fig. 4, we only use the new image energy as the driving force to evolve the curve. As we can see, the results contain the correct intensity information although the shape of the final result is not very good.

In Fig. 5, we illustrate the result using both the new image energy and the shape energy. Compared with Fig. 1, Fig. 2 and Fig. 4, the result clearly indicates an improvement both in terms of segmentation quality and speed of convergence.

In Table 1, we report the quantitative validation of our method by comparing the segmentation results with the ground truth using several different algorithms (including

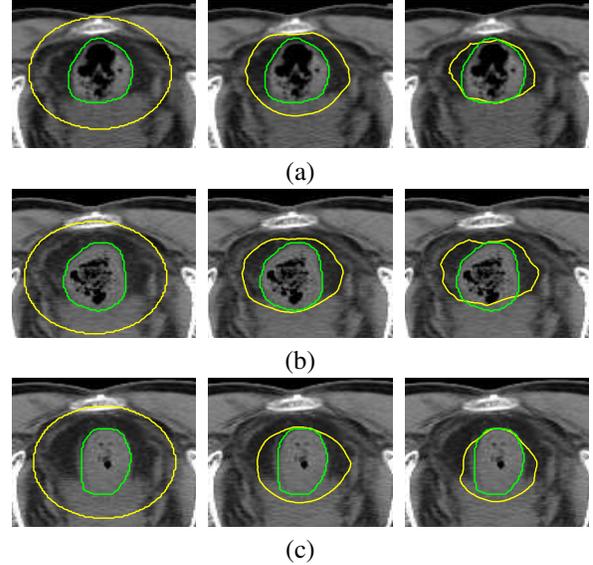


Figure 4. The segmentation result of the rectum using only the image energy. The green curve is the ground truth and the yellow curve is the evolving curve. Images from left to right are initialization, intermediate and final convergence results. The convergence occurs after 900 iterations. (a) Data set 1 (b) Data set 2 (c) Data set 3.

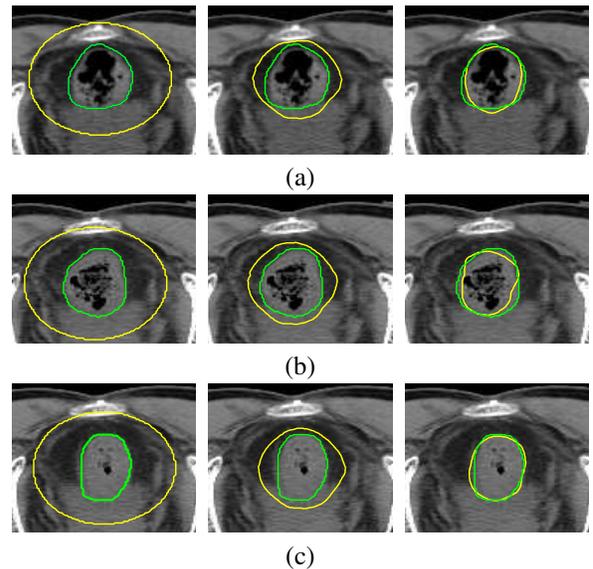


Figure 5. The segmentation result of the rectum using both the shape and image energies. The green curve is the ground truth and the yellow curve is the evolving curve. Images from left to right are initialization, intermediate and final convergence results. The convergence occurs after 300 iterations. (a) Data set 1 (b) Data set 2 (c) Data set 3.

the Chan Vese method [6], the localized active contour method [18] and Chan Vese +  $E_{shape}$  [8] with  $\beta = 10^5$ ). As metrics, we use the symmetric difference between the shapes (18), the CDF distance between the interior/exterior

measure	our $E_{shape} + E_{image}$		our $E_{image}$		Chan Vese method [6]		Localized method [18]		Chan Vese + $E_{shape}$ [8]	
	mean	median	mean	median	mean	median	mean	median	mean	median
$D_\phi$	25	24	30	30	604	630	120	115	104	97
$D_{h_{in}}$	7.87	7.60	20.86	18.12	32.50	31.56	28.24	26.42	25.24	28.54
$D_{h_{out}}$	0.44	0.36	1.15	1.16	2.74	3.14	3.24	3.34	2.14	2.06
$p_d$	0.90	0.92	0.88	0.90	1	1	0.51	0.62	0.91	0.91
$p_{fa}$	0.07	0.07	0.08	0.08	0.62	0.71	0.98	0.94	0.32	0.29
Dice	0.91	0.90	0.88	0.90	0.46	0.42	0.52	0.50	0.71	0.78

Table 1. Average quantitative measures of segmentation results over 10 images.

distributions (19), the probability of detection ( $p_d$ ), the probability of false alarm ( $p_{fa}$ ) and the Dice coefficient. The result indicates that using both the new shape and image energies outperforms the other methods.

