Segmentation of Kidney From Ultrasound Images Based on Texture and Shape Priors

Jun Xie, Yifeng Jiang, and Hung-tat Tsui, *Member, IEEE*

Abstract—**This paper presents a novel texture and shape priors based method for kidney segmentation in ultrasound (US) images. Texture features are extracted by applying a bank of Gabor filters on test images through a two-sided convolution strategy. The texture model is constructed via estimating the parameters of a set of mixtures of half-planed Gaussians using the expectation-maximization method. Through this texture model, the texture similarities of areas around the segmenting curve are measured in the inside and outside regions, respectively. We also present an iterative segmentation framework to combine the texture measures into the parametric shape model proposed by Leventon and Faugeras. Segmentation is implemented by calculating the parameters of the shape model to minimize a novel energy function. The goal of this energy function is to partition the test image into two regions, the inside one with high texture similarity and low texture variance, and the outside one with high texture variance. The effectiveness of this method is demonstrated through experimental results on both natural images and US data compared with other image segmentation methods and manual segmentation.**

Index Terms—**Image segmentation, kidney segmentation, texture and shape prior, ultrasound image processing.**

I. INTRODUCTION

IMAGE segmentation is often the first step for image analysis and is a key basis of many higher-level activities such as visualization, compression, medical diagnosis and other imaging applications. The driving problem discussed in this paper is the segmentation of kidney from medical ultrasound (US) images. US imaging allows faster and more accurate procedures due to its realtime capabilities. Moreover, it is inexpensive and easy to use. The accurate detection of organs or objects from US images plays a key role in many applications. However, compared with other medical imaging modalities [e.g., computed tomography (CT) and magnetic resonance imaging (MRI)], US is particularly difficult to segment since the quality of the images is relatively low [1]. For instance, the speckle fluctuations in the signal are proportional in magnitude to the signal strength. This property leaves US images with significant noise even in very bright regions. Moreover, because of attenuation of the probing sound wave by sound absorbing tissues, the appearance of most tissues change greatly such as the intensity value varies and the boundary is not always complete and prominent.

Bouma *et al.* [2] performed a study of quantitative evaluation of (semi)-automated segmentation of US images and showed that even manual segmentation of noisy US images is not straight forward. On the other hand, reliable semi-automatic segmentation methods offer the potential advantage of making the measurement process more consistent [3]. Therefore, in this paper, we direct our research to develop a semi-automatic segmentation framework, using both texture and shape priors, for kidney segmentation from noisy US images.

Most image segmentation methods focus on region growing or active contours. However, the interference of speckle noise makes region growing methods [4] unreliable to classify image pixels. The active contour methods have been applied to automatically segment the boundaries in US images for the cortex of the brain [5], ovarian follicles [6], and for left-ventricular boundaries in echo-cardiograms [7]. Unfortunately, the basic active contour method is not adequate for our application of kidney segmentation since the tissue-tissue boundaries of kidney are relatively more difficult to localize in US images. In this paper, we present a texture and shape priors based segmentation framework for kidney US segmentation because we believe that a prior model of the expected anatomical structure is a significant advantage for segmenting them from US images.

One of the most used methods to model shape priors is statistical modeling. Cootes *et al.* [8], [9] proposed the active shape models (ASM) which relies on the statistics of an object's shape and gray-level appearances gathered from a training set of manually land-marked instances of the object. They developed a parametric point distribution model for describing the segmenting curve by using linear combinations of the eigenvectors that reflect variations from the mean shape. However, this representation does not contain any explicit information about the point connectivity. Moreover, because the techniques of automatic placement of landmarks are currently still under development (e.g., [10], [11]), landmarks are most frequently obtained manually and it is a time-consuming, error-prone and subjective work.

In [12], Chen *et al.* represented shapes using a collection of points. They applied clustering methods instead of statistical methods to get the shape prior model which is the average shape of n given curves with similar shape, but different size, orientation and translation. However, the similarity of shapes in this method is measured by area information which makes it highly time-consuming. Based on global-to-local registration [13] and prior region statistical properties, Rousson and Paragious [14] showed a method to recover a segmentation map in accordance with the shape prior model as well as a rigid registration between the map and the model. This method is good at accounting for local degrees of variability and local shape variations, but it consists of $o(N^2)$ variables and, thus, is unstable. Using the Gabor

filter bank to characterize the prostate boundaries in a multiscale and multiorientation fashion, Shen *et al.* [15] proposed a statistical shape model for the automatic prostate segmentation in transrectal US images.

In [16], Leventon *et al.* proposed a segmentation method including two steps: initial segmentation and its correction based on a shape prior model. The model is obtained through a principle component analysis (PCA) operation on a collection of signed distance maps of the training shapes. In this framework, the boundaries are represented as the zero level set of a 2-D scalar function [17]. This representation is intrinsic (independent of parameterizations) and is topologically flexible since different topologies of the curve are represented by the scalar function's constant topology. This ability is valuable for realistic image-guided diagnosis. Moreover, this scheme can generally provide smooth surfaces with good local dependence. This property results in a reasonable representative shape model of the training shapes. The same scheme was successfully applied in [18] where segmentation was implemented by calculating parameters of the implicit model that optimize the region-based energy functionals proposed in [19] and [20]. Most recently, Pons *et al.* [21] show a method to overcome the lack of point correspondences in the level set framework. The explicit backward correspondences from the evolving interface to the initial one are maintained by advecting the initial point coordinates with the same speed as the level set function. This makes it possible for the level set scheme to give a more substantial solution to point correspondences. Therefore, we also employ this representation of the segmenting curve in our approach and combine a texture prior into the constructed deformable shape model.

There have been some studies to incorporate texture prior knowledge into the active contour method [22] or the level set scheme [17]. Zhu and Yuille [23] proposed a statistical variational approach for texture image segmentation which combines the geometrical features of a snake/balloon model and the statistical techniques of region growing. However, in this method, orientation information is ignored and the change of topology is handled in an unnatural manner. Chakraborty *et al.* [24] proposed a method incorporating the intensity gradient, region features and shape prior information within a deformable boundary framework. The region information is expressed through the intensity homogeneity which limits this method's applicability to texture images. Based on [25], there are several approaches for supervised bi-modal segmentation (e.g., [19], [26], and [27]) or three-modal segmentation (e.g., [20] and [28]). However, in those statistical approaches, the important boundary information is ignored.

Paragios and Deriche [29] proposed a supervised texture segmentation method based on Geodesic Active Regions [30] and level set method. In this method, the segmentation is implemented by finding the best minimal length geodesic curve that consists of image pixels with high boundary probabilities. But this method needs to model each of the main texture patterns in the input image. This indicates that, for each image to be segmented, a long individual learning phase is indispensable which obstructs its success in real-time applications such as US diagnosis.

In this paper, we present a supervised image segmentation framework for extracting the organs of interest from US images. The texture model is constructed by estimating the
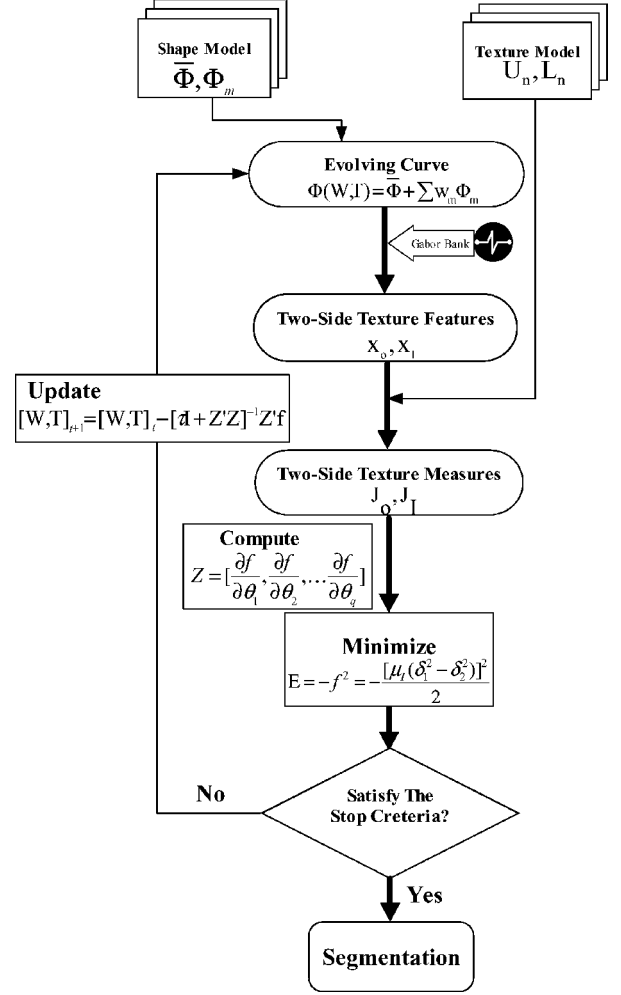

Fig. 1. Summary of the proposed texture and shape priors based segmentation method. $\bar{\Phi}$ is the mean shape and Φ_m $(m = 1, 2, \ldots, M)$ are the major M variability of the training shapes. $\mathbf{x_o}$ and $\mathbf{x_I}$ are the response vectors obtained by applying the filters bank in the outside and inside half planes, respectively. $W = \{w_i\}$ are mixing weights and $T = (s, \beta, T_x, T_y)$ are parameters of a similarity transformation. J_o and J_I are the texture similarity measures calculated as (26). Function E is the energy function define in (27). Parameters $[W, T]$ are updated as described in Section III-B.

parameters of a set of mixtures of half-planed Gaussians for different training directions. Building on the region-based information through the prior texture model, the segmentation is implemented by calculating the parameters of an implicit representation of the shape model [16] to minimize a novel texture-based energy function. A gradient descent algorithm is employed to improve the efficiency of the optimization process.

