Published in final edited form as:

IEEE Trans Med Imaging. 2004 August ; 23(8): 940–948. doi:10.1109/TMI.2004.830802.

Neighbor-Constrained Segmentation With Level Set Based 3-D Deformable Models

Jing Yang[Student Member, IEEE],

Departments of Electrical Engineering and Diagnostic Radiology, Yale University, P.O. Box 208042, New Haven, CT 06520-8042 USA

Lawrence H. Staib[Member, IEEE], and

Departments of Electrical Engineering and Diagnostic Radiology, Yale University, New Haven, CT 06520-8042 USA

James S. Duncan[Fellow, IEEE]

Departments of Electrical Engineering and Diagnostic Radiology, Yale University, New Haven, CT 06520-8042 USA

Jing Yang: j.yang@yale.edu

Abstract

A novel method for the segmentation of multiple objects from three-dimensional (3-D) medical images using interobject constraints is presented. Our method is motivated by the observation that neighboring structures have consistent locations and shapes that provide configurations and context that aid in segmentation. We define a maximum *a posteriori* (MAP) estimation framework using the constraining information provided by neighboring objects to segment several objects simultaneously. We introduce a representation for the joint density function of the neighbor objects, and define joint probability distributions over the variations of the neighboring shape and position relationships of a set of training images. In order to estimate the MAP shapes of the objects, we formulate the model in terms of level set functions, and compute the associated Euler–Lagrange equations. The contours evolve both according to the neighbor prior information and the image gray level information. This method is useful in situations where there is limited interobject information as opposed to robust global atlases. In addition, we compare our level set representation of the object shape to the point distribution model. Results and validation from experiments on synthetic data and medical imagery in two-dimensional and 3-D are demonstrated.

Index Terms

Deformable models; level set; neighbor-constrained segmentation; neighbor prior model; shape prior model; 3-D segmentation

I. Introduction

Segmentation and quantitative analysis of structures in an image have tremendous applications in medical imaging. In order to fully realize the value of medical imaging in both clinical and research settings, information about anatomical structures must be extracted and quantified with accuracy, efficiency, and reproducibility.

Correspondence to: Jing Yang, j.yang@yale.edu.

This paper is an extended version of the authors' IPMI paper [12]. The Guest Editors responsible for coordinating the review of this paper and recommending its publication were C. J. Taylor and J. A. Noble.

Snakes or active contour models (ACMs) [1] have been widely used for segmenting nonrigid objects in a wide range of applications. ACMs are energy minimizing parametric contours with smoothness constraints deformed according to the image data. Unlike level set implementations [2], the direct implementation of this energy model is not capable of handling topological changes of the evolving contour without explicit discrete pixel manipulations. Typically, ACMs detect objects with edges defined by the gradient.

Recently, methods using level set methods and new energy terms have been reported to increase the capture range of deformable models. For example, Chan and Vese [3] have proposed an active contour model that can detect objects whose boundaries are not necessarily defined by gray level gradients.

The incorporation of more specific prior information into deformable models has received some attention. Cootes *et al.* [4] find corresponding points across a set of training images and construct a statistical model of shape variation from the point positions. Staib and Duncan [5] incorporate global shape information into the segmentation process by using an elliptic Fourier decomposition of the boundary and placing a Gaussian prior on the Fourier coefficients. Zeng *et al.* [6] develop a coupled surfaces algorithm to segment the cortex by using a thickness prior constraint. Leventon *et al.* [7] extend Caselles' [8] geodesic active contours by incorporating shape information into the evolution process.

In many cases, objects to be detected have one or more neighboring structures whose location and shape provide information about the local geometry that can aid in the delineation. The relative shape and arrangements among these neighbors can be modeled based on statistical information from a training set [11], [12]. Though applicable in many domains, these models are particularly useful for medical applications. In anatomy, neighboring structures provide a strong constraint on the relative position and shape of a structure. Without a prior model to constrain the segmentation, algorithms often fail due to the difficult challenges of poor image contrast, noise, and missing or diffuse boundaries. Segmentation can be made easier if suitable neighbor prior models are available.