In Table 2, we compare the  $E_{image}$  and  $E_{shape}$  of our new method applied to Data set 1 at different stages of curve evolution. The result indicates that the image energy and shape energy are at the same magnitude level throughout the process; therefore a weighting factor is not needed in our method. However, one can still tune the kernel width  $\sigma$  to balance the contribution of shape energy and image energy if desired.

Energy	Initial	Intermediate	Final
$E_{Shape}$	9.81	2.24	0.91
$E_{Image}$	8.82	2.45	0.33

Table 2. Shape energy and image energy comparison at different stages of curve evolution.

As a final experiment, we applied our segmentation framework to the problem of bone marrow segmentation in Cone Beam CT (CBCT) imagery, illustrated in Fig. 6. CBCT is known for its low image quality and contrast compared with traditional CT, so automatic processing is challenging. Nonetheless, we can see that the complex shape is segmented completely. This was a leave-one-out experiment with 11 datasets.

## 5. Conclusions and Future Work

In this work, we presented a new variational level set segmentation framework that incorporates both shape and intensity prior information. By characterizing the image information using regional intensity distributions, our method not only bypasses the questionable assumption that intensity values at each location are i.i.d., but also the heuristic weighting factor between the image energy and shape energy.

The area ratio  $\alpha$  is a novel aspect of our segmentation framework, acting as a coupling between  $h_{in}$  and  $h_{out}$  and compensating for the traditional independence assumption between them. It also enables the segmentation algorithm to

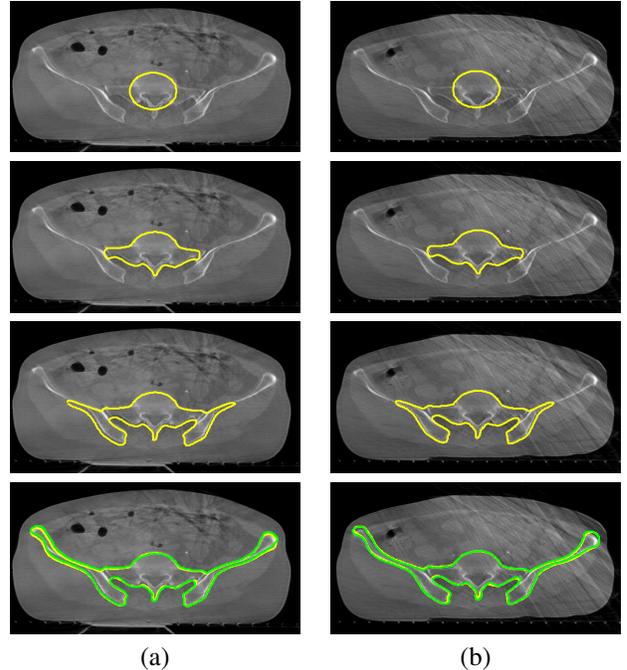


Figure 6. The segmentation result of bone marrow in Cone Beam CT using both the shape and image energies. The green curve is the ground truth and the yellow curve is the evolving curve. Images from top to bottom are initialization, intermediate1, intermediate2 and final convergence results. The convergence occurs after 700 iterations. (a) Data set 1 (b) Data set 2.

gracefully handle changes in object scale. We are currently investigating the role of  $\alpha$  in segmentation.

By using nonparametric density estimation techniques on histograms of regions inside and outside the contour, our method can incorporate learned intensity information into the segmentation framework, outperforming previous methods especially when the image contains highly inhomogeneous intensity variation. However, since histograms discard spatial information entirely, they are highly abstract representations of the image. In future work, we plan to investigate other global descriptors for the image energy. Furthermore, although local minima are not currently problematic for our method, it would be interesting to investigate the global minima of the energy functional we proposed here.