Fig. 1 shows a summary of the proposed segmentation method. In the training phase, a shape model is constructed through the dataset of training shapes. This shape model includes a mean shape $\bar{\Phi}$ and a set of eigenshapes Φ_m. Meanwhile, through a proposed two-sided convolution strategy, the training texture patterns are convoluted by a bank of Gabor filters and the respondences are modeled as a texture model including the upper half plane models U_n and the lower half plane models L_n. In the segmentation phase, the segmenting curve Φ is represented by the mean shape $\bar{\Phi}$ and the eigenshapes Φ_m with the mixing weights $W = \{w_i\}$ and the parameters $T = (s, \beta, T_x, T_y)$ of a similarity transformation.

The same bank of filters are then applied on two sides of the curve through the two-sided convolution strategy. This process provides the texture features I_o and I_I for the outside and inside region, respectively. Based on the texture prior models U_n and L_n, texture measures J_o and J_I of two independent regions are calculated. The novel texture-based energy function E is minimized using the optimization algorithm introduced in Section III-B. Finally the iteratively updated $[W, T]_{t+1}$ results in the deformation of the segmenting curve Φ, until the stop criteria is satisfied.

We evaluate our method using experiments on both natural images and clinical data. We also show the application of our method on kidney segmentation in US images. The experimental results are promising and show that our approach can segment the kidney from US images efficiently and accurately, compared with other image segmentation methods and manual segmentation.

The rest of the paper is organized as follows. In Section II, we briefly review the previous work on filtering theory and present the definition of our two-sided texture model. The shape modeling method of [16] is also introduced as it is a part of our segmentation scheme. In Section III, we describe the proposed texture and shape prior based segmentation method. A novel texture-based energy function and the details of the optimization are presented as well. Experimental results are demonstrated and discussed in Section IV. Our conclusions are presented in Section V.

II. TEXTURE AND SHAPE MODELING

A. Texture Feature Extraction

The first crucial step for texture analysis is to extract proper texture features that have good discrimination properties over a wide range of textures [31]. During the last twenty years, the filtering theory has been successfully applied for texture image processing. In this framework, texture features are extracted by applying the input image with a bank of filters of various orientations and spatial frequencies. Different filters have been proposed including Gabor filters [32], Gaussian derivatives [33] and wavelet transform [34]. Among those different selections of filters, Gabor filters are often used for their good performance in many texture analysis applications including texture classification, texture segmentation, image registration and image retrieval. Daugman [35] found that Gabor filters are closely related to the function of simple cells in the primary visual cortex of primates. Manjunath and Ma [31] have shown that the Gabor decomposition features are optimal in the sense of minimizing the joint two-dimensional uncertainty in space and frequency. Moreover, for US image processing, the Gaussian factor in Gabor filters can smooth the speckle noise in US images.

Because of those good properties, we adopt Gabor filters as the extractor of texture features. Filtered with a bank of Gabor filters, the input image is decomposed into multiple oriented spatial frequency channels, and the channel envelopes (amplitude and phase) are used to form the feature maps. The family of two-dimensional Gabor functions is defined as

$$G(x, y) = \exp\left(-\frac{x'^2}{2\sigma_x^2}\right) \exp\left(-\frac{y'^2}{2\sigma_y^2}\right) \cos\left(\frac{2\pi x'^2}{\lambda} + \varphi\right) \tag{1}$$
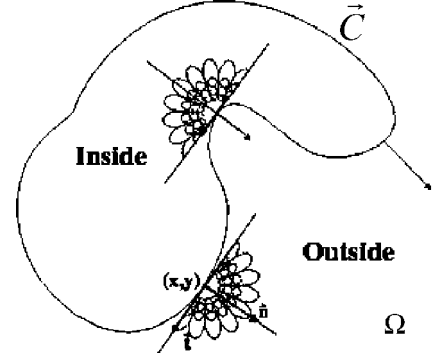

Fig. 2. Schematic explanation of the two-sided texture features extraction. For each step of the deformation of the segmenting curve C, the texture features of inside region and outside region are extracted respectively, using a bank of Gabor filters with different orientations and scales.

where

$$\begin{cases} x' = (x - m_x)\cos(\gamma) - (y - m_y)\sin(\gamma) \\ y' = (x - m_x)\sin(\gamma) + (y - m_y)\cos(\gamma) \end{cases}. \tag{2}$$

Parameter γ specifies the orientation and φ is a phase offset. Parameter pair (m_x, m_y) specifies the center of a receptive field in image coordinates and (σ_x, σ_y) determines the size of the receptive field. Term $1/\lambda$ is the preferred spatial frequency of function G.

In our experiments, following the method in [36], we used the Gabor filters with eight equidistant preferred orientations ($\gamma = \{22.5°, 45°, \ldots, 180°\}$) and three spatial frequencies ($\lambda = \{5.47, 8.20, 10.93\}$), resulting in 24–D feature vectors. The size of the receptive field is determined by the standard deviation $\sigma_y = 2\sigma_x = 1.12\lambda$ as proposed in [37].

When using the above Gabor filter bank to extract texture features, we choose a two-sided convolution strategy. Given a straight line in the 2-D image plane, the space is bisected into two half planes and the convolution is performed on each of them, respectively as illustrated in Fig. 2. This approach has three advantages. First, it is more robust to deform the segmenting curve based on separated texture measures than on that of the integrated region across the segmenting curve. Second, it enables the construction of a texture-based energy function, which depends on the properties of two individual regions inside and outside the curve. Third, it allows the segmenting curve to reach the points near the boundaries of objects of interest more easily.

In our case, we apply the above filter bank on training images in six even-spaced directions. Let L_n denote a straight line along the nth direction through point $P(x, y)$, where $n = (0, 1, \ldots, 5)$. The image plane is then bisected by L_n into two half planes, the upper half plane denoted by H_u and the lower half plane denoted by H_l. We compute the responses R_n^u and R_n^l by applying the Gabor bank on both sides of L_n as

$$R_n^u = \sqrt[2]{\left(r_{n,\varphi=0}^u\right)^2 + \left(r_{n,\varphi=-\frac{1}{2}\pi}^u\right)^2} \tag{3}$$

$$R_n^l = \sqrt[2]{\left(r_{n,\varphi=0}^l\right)^2 + \left(r_{n,\varphi=-\frac{1}{2}\pi}^l\right)^2} \tag{4}$$
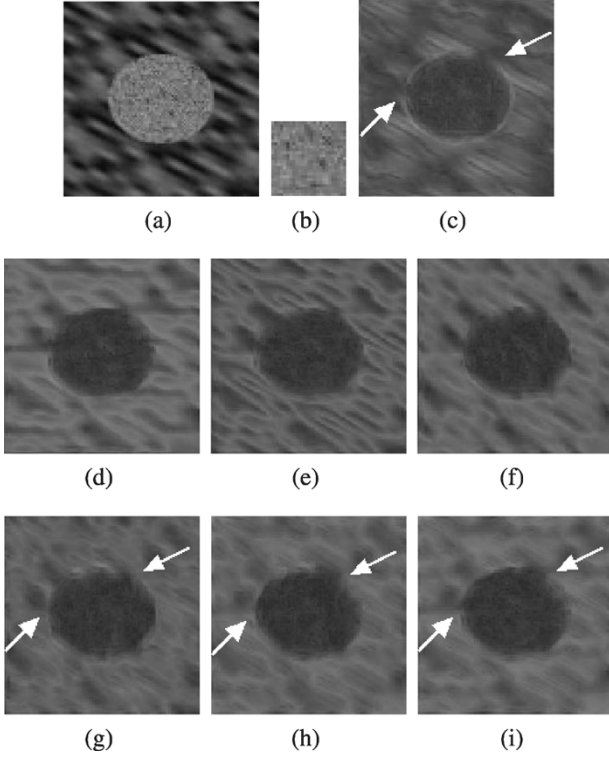
Fig. 3. Texture similarity maps of a synthetic texture image. In those maps, high intensity indicates high difference. (a) The original synthetic texture image. (b) The training texture pattern. (c) The similarity map of the full-region convolution scheme. (d)–(i) The similarity maps of each upper half-plane for six different training directions.

where $r^u_{n,\varphi=0}$ and $r^u_{n,\varphi=-(1/2)\pi}$ are the outputs of the symmetric and antisymmetric filters in the upper half plane defined as

$$r^u_{n,\varphi=0} = \int\int_\Omega \delta_u(x,y)I(x,y)G_{\varphi=0}(x,y)\mathrm{d}x\mathrm{d}y \qquad (5)$$

$$r^u_{n,\varphi=-\frac{1}{2}\pi} = \int\int_\Omega \delta_u(x,y)I(x,y)G_{\varphi=-\frac{1}{2}\pi}(x,y)\mathrm{d}x\mathrm{d}y. \quad (6)$$

The responds in the lower half plane are similarly defined as

$$r^l_{n,\varphi=0} = \int\int_\Omega \delta_l(x,y)I(x,y)G_{\varphi=0}(x,y)\mathrm{d}x\mathrm{d}y \qquad (7)$$

$$r^l_{n,\varphi=-\frac{1}{2}\pi} = \int\int_\Omega \delta_l(x,y)I(x,y)G_{\varphi=-\frac{1}{2}\pi}(x,y)\mathrm{d}x\mathrm{d}y \quad (8)$$

where

$$\delta_u(x,y) = \begin{cases} 1, & \text{if } (x,y) \in H_u \\ 0, & \text{otherwise} \end{cases} \qquad (9)$$

$$\delta_l(x,y) = \begin{cases} 1, & \text{if } (x,y) \in H_l \\ 0, & \text{otherwise} \end{cases}. \qquad (10)$$

After the convolution process, there are 12 half plane feature vectors R^u_n and R^l_n, $n = \{0, 1, \ldots, 5\}$ for each point in the training image. The dimension of each feature vector is 24.