Our model is based on a maximum *a posteriori* (MAP) framework using a neighbor prior constraint. We introduce a representation for the joint density function of the neighbor objects and define the corresponding probability distributions using level sets. In order to relate this effort to other work in the field, we compare this level set distribution model (LSDM) by comparing it with the traditional point distribution model (PDM) [4] using the Chi-square test. Formulating the segmentation as a MAP estimation of the shapes of the objects and modeling in terms of level set functions, we compute the associated Euler-Lagrange equations. The contours evolve both according to the neighbor prior information and the image gray level information. The neighboring objects can be automatically detected simultaneously.

II. Description of the Model

A. Map Framework With Neighbor Prior

A probabilistic formulation is a powerful approach to deformable models. Deformable models can be fit to the image data by finding the model shape parameters that maximize the posterior probability. Consider an image I that has M objects of interest; a MAP framework can be used to realize image segmentation combining neighbor prior information and image information

$$\begin{aligned}
\widehat{\mathcal{S}_i} &= \arg \max_{\mathcal{S}_i} p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_i, \dots, \mathcal{S}_M | I) \\
&= \arg \max_{\mathcal{S}_i} p(I | \mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) \\
&\quad i=1, 2, \dots, M
\end{aligned} \tag{1}$$

where $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M$ are the evolving surfaces of all the objects of interest.

$p(I | \mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M)$ is the probability of producing an image I given $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M$. In three-dimensions, assuming gray level homogeneity within each object, we use the imaging model in (2), shown at the bottom of the page, [3] where c_{1i} and σ_{1i} are the average and variance of I inside $\mathcal{S}_i$, c_{2i} and σ_{2i} are the average and variance of I outside $\mathcal{S}_i$ but also inside a certain domain Ω_i that contains $\mathcal{S}_i$.

$$\begin{aligned}
&p(I | \mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) \\
&= \prod_{i=1}^M \left\{ \prod_{(x,y,z) \text{ inside } (\mathcal{S}_i)} \exp[-(I(x,y,z) - c_{1i})^2 / (2\sigma_{1i}^2)] \prod_{(x,y,z) \text{ outside } (\mathcal{S}_i), \text{ inside } (\Omega_i)} \exp[-(I(x,y,z) - c_{2i})^2 / (2\sigma_{2i}^2)] \right\}
\end{aligned}$$

$p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M)$ is the joint density function of all the M objects. It contains the neighbor prior information such as the relative position and shape among the objects.

B. Neighbor Priors

First we notice that by the chain rule, we have

$$\begin{aligned}
&p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) \\
&= p(\mathcal{S}_M | \mathcal{S}_{M-1}, \mathcal{S}_{M-2}, \dots, \mathcal{S}_1) \\
&\times p(\mathcal{S}_{M-1} | \mathcal{S}_{M-2}, \mathcal{S}_{M-3}, \dots, \mathcal{S}_1), \dots \\
&p(\mathcal{S}_3 | \mathcal{S}_2, \mathcal{S}_1) p(\mathcal{S}_2 | \mathcal{S}_1) p(\mathcal{S}_1).
\end{aligned} \tag{3}$$

Let us now define a binary matrix $R_{M \times M}$ where each element r_{ij} describes the independence of $\mathcal{S}_i$ and $\mathcal{S}_j$. r_{ij} has value zero when $\mathcal{S}_i$ and $\mathcal{S}_j$ are independent and has value one otherwise. Obviously, the more ones there are in R , the more neighbor prior information is incorporated in the MAP segmentation model.

When

$$R = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ & & \dots & \\ 1 & 1 & \dots & 1 \end{bmatrix} \tag{4}$$

all of the M objects are related to each other. The shape self prior, as well as the neighbor relation prior of all the neighbors, are included. In this case, (3) cannot be simplified and finding the joint density function of all the M objects can be complicated. This case incorporates the most neighbor prior information (full order prior) but with a corresponding loss of computational efficiency.

If all the M objects are independent to each other, i.e.,

$$R = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (5)$$

then (3) can be simplified to

$$p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) = p(\mathcal{S}_M) p(\mathcal{S}_{M-1}) \dots p(\mathcal{S}_2) p(\mathcal{S}_1) \quad (6)$$

No neighboring information is included here since all the objects are independent to each other. The only prior information contained in the MAP model for each object is the object's self shape prior $p(\mathcal{S}_i)$, which we designate the first-order prior. In this case, (3) can be most simplified; we can achieve good computational efficiency but with no neighbor prior information. Each object in the image can be segmented independently according to its shape prior and image gray level information. This formulation corresponds to previous work [7], [9].