## References

- [1] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson. Image Segmentation Using Active Contours: Calculus of Variations or Shape Gradients? *SIAM Journal On Applied Mathematics*, 63(6):2128–2154, 2003.
- [2] X. Bresson, P. Vanderghenst, and J. Thiran. A Priori Information in Image Segmentation: Energy Functional Based on Shape Statistical Model and Image Information. In *Proc. International Conference on Image Processing*, 2003.
- [3] T. Brox and D. Cremers. On the Statistical Interpretation of the Piecewise Smooth Mumford-Shah Functional. In *Proc. SSVM*, 2007.
- [4] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic Active Contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
- [5] T. Chan, S. Esedoglu, and K. Ni. Histogram Based Segmentation Using Wasserstein Distances. In *Proc. SSVM*, 2007.
- [6] T. Chan and L. Vese. Active Contours Without Edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
- [7] G. Charpiat, O. Faugeras, and R. Keriven. Approximations of Shape Metrics and Application to Shape Warping and Empirical Shape Statistics. *Foundations of Computational Mathematics*, 5(1):1–58, 2005.
- [8] D. Cremers, S. Osher, and S. Soatto. Kernel Density Estimation and Intrinsic Alignment for Shape Priors in Level Set Segmentation. *International Journal of Computer Vision*, 69(3):335–351, 2006.
- [9] D. Cremers, M. Rousson, and R. Deriche. A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape. *International Journal of Computer Vision*, 72(2):195–215, 2007.
- [10] D. Cremers and S. Soatto. A Pseudo-Distance for Shape Priors in Level Set Segmentation. In *Proc. VLISM*, 2003.
- [11] S. Dambreville, Y. Rathi, and A. Tannenbaum. A Framework for Image Segmentation Using Shape Models and Kernel Space Shape Priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(8):1385–99, 2008.
- [12] M. Heiler and C. Schnörr. Natural Image Statistics for Natural Image Segmentation. *International Journal of Computer Vision*, 63(1):5–19, 2005.
- [13] A. Herbulot, S. Jehan-Besson, M. Barlaud, and G. Aubert. Shape Gradient for Image Segmentation Using Information Theory. In *Proc. IEEE ICASSP*, 2004.
- [14] S. Jehan-Besson, M. Barlaud, G. Aubert, and O. Faugeras. Shape Gradients for Histogram Segmentation Using Active Contours. In *Proc. IEEE International Conference on Computer Vision*, 2003.
- [15] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *International Journal of Computer Vision*, 1(4):321–331, 1988.
- [16] J. Kim, M. Çetin, and A. Willsky. Nonparametric Shape Priors for Active Contour-Based Image Segmentation. *Signal Processing*, 87(12):3021–3044, 2007.
- [17] J. Kim, J. Fisher III, A. Yezzi, M. Cetin, and A. Willsky. A Nonparametric Statistical Method for Image Segmentation Using Information Theory and Curve Evolution. *IEEE Trans. on Image Processing*, 14(10):1486–1502, 2005.
- [18] S. Lankton and A. Tannenbaum. Localizing Region-Based Active Contours. *IEEE Transactions on Image Processing*, 17(11):1, 2008.
- [19] M. Leventon, W. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. In *Proc. IEEE CVPR*, 2000.
- [20] C. Li, C. Kao, J. Gore, and Z. Ding. Minimization of Region-Scalable Fitting Energy for Image Segmentation. *IEEE Trans. on Image Processing*, 17(10):1940–1949, 2008.
- [21] O. Michailovich, Y. Rathi, and A. Tannenbaum. Image Segmentation Using Active Contours Driven by the Bhat-tacharyya Gradient Flow. *IEEE Transactions on Image Processing*, 16(11):2787–2801, 2007.
- [22] D. Mumford and J. Shah. Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
- [23] D. Peng, B. Merriman, S. Osher, H. Zhao, and M. Kang. A PDE-Based Fast Local Level Set Method. *Journal of Computational Physics*, 155(2):410–438, 1999.
- [24] M. Rousson and D. Cremers. Efficient Kernel Density Estimation Of Shape and Intensity Priors for Level Set Segmentation. In *Proc. Medical Image Computing and Computer Assisted Intervention*, 2005.
- [25] M. Rousson and R. Deriche. A Variational Framework for Active and Adaptive Segmentation of Vector Valued Images. In *Proc. IEEE Workshop on Motion and Video Computing*, 2002.
- [26] M. Rousson and N. Paragios. Shape Priors for Level Set Representations. In *Proc. European Conference on Computer Vision*, 2002.
- [27] J. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 1999.
- [28] A. Tsai, A. Yezzi, and A. Willsky. Curve Evolution Implementation of the Mumford-Shah Functional for Image Segmentation, Denoising, Interpolation, and Magnification. *IEEE Trans. on Image Processing*, 10(8):1169–1186, 2001.
- [29] A. Tsai, A. Yezzi Jr, W. Wells, C. Tempany, D. Tucker, A. Fan, W. Grimson, and A. Willsky. A Shape-Based Approach to The Segmentation of Medical Imagery Using Level Sets. *IEEE Trans. on Medical Imaging*, 22(2):137–154, 2003.
- [30] L. Vese and T. Chan. A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model. *Int. J. Comput. Vis.*, 50(3):271–293, 2002.
- [31] A. Yezzi Jr, A. Tsai, and A. Willsky. A Statistical Approach to Snakes for Bimodal and Trimodal Imagery. In *Proc. IEEE International Conference on Computer Vision*, 1999.
- [32] H. Zhao, T. Chan, B. Merriman, and S. Osher. A Variational Level Set Approach to Multiphase Motion. *Journal of Computational Physics*, 127(1):179–195, 1996.
- [33] S. Zhu and A. Yuille. Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.