Fig. 3 shows the texture similarity maps of a synthetic texture image. In these similarity maps, the high intensity indicates high difference between the corresponding region and the training texture patterns. Because the full-region convolution scheme considers the similarity of a given location P on the

integrated patch around it, there are several obvious gaps on the desired boundary in its similarity map [e.g., the regions highlighted by arrows in Fig. 3(c)]. This problem is overcome by the half-planed convolution scheme where the similarity of the given patch is measured separately on two sides. This leads to the disappearance of the gaps in some maps of the half-planed convolution scheme, as indicated in Fig. 3(g), (h). Fig. 4 is another case on a natural image. The full-region convolution result is shown in Fig. 4(c) in which the water region obtained a very high similarity indicated by the low intensity. This makes it hard to archive a good segmentation of the tiger hide without the background of water. Oppositely, the similarity of water is very low in Fig. 4(d), (e) which makes it possible for the segmenting curve to stop on the boundary of the water region. These two experiments show that the two-sided convolution strategy can increase the accuracy of texture feature extraction on texture's edge areas.

B. Two-Sided Texture Model

After texture features extraction, a good similarity measure is necessary for discriminating different textures. The windowed histogram model is one of the most common models in filtering theory. But it is hard to choose the proper size and the cell number of the histogram. Another common model is the finite mixture (FM) model which has a number of elegant features and is mathematically simple. It has been successfully employed to segment natural images [38] and medical images [39]. In our method, we model the training textures in the feature vector space as a mixture of K Gaussians. Parameter K should be chosen large enough to be able to capture the prominent texture variations present in the training dataset, but not too large that the texture model is too complex and unstable. In our experiments, we set $K = 3$ empirically and it can increase when the texture patterns become more complex.

For point P in the image, its texture similarity probability density in the upper half plane for the nth direction is defined as

$$U_n(\mathbf{x}_P|\Theta_n) = \sum_{i=1}^K \alpha_{ni} h_{ni}(\mathbf{x}_P|\theta_{ni}) \qquad (11)$$

where $\mathbf{x}_P = (x_1, x_2, \ldots, x_D)'$ is a feature vector and $D = 24$ in our case. α_{ni}'s are mixing weights $(\sum_{i=1}^K \alpha_{ni} = 1)$ and Θ_n refers to the collection of parameters $\{\alpha_{n1}, \ldots, \alpha_{nK}; \theta_{n1}, \ldots, \theta_{nK}\}$. Function h_{ni} is a multivariate Gaussian density defined in the following form

$$h_i(\mathbf{x}|\theta_i) = \frac{1}{(2\pi)^{\frac{D}{2}} \det \Sigma_i^{\frac{1}{2}}} \exp -\frac{1}{2}(\mathbf{x} - \mu_i)^T \Sigma_i^{-1}(\mathbf{x} - \mu_i). \qquad (12)$$

In order to determine the maximum-likelihood (ML) parameters of this model, we make use of the expectation-maximization (EM) algorithm proposed in [40]. Given an initialization of the K mean vectors $\mu_1, \ldots, \mu_K$ and K covariance matrices $\Sigma_1, \ldots, \Sigma_K$, the update equations are

$$\alpha_i^{t+1} = \frac{1}{N} \sum_{j=1}^N p(i|\mathbf{x}_j, \Theta^t) \qquad (13)$$

$$\mu_i^{t+1} = \frac{\sum_{j=1}^N \mathbf{x}_j p(i|\mathbf{x}_j, \Theta^t)}{\sum_{j=1}^N p(i|\mathbf{x}_j, \Theta^t)} \qquad (14)$$
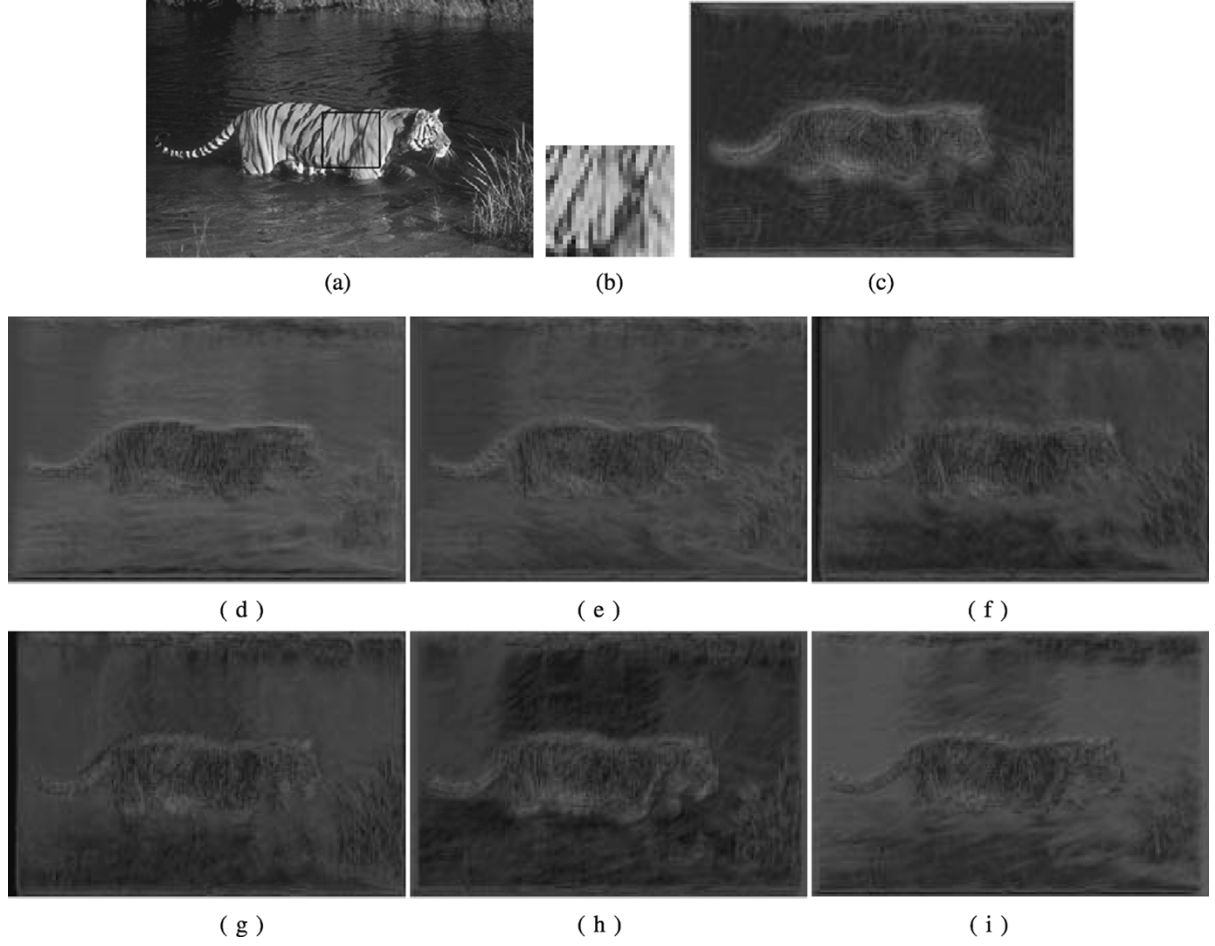
Fig. 4. Texture similarity maps of a real world image. In those maps, high intensity indicates high difference. (a) The original image. (b) The training texture pattern. (c) The similarity map of the full-region convolution scheme. (d–i) The similarity maps of each upper half-plane for six different training directions.

$$\Sigma_i^{t+1} = \frac{\sum_{j=1}^{N} p(i|\mathbf{x}_j, \Theta^t)\left(\mathbf{x}_j - \mu_i^{t+1}\right)\left(\mathbf{x}_j - \mu_i^{t+1}\right)^T}{\sum_{j=1}^{N} p(i|\mathbf{x}_j, \Theta^t)} \quad (15)$$

where N is the number of pixels in the training image and $p(i|\mathbf{x}_j, \Theta)$ is the probability that the ith Gaussian fits $\mathbf{x}_j$ with the given Θ

$$p(i|\mathbf{x}_j, \Theta) = \frac{\alpha_i h_i(\mathbf{x}_j|\theta_i)}{\sum_{r=1}^{K} \alpha_r h_r(\mathbf{x}_j|\theta_r)} \quad (16)$$

these update processes are repeated until the log likelihood

$$\log \mathcal{L}(\Theta_n|\mathcal{X}) = \log \prod_{i=1}^{N} U_n(\mathbf{x}_i|\Theta_n) \quad (17)$$

reaches the maximum value. Similarly, we define the probability density $L_n(\mathbf{x}|\Upsilon_n)$ for the lower half plane and estimate parameters Υ_n via the same procedure.

Once we get the ML parameters Θ^* and Υ^*, $U(\mathbf{x}|\Theta^*)$ and $L(\mathbf{x}|\Upsilon^*)$ are powerful models to describe the training texture for the given directions and half planes. In our case, there are totally 12 such mixture of half-planes Gaussians $U_n(\mathbf{x}|\Theta_n^*)$ and $L_n(\mathbf{x}|\Upsilon_n^*)$, where $n = (0, 1, \ldots, 5)$.