In order to achieve a better segmentation result by using neighboring information, as well as good computational efficiency, we assume that neighbors are related to one object at a time. Thus, we only consider second-order prior information, i.e., the neighboring information from only one neighbor and the first-order prior, i.e., the self shape information. Let us consider a simple case where each object is related to object 1 independently; the corresponding R is

$$R = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 0 \\ & & \cdots & \\ 1 & 0 & \cdots & 1 \end{bmatrix}. \quad (7)$$

The joint density function can be simplified to

$$\begin{aligned} & p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) \\ &= p(\mathcal{S}_M | \mathcal{S}_1) p(\mathcal{S}_{M-1} | \mathcal{S}_1) \dots p(\mathcal{S}_2 | \mathcal{S}_1) p(\mathcal{S}_1) \\ &= p(\Delta_{M,1}) p(\Delta_{M-1,1}) \dots p(\Delta_{2,1}) p(\mathcal{S}_1) \end{aligned} \quad (8)$$

where $\Delta_{i,1}$ is a function of the difference between object i and object 1 (to be defined in the next section). The process of defining the joint density function $p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M)$ is simplified to building only the shape prior, $p(\mathcal{S}_i)$, and the local neighbor prior $p(\Delta_{i,1})$, $i = 2, 3, \dots, M$. In our MAP model, we consider this case for the remainder of this paper.

C. Neighbor Prior Model

To build a model for the neighbor prior and shape prior, we choose level sets as the representation of the objects, and then define the joint probability density function in (8).

Consider a training set of n aligned images, with M objects or structures in each image. Each object in the training set is embedded as the zero level set of a higher dimensional level set Ψ . For object 1, the training set consists of a set of level set functions $\{\Psi_{1,1}, \Psi_{2,1}, \dots, \Psi_{n,1}\}$. We can use the difference between the two level sets, $\Psi_i - \Psi_1$ as the representation of the neighbor difference $\Delta_{i,1}$, $i = 2, 3, \dots, M$. Thus, the corresponding training set is $\{\Psi_{1,i} - \Psi_{1,1}, \Psi_{2,i} - \Psi_{2,1}, \dots, \Psi_{n,i} - \Psi_{n,1}\}$, $i = 2, 3, \dots, M$. Our goal is to build the shape model

and neighbor difference model over these distributions of the level set functions and level sets differences.

The mean and variance of the boundary of object 1 can be computed using principal component analysis (PCA) [7]. The mean level set, $\bar{\Psi}_1 = (1/n) \sum_{l=1}^n \Psi_{l,1}$, is subtracted from each $\Psi_{l,1}$ to create the deviation from the mean. Each such deviation is placed as a column vector in a $N^d \times n$ -dimensional matrix Q where d is the number of spatial dimensions and N^d is the number of samples of each level set function. Using Singular Value Decomposition (SVD), $Q = U \Sigma V^T$. U is a matrix whose column vectors represent the set of orthogonal modes of shape variation and Σ is a diagonal matrix of corresponding singular values. An estimate of the object shape Ψ_1 can be represented by k principal components and a k -dimensional vector of coefficients (where $k < n$), α_1 [7]

$$\tilde{\Psi}_1 = U_k \alpha_1 + \bar{\Psi}_1. \quad (9)$$

Under the assumption of a Gaussian distribution of object represented by α_1 , we can compute the probability of a certain shape

$$p(\alpha_1) = \frac{1}{\sqrt{(2\pi)^k |\Sigma_k|}} \exp \left[-\frac{1}{2} \alpha_1^T \Sigma_k^{-1} \alpha_1 \right]. \quad (10)$$

Similarly, an estimate of the neighbor difference $\Delta_{i,1}$ can be represented from the mean neighbor difference $\bar{\Delta}_{i,1}$ and k principal components P_{ik} and a k -dimensional vector of coefficients, $\beta_{i,1}$

$$\tilde{\Delta}_{i,1} = P_{ik} \beta_{i,1} + \bar{\Delta}_{i,1}. \quad (12)$$

The neighbor difference $\Delta_{i,1}$ can also be assumed to be Gaussian distributed over $\beta_{i,1}$

$$p(\beta_{i,1}) = \frac{1}{\sqrt{(2\pi)^k |\Lambda_{i,1k}|}} \exp \left[-\frac{1}{2} \beta_{i,1}^T \Lambda_{i,1k}^{-1} \beta_{i,1} \right]. \quad (12)$$

Fig. 1 shows a training set of four subcortical structures: left and right amygdalae and hippocampi, where we assume the left amygdala is related to all of the other three structures, independently. By using PCA, we can build the shape model of the left amygdala and the neighbor difference models of the other three structures. Fig. 2 shows the zero level sets corresponding to the mean and three primary modes of variance of the distribution of the level set of the left amygdala. Fig. 3 shows the zero level sets of the three primary modes of variance of the neighbor difference between the level sets of the left hippocampus and the left amygdala added to the mean level set of left amygdala, i.e., the zero level set of $\bar{\Delta}_{2,1} + \bar{\Psi}_1$. We add to the mean left amygdala to allow visualization of the effect of the difference variance.

D. Energy Functional of the Model

In our deformable model, we also add some regularizing terms [10]: a general boundary

smoothness prior, $p_B(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) = \prod_{i=1}^M e^{-\mu_i \oint_{\mathcal{S}_i} ds}$, and a prior for the size of the region,

$p_A(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) = \prod_{i=1}^M e^{-v_i A_i^c}$, where A_i is the size of the region of shape i , c is a constant and μ_i and v_i are scalar factors. Here we assume the boundary smoothness and the region size of all the objects are independent. Thus, the prior joint probability $p(\mathcal{S}_1, \mathcal{S}_2 \dots \mathcal{S}_M)$ in (8) can be approximated by a product of the following probabilities:

$$p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) = \left[\prod_{i=2}^M p(\beta_{i,1}) \right] \cdot p(\alpha_1) \cdot p_B(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) \cdot p_A(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M) \quad (13)$$

since

$$\begin{aligned} \widehat{\mathcal{S}}_i &= \arg \max_{\mathcal{S}_i} p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_i, \dots, \mathcal{S}_M | I) \\ &= \arg \min_{\mathcal{S}_i} [-\ln p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_i, \dots, \mathcal{S}_M | I)] \end{aligned} \quad (14)$$

$i=1, 2, \dots, M.$

Combining (1), (2), and (13), we introduce the energy functional E defined by

$$\begin{aligned} E &= -\ln p(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_i, \dots, \mathcal{S}_M | I) \\ &\propto \sum_{i=1}^M \left\{ \lambda_{1i} \cdot \int_{(x,y,z) \text{ inside } (\mathcal{S}_i)} |I(x, y, z) - c_{1i}|^2 dx dy dz \right. \\ &\quad \left. + \lambda_{2i} \cdot \int_{(x,y,z) \text{ outside } (\mathcal{S}_i), \text{ inside } (\Omega_i)} |I(x, y, z) - c_{2i}|^2 dx dy dz \right\} \\ &\quad + \sum_{i=1}^M \mu_i \oint_{\mathcal{S}_i} ds + \sum_{i=1}^M v_i A_i^c \\ &\quad + \sum_{i=2}^M \omega_i \beta_{i,1}^T \Lambda_{i,1k}^{-1} \beta_{i,1} + \omega_1 \alpha_1^T \Sigma_k^{-1} \alpha_1. \end{aligned} \quad (15)$$

The MAP estimation of the objects in (1), $\mathcal{S}_i (i=1, 2, \dots, M)$, is also the minimizer of the above energy functional E . This minimization problem can be formulated and solved using the level set method and we can realize the segmentation of multiple objects simultaneously.