C. Shape Modeling

To model the shape prior, we apply a similar shape model construction method to that in [16] and [18]. Suppose the training set consists of B gray images $\{I_1, I_2, \ldots, I_B\}$. Let $\Psi : \Omega \to$ $\Re^+$ be a Lipchitz function that refers to a level set representation for a given shape C. This shape defines a region $\Re_C$ in the image plane Ω. Given these definitions the following shape representation is considered

$$\Psi_C(x, y) = \begin{cases} 0, & (x, y) \in C \\ D\left((x, y), C\right), & (x, y) \in \Re_C \\ -D\left((x, y), C\right), & (x, y) \in [\Omega - \Re_C] \end{cases}$$

where $D((x, y), C)$ refers to the minimum Euclidean distance between grid location (x, y) and shape C. The B images are aligned by minimizing the difference between any pair of images in the training set through the following similarity transformation

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = T_S \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s\cos\beta & -s\sin\beta & T_x \\ s\sin\beta & s\cos\beta & T_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

where β is the rotation angle, s is the scale factor, and (T_x, T_y) are the translation parameters in X and Y directions. Let $\{\Psi_1, \Psi_2, \ldots, \Psi_B\}$ denote the level set functions of the B aligned training images. The mean shape function is defined as

$$\bar{\Phi} = \frac{1}{B} \sum_{b=1}^{B} \Psi_b. \quad (18)$$

Then the mean-offset functions $S = [\Psi_1 - \bar{\Phi}, \Psi_2 - \bar{\Phi}, \ldots, \Psi_B - \bar{\Phi}]$ are used to capture the variabilities of the training shapes by

decomposing the covariance matrix $\Sigma = (1/B)SS^T$. In particular, the Singular Value Decomposition (SVD) is employed to factor $(1/B)S^T S$ as

$$V \Lambda V^T = \left(\frac{1}{B} \right) S^T S. \tag{19}$$

Next, the eigenvector matrix of the covariance matrix $(1/Y)SS^T$ is calculated as $A = SV$. The vectors of A are arranged back into the rectangular image grid to yield the principal modes or eigenshapes Φ_m. Finally, the shape prior model is constructed as

$$\Phi[W,T](x,y) = \bar{\Phi}(u,v) + \sum_{n=1}^{M} w_m \Phi_m(u,v) \tag{20}$$

with

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = T_S(s, \beta, T_x, T_y) \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{21}$$

where w_m's are mixing weights and M is the selected number of modes to consider. This shape model can deform to be an object's shape with different orientation, scale or location by changing parameters $W = \{w_m\}$ and $T = \{s, \beta, T_x, T_y\}$.

III. TEXTURE AND SHAPE PRIORS BASED SEGMENTATION

A. Energy Function

Let curve C be the zero level set of a Lipschitz function Φ such that

$$C = \left\{ (x,y) \in R^2 : \Phi(x,y) = 0 \right\}. \tag{22}$$

After the initial contour C_0 is manually positioned on the test image, the image is partitioned into two regions:

$$R^I = \left\{ (x,y) \in R^2 : \Phi(x,y) \geq 0 \right\} \tag{23}$$
$$R^o = \left\{ (x,y) \in R^2 : \Phi(x,y) < 0 \right\}. \tag{24}$$

We calculate the parameters

$$\Theta = \{W, T\}$$

of the shape model to update curve C. During the deformation process, for point P on the segmenting curve, the tangent line at P bisects the image plane into two half planes. Assume the slope of the tangent line is k and

$$\frac{(n-0.5)\pi}{6} \leq \arctan(k) < \frac{(n+0.5)\pi}{6}. \tag{25}$$

The half plane containing R^I is called the inside plane and the another plane is called the outside plane. Let $\mathbf{x}_o(P)$ denote the response vector obtained by applying the filter bank in the outside plane and $\mathbf{x}_I(P)$ denote the inside response vector as shown in Fig. 2. The inside and outside texture similarities $J_I(P)$ and $J_o(P)$ are defined as

$$\begin{cases} J_I(P) = U_n \left(\mathbf{x}_I(P) | \Theta_n^* \right) \\ J_o(P) = L_n \left(\mathbf{x}_o(P) | \Upsilon_n^* \right) \end{cases}. \tag{26}$$

Then, the energy function for our segmentation method is

$$E = -\frac{1}{2} f^2(\Theta) = -\frac{1}{2} \left[\mu_I \left(\sigma_I^2 - \sigma_o^2 \right) \right]^2 \tag{27}$$

where σ_I^2 and σ_o^2 are the sample variances in R^I and R^o, respectively. μ_I is the inside average texture similarity of the segmenting curve C. The objective of this energy function is to partition the image into two regions, the inside region with high average texture similarity and low texture variance, and the outside region with high texture variance. The two sample variances and the inside average texture similarity are defined as

$$\sigma_I^2 = \frac{Q_I}{L} - \left(\frac{S_I}{L} \right)^2 \tag{28}$$

$$\sigma_o^2 = \frac{Q_o}{L} - \left(\frac{S_o}{L} \right)^2 \tag{29}$$

$$\mu_I = \frac{S_I}{L} \tag{30}$$

where $L = \int_C dl$ and

$$S_I = \int \int_\Omega J_I \delta(\Phi) dA \tag{31}$$

$$S_o = \int \int_\Omega J_o \delta(\Phi) dA \tag{32}$$

$$Q_I = \int \int_\Omega J_I^2 \delta(\Phi) dA \tag{33}$$

$$Q_o = \int \int_\Omega J_o^2 \delta(\Phi) dA \tag{34}$$

and δ is the one-dimensional Dirac measure defined as

$$\delta(\Phi) = \begin{cases} 1, & \text{if } \Phi = 0 \\ 0, & \text{else} \end{cases}. \tag{35}$$

B. Optimization

Many iterative optimization algorithms for minimizing n-dimensional functions have been suggested such as the simplex, steepest descent and Newton–Raphson methods. In the case of a least squares function such as (27), the steps involved in the latter two methods may be simplified somewhat by defining a matrix $\mathbf{Z}$ given by

$$\mathbf{Z} = \begin{bmatrix} \dfrac{\partial f}{\partial \theta_1} & \dfrac{\partial f}{\partial \theta_2} & \cdots & \dfrac{\partial f}{\partial \theta_q} \end{bmatrix} \tag{36}$$

where q is the number of parameters in the function. By differentiating (27), the elements of the gradient vector necessary for the application of steepest descent are given by

$$\frac{\partial E}{\partial \theta_{\mathbf{k}}} = -f \frac{\partial f}{\partial \theta_k} \tag{37}$$

so that the gradient vector $\mathbf{g}$

$$\mathbf{g} = \begin{bmatrix} \frac{\partial E}{\partial \theta_1} \\ \vdots \\ \frac{\partial E}{\partial \theta_q} \end{bmatrix} = - \begin{bmatrix} \frac{\partial f}{\partial \theta_1} \\ \vdots \\ \frac{\partial f}{\partial \theta_q} \end{bmatrix} f \tag{38}$$

that is

$$\mathbf{g} = -\mathbf{Z}' f. \tag{39}$$

Now differentiating (37) with respect to θ_j we get

$$\frac{\partial^2 E}{\partial \theta_k \partial \theta_j} = -\frac{\partial f}{\partial \theta_j} \frac{\partial f}{\partial \theta_k} - f \frac{\partial^2 f}{\partial \theta_j \partial \theta_k}. \tag{40}$$
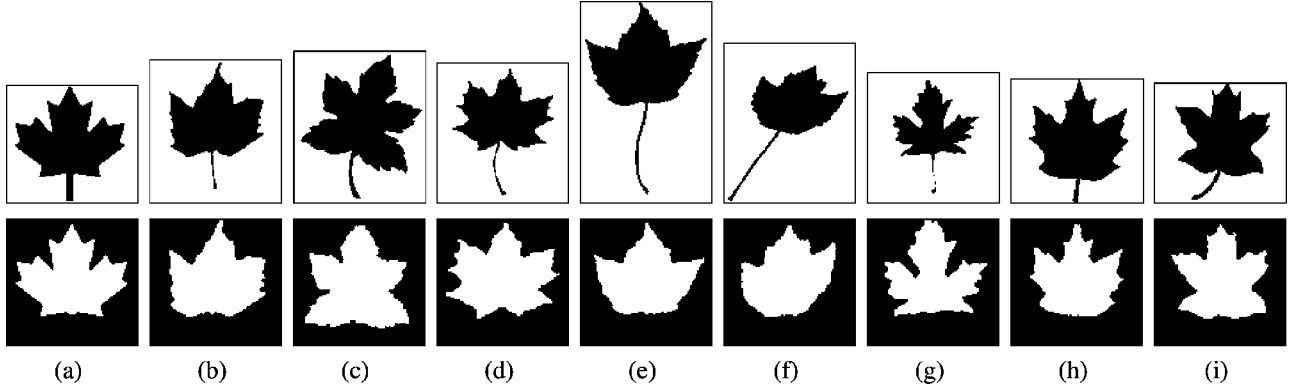
Fig. 5. The training set of maple leaf shapes. Top row shows the original maple leaf images and bottom row shows the aligned edition of the training shapes.

If we assume that the second term in (40) may be negligible we have

$$\frac{\partial^2 E}{\partial \theta_k \partial \theta_j} \approx -\frac{\partial f}{\partial \theta_j} \frac{\partial f}{\partial \theta_k}. \tag{41}$$

These are the elements of the Hessian matrix $\mathbf{H}$ which may, therefore, be written in the form

$$\mathbf{H} = \left[\frac{\partial^2 \mathbf{f}}{\partial \theta_{\mathbf{k}} \partial \theta_{\mathbf{j}}}\right] \approx -\mathbf{Z}'\mathbf{Z}. \tag{42}$$

Then, the update equation for our optimization is given by

$$\Theta_{i+1} = \Theta_i - [\tau \mathbf{I} + \mathbf{Z}'\mathbf{Z}]^{-1}\mathbf{Z}' \left[\mu_I \left(\sigma_I^2 - \sigma_o^2\right)\right] \tag{43}$$

where τ is a scalar which may be adjusted to control the sequence of iterations and $\mathbf{I}$ is the ($q \times q$) identity matrix. In general $\mathbf{Z}'\mathbf{Z}$ is positive definite so the procedure should converge. We use τ to control the iterative procedure at different situations. When the energy function is still far from the minimum, τ is set to a large value to take advantage of the reliable improvement given by steepest descent. After the searching is close to the minimum, τ is assigned a small value such that our method has the rapid convergence of the Newton–Raphson method.