E. Level Set Formulation of the Model

In the level set method, $\mathcal{S}_i$ is the zero level set of a higher dimensional level set ψ_i corresponding to the i th object being segmented, i.e., $\mathcal{S}_i = \{(x, y, z) | \psi_i(x, y, z) = 0\}$. The evolution of surface $\mathcal{S}_i$ is given by the zero-level surface at time t of the function $\psi_i(t, x, y, z)$. We define ψ_i to be positive outside $\mathcal{S}_i$ and negative inside $\mathcal{S}_i$. Each of the M objects being segmented in the image has its own $\mathcal{S}_i$ and ψ_i .

For the level set formulation of our model, we replace $\mathcal{S}_i$ with ψ_i in the energy functional in (15) using regularized versions of the Heaviside function H and the Dirac function δ , denoted by H_ϵ and δ_ϵ [3] (described below)

$$\begin{aligned}
& E(c_{1i}, c_{2i}, \psi_i | i=1, 2, \dots, M) \\
&= \sum_{i=1}^M \left[\lambda_{1i} \int_{\Omega} |I(x, y, z) - c_{1i}|^2 \right. \\
&\quad \times (1 - H_{\varepsilon}(\psi_i(x, y, z))) dx dy dz \\
&\quad + \lambda_{2i} \int_{\Omega_i} |I(x, y, z) - c_{2i}|^2 H_{\varepsilon}(\psi_i(x, y, z)) dx dy dz \\
&\quad + \mu_i \int_{\Omega} \delta_{\varepsilon}(\psi_i(x, y, z)) |\nabla \psi_i(x, y, z)| dx dy dz \\
&\quad + \nu_i \int_{\Omega} (1 - H_{\varepsilon}(\psi_i(x, y, z))) dx dy dz \Big] \\
&\quad + \sum_{i=2}^M \omega_i [G(\psi_i - \psi_1) - \bar{\Delta}_{i,1}]^T P_{ik} \Lambda_{i,1k}^{-1} \\
&\quad \cdot P_{ik}^T [G(\psi_i - \psi_1) - \bar{\Delta}_{i,1}] \\
&\quad + \omega_1 [G(\psi_1 - \bar{\psi}_1)]^T U_k \Sigma_k^{-1} U_k^T [G(\psi_1 - \bar{\psi}_1)]
\end{aligned} \tag{16}$$

where Ω denotes the image domain. In this paper, we define Ω_i to be the domain where the level set distance is no more than five pixels/voxels considering that the size of objects of interest in our data is less than 20–30 pixels/voxels, i.e., $\Omega_i = \{(x, y, z) \mid \psi_i(x, y, z) \leq 5\}$. $G(\cdot)$ is an operator to generate the vector representation (as in (9)) of a matrix by column scanning. $g(\cdot)$ is the inverse operator of $G(\cdot)$. To compute the associated Euler-Lagrange equation for each unknown level set function ψ_i , we keep c_{1i} and c_{2i} fixed, and minimize E with respect to ψ_i ($i=1, 2, \dots, M$) respectively. Parametrizing the descent direction by artificial time $t \geq 0$, the evolution equations in $\psi_i(t, x, y, z)$, $i=1, 2, \dots, M$ are

$$\begin{aligned}
\frac{\partial \psi_1}{\partial t} = & \delta_{\varepsilon}(\psi_1) \left[\lambda_{11} |I - c_{11}|^2 - \lambda_{21} |I - c_{21}|^2 \right. \\
& \left. + \mu_1 \cdot \operatorname{div} \left[\frac{\nabla \psi_1}{|\nabla \psi_1|} \right] + \nu_1 \right] \\
& - \sum_{i=2}^M \omega_i \cdot g \left\{ P_{ik} \Lambda_{i,1k}^{-1} P_{ik}^T [G(\psi_i - \psi_1) - \bar{\Delta}_{i,1}] \right\} \\
& - \omega_1 \cdot g \{ U_k \Sigma_k^{-1} U_k^T [G(\psi_1 - \bar{\psi}_1)] \}
\end{aligned} \tag{17}$$

and

$$\begin{aligned}
\frac{\partial \psi_i}{\partial t} = & \delta_{\varepsilon}(\psi_i) \left[\lambda_{1i} |I - c_{1i}|^2 - \lambda_{2i} |I - c_{2i}|^2 \right. \\
& \left. + \mu_i \cdot \operatorname{div} \left[\frac{\nabla \psi_i}{|\nabla \psi_i|} \right] + \nu_i \right] \\
& - \omega_i \cdot g \left\{ P_{ik} \Lambda_{i,1k}^{-1} P_{ik}^T [G(\psi_i - \psi_1) - \bar{\Delta}_{i,1}] \right\} \\
& i=2, 3, \dots, M.
\end{aligned} \tag{18}$$

F. Evolving the Surface

We approximate H_{ε} and δ_{ε} as follows [3]

$$H_{\varepsilon}(z) = \frac{1}{2} \left[1 + \frac{2}{\pi} \arctan \left(\frac{z}{\varepsilon} \right) \right] \tag{19}$$

$$\delta_{\varepsilon}(z) = \frac{\varepsilon}{\pi(\varepsilon^2 + z^2)}. \tag{20}$$

The constants c_{1i} and c_{2i} are defined by [3], [12]

$$c_{1i}(\psi_i) = \frac{\int_{\Omega} I(x, y, z) \cdot (1 - H(\psi_i(x, y, z))) dx dy dz}{\int_{\Omega} (1 - H(\psi_i(x, y, z))) dx dy dz} \quad (21)$$

$$c_{2i}(\psi_i) = \frac{\int_{\Omega_i} I(x, y, z) \cdot H(\psi_i(x, y, z)) dx dy dz}{\int_{\Omega_i} H(\psi_i(x, y, z)) dx dy dz}. \quad (22)$$

Given the surfaces $\psi_i (i = 1, 2, \dots, M)$ at time t , we seek to compute the evolution steps that bring all the zero level set curves to the correct final segmentation based on the neighbor prior information and image information. We first set up $p(\alpha_1)$ and $p(\beta_{i,1})$, $i = 2, 3, \dots, M$ from the training set using PCA. At each stage of the algorithm, we recompute the constants $c_{1i}(\psi_i^t)$ and $c_{2i}(\psi_i^t)$ and update ψ_i^{t+1} . This is repeated until convergence.

To simplify the complexity of the segmentation system, we generally choose the parameters in our experiments as follows: $\lambda_{1i} = \lambda_{2i} = \lambda_i$, $\mu_i = 0.000\ 05 \cdot 255^2$, and $\nu_i = 0$ [3]. This leaves us only two free parameters (ω_i and λ_i) to balance the influence of two terms, the image data term and the neighbor prior term for each object. The tradeoff between neighbor prior and image information depends on the strength of the neighbor prior and the image quality for a given application. We set these parameters empirically for particular segmentation tasks.

III. Experimental Results

A. Evaluation of Level Set Distribution Models

The LSDM presented in Section II-C can be used as an approximation of the distribution of the object self shape. Since the level set function is a nonlinear function of the corresponding zero level set (i.e., the surface of the object), it is very difficult to theoretically analyze the performance of the LSDM as an approximation of the shape model. A practical way to evaluate it is to compare the LSDM with the PDM [4].

In a PDM, each object is represented by a set of points on the boundary. The labeling of the points is important. Each point represents a particular part of the object boundary. The point must be located in the same way on each of the training images, which may be difficult for three-dimensional (3-D) objects although minimum description length [15] approaches can be used to define point correspondences.

Let $X_i = (x_{i1}, y_{i1}, x_{i2}, y_{i2}, \dots, x_{iL}, y_{iL})^T$ be a vector describing the set of L points of the i th object in the set. (x_{ij}, y_{ij}) is the j th point of the i th object. Table I shows the process of getting the LSDM and PDM over a set of n aligned training images.

By using PCA of the level sets of the left ventricles in Fig. 4, we can build a model of the shape of left ventricles. Fig. 5 (top) illustrates zero level sets corresponding to the mean and three primary modes of variance of the distribution of the level set of the left ventricle. Note that the zero level sets of the mean level set and primary modes appear to be representative shapes of the class of object being learned. Fig. 5 (bottom) shows the three PDM-based primary modes of variance of the left ventricle using the same training set as in Fig. 4 where the correspondences between the boundary points were manually determined. Compared to the PDM, the zero level set tends to have a smoother boundary over the distribution. The two models are quite similar in describing the main variations of the shape.