The matrix $\mathbf{Z}$ for (27) can be derived as

$$\nabla_W f = \nabla_W \mu_I \left(\sigma_I^2 - \sigma_o^2\right) + \mu_I \left(\nabla_W \sigma_I^2 - \nabla_W \sigma_o^2\right) \tag{44}$$
$$\nabla_T f = \nabla_T \mu_I \left(\sigma_I^2 - \sigma_o^2\right) + \mu_I \left(\nabla_T \sigma_I^2 - \nabla_T \sigma_o^2\right) \tag{45}$$

where

$$\nabla_W \mu_I = \frac{\nabla_W S_I - \mu_I \nabla_W L}{L} \tag{46}$$

$$\nabla_T \mu_I = \frac{\nabla_T S_I - \mu_I \nabla_T L}{L} \tag{47}$$

$$\nabla_W \sigma_I^2 = \frac{\nabla_W Q_I L - Q_I \nabla_W L}{L^2} - \frac{2S_I(\nabla_W S_I L - S_I \nabla_W L)}{L^3}$$
$$\nabla_W \sigma_o^2 = \frac{\nabla_W Q_o L - Q_o \nabla_W L}{L^2} - \frac{2S_o(\nabla_W S_o L - S_o \nabla_W L)}{L^3}$$
$$\nabla_T \sigma_I^2 = \frac{\nabla_T Q_I L - Q_I \nabla_T L}{L^2} - \frac{2S_I(\nabla_T S_I L - S_I \nabla_T L)}{L^3}$$
$$\nabla_T \sigma_o^2 = \frac{\nabla_T Q_o L - Q_o \nabla_T L}{L^2} - \frac{2S_o(\nabla_T S_o A_o - S_o \nabla_T L)}{L^3}.$$

IV. Experiments and Discussions

A. Maple Leaf Segmentation

Before applying the algorithm to clinical data, we first applied it to natural images to identify its characteristics. The training dataset of this experiment consists of nine maple leaves of different shapes and sizes as shown in the top row of Fig. 5. To capture the shape prior knowledge, nine interactive segmentations of the maple leaves were aligned using the method described in Section II-C. The bottom row of Fig. 5 shows the aligned version of the shapes. Fig. 6(a) is the mean shape and Fig. 6(b) is the shape overlap map of the aligned shapes. Fig. 6(c)–(n) shows the effects of varying the first three modes with values in the range of ± 2 standard deviations. Although the shape representation introduced in Section II-C is not inconsistent framework since the convex linear combination of different distance functions is not a distance function, we notice that both the mean shape and primary modes in Fig. 6 still appear to be reasonable representative shapes of the training shapes, despite the shapes at two standard deviations away from the mean which are a little unfamiliar compared with the training shapes. This observation consists with the description of Leventon [16] and indicates that in a limited distance from the mean shape, this shape representation method can generally construct shapes consistent with the original training shapes. For comparison, we derived another shape model for the maple dataset using the ASM sahpe modeling approach in [9] and the resulting shapes (appearances) are shown in Fig. 7. For each leaf, 23 marks were annotated manually on the leaf boundary, to describe the corresponding parts across the training leaves. From Fig. 6 and 7, we can observe that the changes of shapes constructed by the two shape modeling approaches, at the same distance away from the mean shape, are comparable with each other and acceptable compared with the training set.

To segment the maple leaf from the test image in Fig. 8(a), two different texture patterns [Fig. 8(b) and (c)] were extracted from the maple leaf in the test image as shown in Fig. 8(d). Fig. 8(e) is the similarity map of Fig. 8(a) using the texture extractor introduced in [36]. The high intensity indicates high similarity between the local texture and the training patterns. We
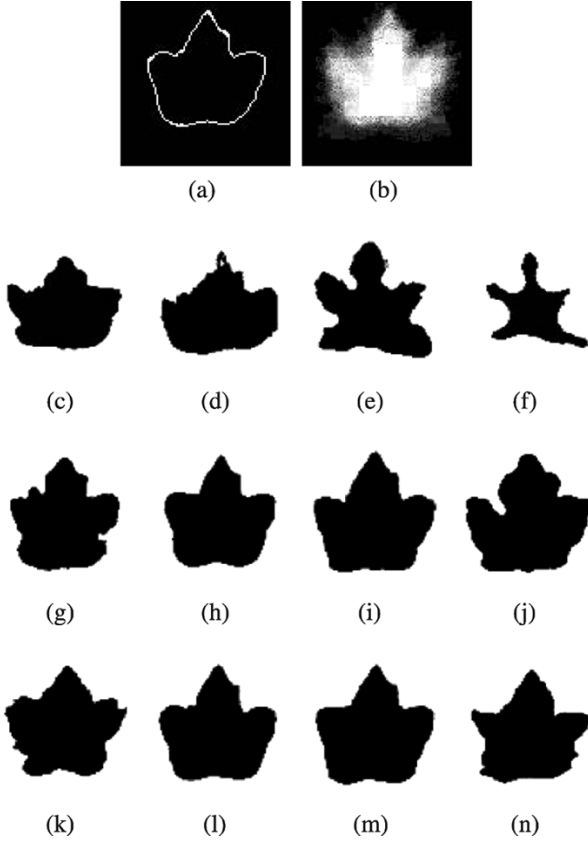
Fig. 6. Shape variability of the maple leaves. (a) The mean shape. (b) The shape overlap map. (c) $-2\sigma_1$ variation of the first principal mode. (d) $-1\sigma_1$ variation of the first principal mode. (e) $+1\sigma_1$ variation of the first principal mode. (f) $+2\sigma_1$ variation of the first principal mode. (g) $-2\sigma_2$ variation of the second principal mode. (h) $-1\sigma_2$ variation of the second principal mode. (i) $+1\sigma_2$ variation of the second principal mode. (j) $+2\sigma_2$ variation of the second principal mode. (k) $-2\sigma_3$ variation of the third principal mode. (l) $-1\sigma_3$ variation of the third principal mode. (m) $+1\sigma_3$ variation of the third principal mode. (n) $+2\sigma_3$ variation of the third principal mode.

observe that although most parts of the leaf have higher similarity than the background, the interior property of the leaf is not homogenous and the boundary is not complete. It is difficult to obtain a correct segmentation only using this texture feature map.

In our method, we applied the Gabor filter bank on the training patterns in 6 directions through the two-sided convolution strategy introduced in Section II-B. Their responses were modeled as 12 mixtures of half-planed Gaussians. The class number K is chosen as three for each mixture. After the training process, the mean shape [Fig. 6(a)] was applied as the initialization contour manually positioned on the test image as shown in Fig. 8(f). Then, the parameters of the model were updated to minimize the energy function defined in (27). Fig. 8(g) shows a middle step of the deformation process. The segmentation result is shown in Fig. 8(h) and our interactive segmentation is shown in Fig. 8(i). We can see that the boundary of the leaf is extracted correctly in our automatic segmentation result, in spite of the existence of the grassy textured background.

To obtain an accurate estimate for the generalization ability of the proposed method, we conduct a leave-one-out test on this maple dataset. We performed the segmentation on the test image [Fig. 8(a)] nine times. For each time, we removed one maple
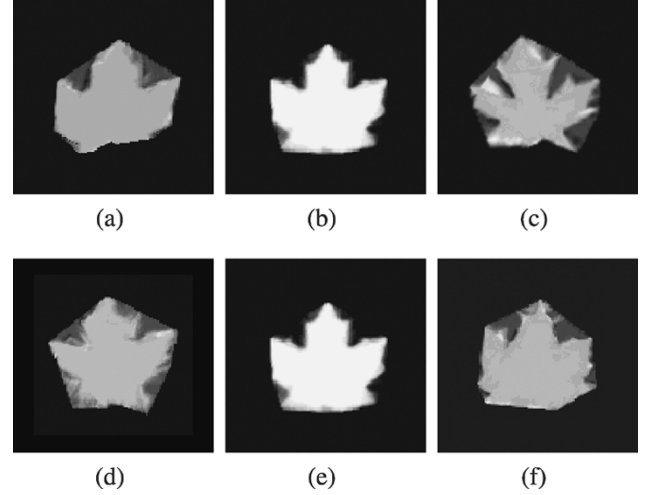

Fig. 7. Shape variability of the maple leaves using the active appearance model proposed in [9]. (a) The mean shape $-2\sigma_1$ variation of the first principal mode. (b) The mean shape. (c) The mean shape $+2\sigma_1$ variation of the first principal mode. (d) The mean shape $-2\sigma_2$ variation of the second principal mode. (e) The mean shape. (f) The mean shape $+2\sigma_2$ variation of the second principal mode.

sample from the training set and measured the mean pixel distance (MPD) and area error (AE) between the manually segmented and automatically generated boundary. MPD is defined as the mean of the minimum Euclidean distances of all points on the automatically generated curve to the manually segmented boundary. AE is calculated as the percentage of the automatically generated shape area that differs from the manually segmented area.

Table I shows the estimate result where all leave-one-out segmented curves had a mean distance of 1.36–2.43 pixels (average 1.744 pixels) and an AE of 1.2%–3.1% (average 2.08%). The mean distance of the leave-all-in experiment is 1.42 pixels and the corresponding AE is 1.8%. We notice that besides three cases [Fig. 5(c), (d), (g)], where results are unsatisfying due to large deviations, most of the training maple images could be segmented in a leave-one-out test with a good quality. On the other hand, the mean distance and AE of the leave-all-in experiment are all lower than those of the leave-one-out tests. This indicates that increasing the number of training data can improve the segmentation results.

B. Left Kidney Segmentation

To segment kidney in US images, we have tried to derive the shape prior model from US images directly. However, generally the structures of kidney are not complete or prominent in US images and it is very hard to segment these structures correctly due to the speckle noise. Therefore, we chose to derive the shape prior from 20 two-dimensional (2-D) left kidney CT images of different patients. The interactively segmented kidneys formed the shape training set. These kidney images were sliced from the same spacial position from each patient so that the shapes of these left kidneys were similar with some local differences. We constructed the shape prior model through the same procedure as done in the maple experiment. Fig. 9 shows the aligned version of the training left kidney images and Fig. 10(a) is the mean shape. Fig. 10(c)–(n) demonstrate the effects of varying
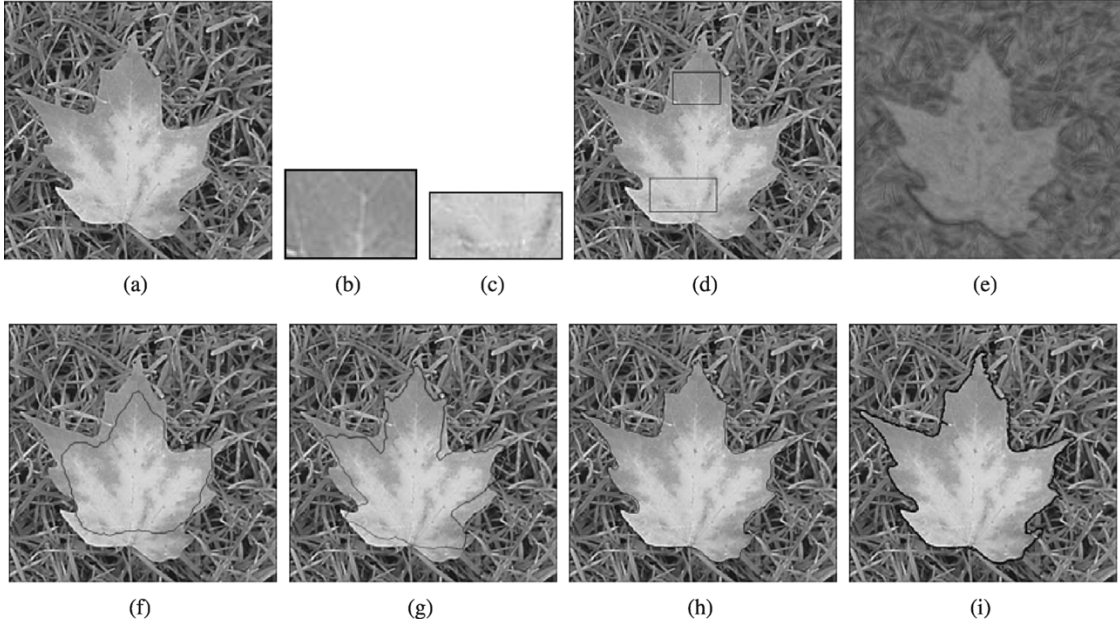
Fig. 8. Experimental results on maple leaf segmentation. (a) The test image. (b) The first training texture pattern. (c) The second training texture pattern. (d) Demonstration of the positions where the two training texture patterns are selected. (e) The texture similarity map using the texture extractor proposed in [36]. High intensity indicates high similarity. (f) The initialization of the segmentation process where the initial curve (black) is the mean shape of the training data. (g) One step of the deformation of the segmenting curve. (h) The segmentation result of the proposed method. (i) The manual segmentation result.
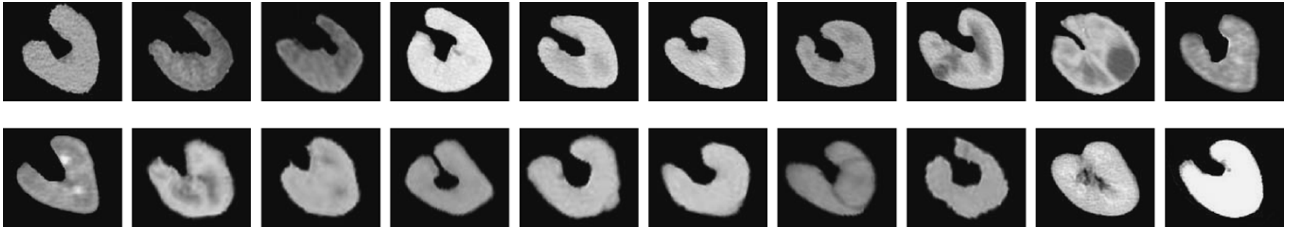

Fig. 9. The aligned version of the kidney training database.

the first three modes with values in the range of ± 2 standard deviations and Fig. 11 are shapes derived using the ASM model with 20 marks for each training kidney. We can see both the two shape modeling approaches have constructed reasonable representative shapes of the training data, with limited distances from the mean shape.

After the shape prior model of left kidney was constructed, we applied it to segment left kidneys in other patients' US images. The first test image [Fig. 12(a)] is an US image with a left kidney which has a 2.5×1.8 ectogenic mass in the longitudinal view. Fig. 12(c) is the texture similarity map using the texture feature extractor in [36]. It shows that it is difficult to segment such US images using only texture features. In our segmentation phase, the mean shape was put on the test image as shown in Fig. 12(d) and the texture similarity of regions inside and outside the curve were measured. The result of our automatic method is shown in Fig. 12(f). Compared with the manual segmentation result by physicians [Fig. 12(g)], the automatic segmentation converges to a satisfactory result and comes close to the manual expert segmentation.

The second experiment on real medical US images is shown in Fig. 13. Three segmentation results, including the manual result by experts [Fig. 13(b)], the result using the basic level set segmentation [Fig. 13(c)] and our algorithm segmentation [Fig. 13(d)], are presented to assess the performance of the

two segmentation algorithms. From Fig. 13(c) we observe that, at locations of missing or fuzzy boundaries, the level set scheme failed to maintain global smoothness and leaked through the gaps. This is because the propagation force of the level set method is only opposed by strong gradient magnitude at image discontinuities. Thus, at the weak tissue-tissue boundary in the noisy kidney US image, this method cannot yield correct convergence and the evolution has to be halted manually. Oppositely, using the shape prior model, our method can guarantee a reasonable representative shape of the training shapes. Meanwhile, by minimizing the novel texture-based energy function, the proposed deformable model can converge to the kidney's boundary although the boundary did not lie on points with high intensity gradient.

To evaluate our method quantitatively, we compared our method's results with manually segmented left kidneys in terms of MPD and AE. Six visual results are shown in Fig. 14. These images were captured in a HDI5000 manufactured by ATL Ultrasound. The pixel sizes are 0.218 mm and 0.225 mm in X direction and Y direction, respectively. In each of these images, the white contour is the manual segmentation result and the black contour is the result of our method. From the values in Table II, we observe that the MPD's of the six experiments vary from 0.501 mm to 1.692 mm with average of 1.0182 mm. The AE's are in the range from 6.1% to 1.8%
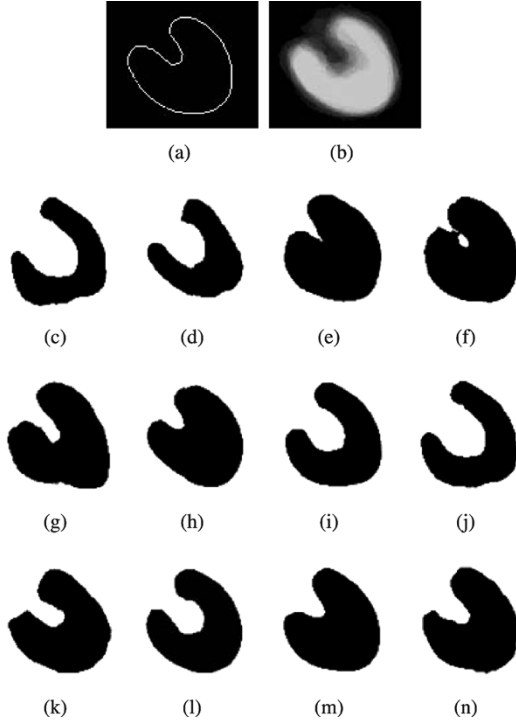
Fig. 10. Shape variability of the kidney database. (a) The mean shape. (b) The shape overlap map. (c) $-2\sigma_1$ variation of the first principal mode. (d) $-1\sigma_1$ variation of the first principal mode. (e) $+1\sigma_1$ variation of the first principal mode. (f) $+2\sigma_1$ variation of the first principal mode. (g) $-2\sigma_2$ variation of the second principal mode. (h) $-1\sigma_2$ variation of the second principal mode. (i) $+1\sigma_2$ variation of the second principal mode. (j) $+2\sigma_2$ variation of the second principal mode. (k) $-2\sigma_3$ variation of the third principal mode. (l) $-1\sigma_3$ variation of the third principal mode. (m) $+1\sigma_3$ variation of the third principal mode. (n) $+2\sigma_3$ variation of the third principal mode.

TABLE I
THE RESULT OF LEAVE-ONE-OUT ANALYSIS OF MAPLE EXPERIMENT

Removed Image	MPD (Pixels)	AE (%)
Fig. 5(a)	1.52	1.2
Fig. 5(b)	1.43	1.4
Fig. 5(c)	2.43	3.1
Fig. 5(d)	2.17	2.7
Fig. 5(e)	1.36	1.8
Fig. 5(f)	1.63	1.7
Fig. 5(g)	2.23	2.9
Fig. 5(h)	1.45	1.6
Fig. 5(i)	1.48	2.3
Mean	1.744	2.078
Std.	0.411	0.692
Leave-All-In	1.42	1.8

with average of 3.61%. This evaluation shows that our method can segment kidney in US images with a quality that is close to satisfactory for image-guided diagnosis. Our experiments were performed on a 1.7-GHz Pentium IV PC. The time required for segmentation depends on the input images and the initialization contours. After the initialization, the average time for curve deformation is 40.3 s, including the time for the two-sided texture feature extraction and that for parameters optimization.

V. CONCLUSION

We have presented a novel segmentation method based on both texture and shape priors. One major novelty of the paper is
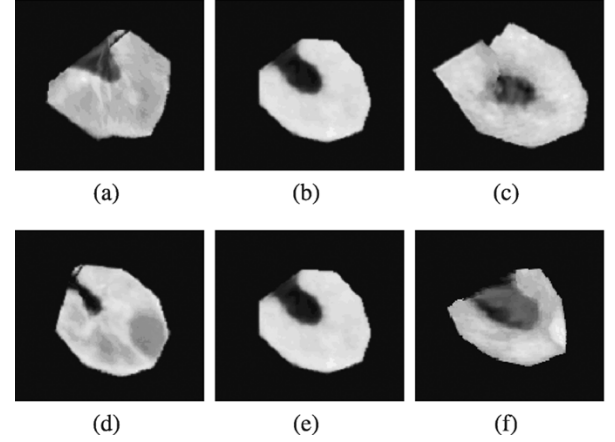

Fig. 11. Shape variability of the kidney database using the active appearance model proposed in [9]. (a) The mean shape $-2\sigma_1$ variation of the first principal mode. (b) The mean shape. (c) The mean shape $+2\sigma_1$ variation of the first principal mode. (d) The mean shape $-2\sigma_2$ variation of the second principal mode. (e) The mean shape. (f) The mean shape $+2\sigma_2$ variation of the second principal mode.

the proposed segmentation framework which combines the texture priors into the shape prior segmentation scheme proposed in [16]. This provides our method with the ability to deal with textured objects with incomplete boundaries. Another important contribution of this paper is the proposed two-sided convolution strategy through which the texture priors are modeled as a set of mixtures of half-planed Gaussians. Texture features are extracted by applying a bank of Gabor filters on test images through a two-sided convolution strategy. Texture similarity analysis on both simulated and real images shows the proposed texture feature is more robust to deform the segmenting curve than that in [36]. Segmentation is implemented by updating the parameters of an implicit representation of the segmenting curve to minimize a texture-based energy function. The goal of the novel energy function is to partition the test image into two regions, the region inside the curve with a high texture similarity and a low variance of texture measure, and the outside region with a high variance. The optimization is implemented via an improved gradient descent algorithm which makes the searching method faster and more robust. Compared with the segmentation method in [29], our method has an obvious advantage that it does not depend on the training for the background in test images. Moreover, the shape constraint combined in our framework makes the segmentation more correct and reasonable for certain objects of interest.

We evaluated the proposed method via experiments on both natural images and US images compared with manual raters and other image segmentation methods (e.g., [17] and [36]). The experiments on the maple leaf dataset shows the valid and efficiency of the proposed segmentation framework. Additionally, the experimental results on the medical US data are promising and show that our approach can segment kidney in US images accurately and efficiently.

For all experiments in the paper, the initial segmenting curves were positioned manually. Ideally a fully automated technique would be developed such that the prior model can be initialized on the test medical image in some sense optimal. However, it is a challenging perspective because it is hard to define what is optimal and the speckle patterns in US images are too complex.
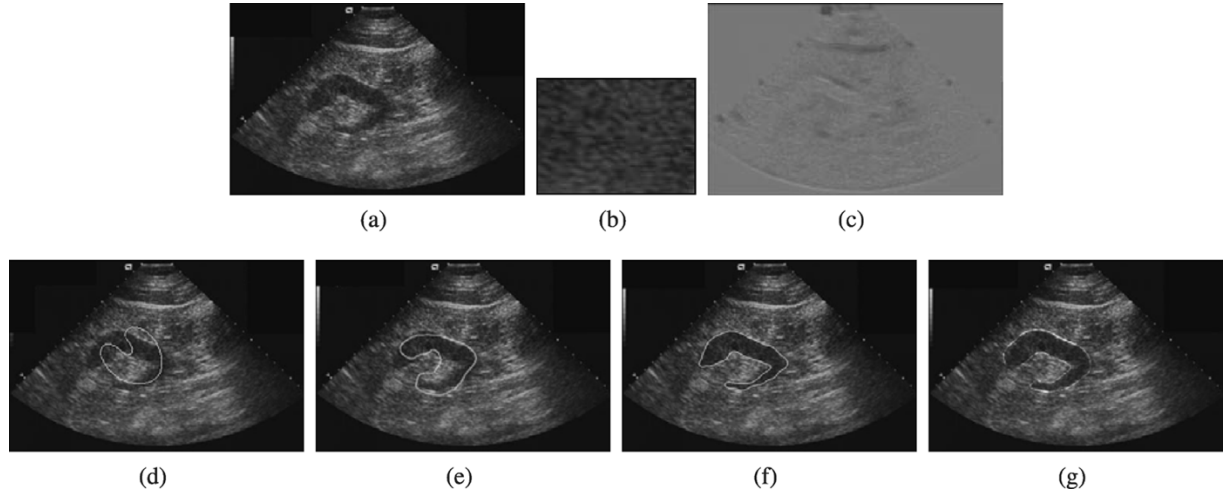
Fig. 12.   Experimental results on an US image. (a) The test image. (b) The training texture pattern. (c) The texture similarity map using the texture feature extractor in [36]. (d) The initialization of the segmentation process where the initial curve (white) is the mean shape of the training data. (e) One step of the evolution of the segmenting curve. (f) The segmentation result of the proposed method. (g) The manual segmentation result.
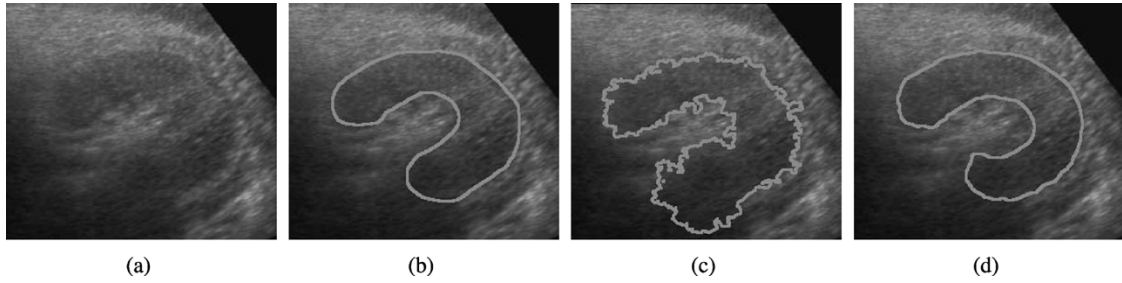

Fig. 13.   Experimental results on an US kidney image. (a) The test image. (b) The manual segmentation by experts. (c) The segmentation result using the basic level set segmentation method proposed in [17]. (d) The segmentation result using the proposed method.
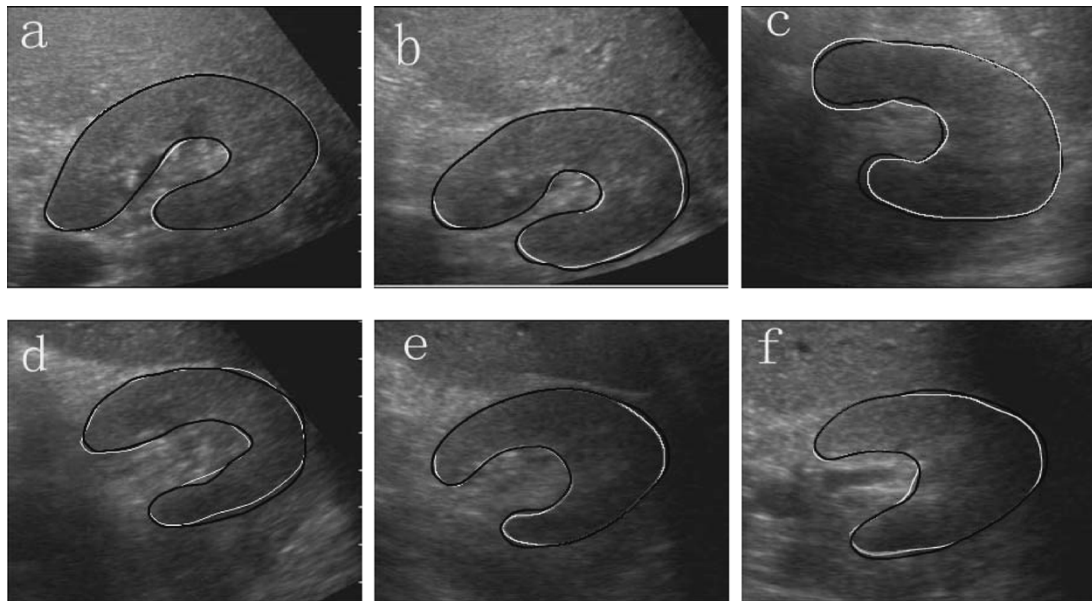

Fig. 14.   Comparison of segmentations of the proposed method and manual segmentations. The bright contour in each image is the manual segmentation and the black contour is the segmentation of our method.

A possible solution for this limit is to use a larger respective region of the Gabor filter bank because the texture similarity is measured in the nearby region beside the segmenting curve. However, this will result in an increase of the computation com-plex of both the training and segmenting procedures. An alternative method is using a multiresolution deformation framework. That is in the training procedure, a resolution pyramid of the prior texture model is constructed. Then in the segmentation

TABLE II
EVALUATION OF THE PROPOSED METHOD.
PIXEL SIZE IS 0.218 mm $\times$ 0.225 mm

Image	MPD		Object Size		AE
	Pixels	mm	Pixels	mm^2	%
Fig. 14(a)	1.6	0.501	18157	890.6	1.8
Fig. 14(b)	1.9	0.595	16622	815.3	2.1
Fig. 14(c)	5.2	1.630	19926	977.4	5.2
Fig. 14(d)	5.4	1.692	13671	670.6	6.1
Fig. 14(e)	3.3	1.034	14663	719.2	2.3
Fig. 14(f)	2.1	0.658	16845	826.2	4.2
Average	3.25	1.0182	16647.3	816.55	3.61
Std.	1.69	0.5295	2274.4	111.56	1.80

procedure, the texture similarity is measured in different resolutions. We believe this technique can enable our methodology to adaptively focus on particular texture features in the various deformation stages.

On the other hand, although the proposed method in this paper stands on its performance in practice, the shape modeling technique employed in this method is not an inconsistent framework which may introduce shape degeneracy in the segmentation approach when the shapes are derived at a long distance away from the mean shape. This deformation characteristic implies that a good initial placement of the segmentation curve is indispensable for a good segmentation. Thus, we are very interested in developing a hybrid shape modeling technique using both the distance function and the landmarks in the future, to combine the attractive properties of these two shape modeling methods.

To further improve our method for the particular application of medical US image segmentation, we are interested in investigating a more stable texture prior model by including a local speckle analysis in our texture feature extractor. Since B-mode US is most sensitive to the surfaces of structures normal to the beam, scattering shows obvious view-dependent property and its appearances in the image for a certain patch of tissues, are variant to the relative positions of the US transducer. The performance of our algorithm with a removal of the view-dependent speckle pattern will be tested in the future. Currently, our work is focused on kidney US images only. The performance of this method on other applications such as cardiac segmentation will be studied in the future. In this paper we have only dealt with 2-D images. However, this method can be easily expanded to a three–dimensional one using a procedure similar to that in [41]. This system is also under development in our laboratory.

ACKNOWLEDGMENT

REFERENCES

[1] B. Levienaise-Obadia and A. Gee, "Adaptive segmentation of ultrasound images," *Image Vis. Computing*, vol. 17, no. 8, pp. 583–588, Jun. 1999.

[2] C. J. Bouma, K. J. Zuiderveld, and M. A. Viergever, "Quantitative evaluation of (semi)-automated segmentation of intravascular ultrasound subtraction images," presented at the *11th Symp. Echocardiology*, Rotterdam, The Netherlands, 1995.

[3] S. P. abd, V. Chalana, and Y. Kim, "Interactive automatic fetal head measurements from ultrasound images using multimedia computer technology," *Ultrasound Med. Biol.*, vol. 23, p. 665C673, 1997.

[4] X. Hao, C. Bruce, C. Pislaru, and J. F. Greenleaf, "A novel region growing method for segmenting ultrasound images," in *Proc. IEEE Ultrason. Symp.*, vol. 2, 2000, pp. 1717–1720.

[5] C. Davatzikos and R. Bryan, "Using a deformable surface model to obtain a shape representation of the cortex," *IEEE Trans. Med. Imag.*, vol. 15, no. 6, pp. 785–795, Dec. 1996.

[6] R. Muzzolini, Y. H. Yang, and R. Pierson, "Multiresolution texture segmentation with application to diagnostic ultrasound images," *IEEE Trans. Med. Imag.*, vol. 12, no. 1, pp. 108–123, Mar. 1993.

[7] V. Chalana, D. Linker, D. Haynor, and Y. Kim, "A multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Trans. Med. Imag.*, vol. 15, no. 3, pp. 290–298, Jun. 1996.

[8] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models-their training and application," *Comput. Vis. Image Understanding*, vol. 61, pp. 38–59, 1995.

[9] T. Cootes, C. Beeston, G. Edwards, and C. Taylor, "A unified framework for atlas matching using active appearance models," in *Proc. Int. Conf. on Image Processing in Medical Imaging*, 1999, pp. 322–333.

[10] R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor, "A minimum description length approach to statistical shape modeling," *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 525–537, May 2002.

[11] H. H. Thodberg and H. Olafsdottir, "Adding curvature to minimum description length shape models," in *Proc. Br. Machine Vision Conf.*, Sep. 2003.

[12] Y. Chen, H. Thiruvenkadam, and D. Wilson *et al.*, "On the incorporation of shape priors into geometric active contours," in *Proc. IEEE Int. Workshop Variational and Level Set Methods*, Vancouver, BC, Canada, Jul. 2001, pp. 145–152.

[13] N. Paragios and M. Rousson, "Matching distance functions: a shape-to-area variational approach for global-to-local registration," in *Proc. European Conf. on Computer Vision*, Copenhagen, Denmark, May 2002, pp. 78–92.

[14] M. Rousson and N. Paragios, "Shape prior for level set representation," in *Proc. IEEE Int. Conf. on Image Processing*, Kobe, Japan, Oct. 1999, pp. 188–192.

[15] D. Shen, Y. Zhang, and C. Davatzikos, "Segmentation of prostate boundaries from ultrasound images using statistical shape model," *IEEE Trans. Med. Imag.*, vol. 22, no. 4, pp. 539–551, Apr. 2003.

[16] M. Leventon, E. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proc. IEEE Comput. Vis. Pattern Recogn.*, Hilton Head, S.C., Jun. 2000, pp. 316–322.

[17] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations," *Int. J. Comp. Phys.*, vol. 79, pp. 12–49, 1988.

[18] A. Tsai, A. Yezzi, W. Wells, and A. Willsky, "Model-based curve evolution technique for image segmentation," in *Proc. IEEE Comput. Vis. Pattern Recogn.*, vol. I, Kauai, HI, Dec. 2001, pp. 463–468.

[19] T. Chan and L. Vese, "A level set algorithm for minimizing the Mumford-Shah functional in image processing," in *Proc. IEEE Int. Workshop Variational and Level Set Methods*, Vancouver, BC, Canada, Jul. 2001, pp. 161–168.

[20] A. Yezzi, A. Tsai, and A. Willsky, "A statistical approach to snakes for bimodal and trimodal imagery," in *Proc. IEEE Int. Conf. Computer Vision*, vol. 2, Kerkyra, Greece, Sep. 1999, pp. 898–903.

[21] J.-P. Pons, G. Hermosillo, R. Keriven, and O. Faugeras, "How to deal with point correspondences and tangential velocities in the level set framework," in *Proc. IEEE Int. Conf. Computer Vision*, Nice, France, 2003, pp. 894–899.

[22] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active shape models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1987.

[23] S. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 9, pp. 884–900, Sep. 1996.

[24] A. Chakraborty, H. Staib, and J. Duncan, "Deformable boundary finding in medical images by integrating gradient and region information," *IEEE Trans. Med. Imag.*, vol. 15, no. 6, pp. 859–870, Dec. 1996.

[25] D. Mumford and J. Shah, "Boundary detection by minimizing functionals," in *Proc. IEEE Comput. Vis. Pattern Recogn.*, San Francisco, CA, Jun. 1985, pp. 22–26.

[26] T. Chan and L. Vese, "An active contour model without edges," in *Proc. Int. Conf. Scale-Space Theories in Computer Vision*, Kerkyra, Greece, Sep. 1999, pp. 141–151.

[27] C. Samson, L. B. Feraud, G. Aubert, and J. Zerubia, "A level set model for image classification," in *Proc. Int. Conf. Scale-Space Theories in Computer Vision*, Corfu, Greece, Sep. 1999, pp. 306–317.

[28] A. Tsai, A. Yezzi, and A. Willsky, "A curve evolution approach to smoothing and segmentation using the Mumford-Shah functional," in *Proc. IEEE Comput. Vis. Pattern Recogn.*, Hilton Head, SC, Jun. 2000, pp. 119–124.

[29] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *Int. J. Comput. Vis.*, vol. 46, no. 3, pp. 222–247, 2002.

[30] ——, "Geodesic active regions for supervised texture segmentation," in *Proc. IEEE Int. Conf. Computer Vision*, Kerkyra, Greece, Sep. 1999, pp. 926–932.

[31] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.

[32] J. Puzicha, J. M. Buhmann, Y. Rubner, and C. Tomasi, "Empirical evaluation of dissimilarity measures for color and texture," in *Proc. IEEE Int. Conf. Computer Vision*, Kerkyra, Greece, Sep. 1999, pp. 1165–1173.

[33] C. H. Chen, L. F. Pau, and P. Wang, *Handbook of Pattern Recognition and Computer Vision*, 2nd ed, Singapore: World Scientific, 1998.

[34] P. Rivaz and N. Kingsbury, "Fast segmentation using level set curves of complex wavelet surfaces," in *Proc. IEEE Int. Conf. on Image Processing*, Vancouver, BC, Canada, Sep. 2000, pp. 592–595.

[35] J. G. Daugman, "Two-dimentional spectral analysis of cortical receptive field profiles," *Vis. Res.*, vol. 20, pp. 847–856, 1980.

[36] P. Kruizinga and N. Petkov, *Nonlinear Operator for Oriented Texture*, vol. 8, no. 10, pp. 1395–1407, Oct. 1999.

[37] R. L. Devalois, D. G. Albrecht, and L. G. Thorell, "Spatial frequency selectivity of cells in macaque visual cortex," *Vis. Res.*, vol. 22, pp. 545–559, 1982.

[38] S. Belongie and C. Carson, "Color- and texture-based image segmentation using em and its application to content-based image retrieval," in *Proc. IEEE Int. Conf. Computer Vision*, Bombay, India, Jan. 1998, pp. 675–682.

[39] Y. Zhang, B. Michael, and S. Stephen, "Segmentation of brain images through a hidden markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imag.*, vol. 20, no. 1, pp. 45–57, Jan. 2001.

[40] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc.*, vol. 39, pp. 1–38, 1977.

[41] D. G. Gobbi and T. M. Peters, "Interactive intra-operative 3D ultrasound reconstruction and visualization," in *Proc. Medical Image Computing and Computer Assisted Intervention Conf (MICCAI'02)*, Tokyo, Japan, Sep. 2002, pp. 156–163.