We then estimate a number of test left ventricles using LSDM and PDM as shown in Table I (parametric model). Fig. 6 shows the two estimates of the left ventricles (with five primary modes). Most parts of the two curves overlap well. To compare the two estimates, we compute the undirected average distance between the zero level set A (N_A points) and the sample points B

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (23)$$

where

$$h(A, B) = \frac{1}{N_A} \sum_{a \in A} \min_{b \in B} \|a - b\|. \quad (24)$$

The distances of the estimates for the six cases in Fig. 6 are shown in Table II (unit: pixel). Virtually all the points in the zero level set lie within one or two pixels of the PDM estimate.

Next, we evaluate the two distributions using a Chi-square test, again using the left ventricle as the example. Since the PDM is built in the sample point vector space, while the LSDM is in the level set space, we need to convert the two distributions to the same space first in order to compare them. We first randomly generate a set of samples $\{X_1, X_2, \dots, X_m, m = 10^5\}$ (each sample corresponds to a shape represented by a sample point vector) based on the PDM. Each X_i can be embedded as the zero level set of a higher dimensional level set Ψ_i , the projection of X_i in the level set space. Then we compare the set of projections $\{\Psi_1, \Psi_2, \dots, \Psi_m, m = 10^5\}$ with the LSDM shown in (10) using the Chi-square test, with the null hypothesis, H_0 , that the projections have the same distribution defined by the LSDM. Consider the first five principal components in the LSDM, with five bins along each of the principal modes, the Chi-square statistic is given by

$$\chi^2 = \sum \frac{(\text{Ob.} - \text{Ex.})^2}{\text{Ex.}} \quad (25)$$

where the sum is over all the bins considered, *Ob.* is the observed frequency of Ψ s falling in each bin, *Ex.* is the corresponding expected frequency which can be calculated by integrating the probability function in (10) over the bin. For our test $[(5 - 1)^5 = 1024 \text{ degrees of freedom}]$, $\chi^2 = 981$, with a upper tail probability $\Pr(\chi^2 > 981) = 0.82$. Therefore, there is a 0.82 probability that it is wrong to reject H_0 . Thus, the two distributions are statistically indistinguishable.

The level set representation of shape is an efficient way to formulate a distribution model for both 2-D and 3-D shapes, and appears to be an accurate approximation to the corresponding PDM. Moreover, it does not require explicit point correspondences (since the object shape is embedded in a level set).

B. Applications to Medical Imagery

We have used our model on various synthetic and real images, with at least two different types of contours and shapes. We rigidly align the training data before performing PCA. Then we construct the LSDM for the prior.

In Fig. 7 (top), we show the segmentation of the left and right ventricles using only image information, by which the curves cannot lock onto the shapes of the objects. In Fig. 7 (bottom), we show the results obtained using our model. The curves are able to converge on

the desired boundaries even though some parts of the boundaries are too blurred to be detected using only gray level information. Both of the segmentations converged in a couple of minutes on a 2.00-GHz Intel XEON CPU.

In Fig. 8, we show that our model can detect multiple objects of different intensities and with blurred boundaries. Fig. 8 (top) shows the results of using only gray level information. Only the lower (posterior) portions of the lateral ventricles can be segmented perfectly since they have clearer boundaries. Fig. 8 (bottom) shows the results obtained using our neighbor prior model. Segmenting all eight subcortical structures took no more than five minutes.

Fig. 9 shows the segmentation of the right amygdala and hippocampus in a 2-D MR image. In Fig. 9 (top), we show results of using only gray level information. The segmentations are poor since both structures have very poorly defined boundaries. The middle row in Fig. 9 shows the results of using the shape prior but with no neighbor prior. The results are much better, but the boundaries of the amygdala and the hippocampus overlap where the two structures are connected. This is because the two structures are treated independently here without the constraint of the neighbor. In Fig. 9 (bottom), we show results of using our neighbor prior model. The two structures can be clearly segmented, and there is no overlap of the boundaries due to the use of the neighbor constraint. Since the neighbor constraint is based on statistical prior information from a training set, there is a possibility of getting intersections of the neighboring objects under some applications. However, compared to the shape prior model, our neighbor prior model can improve and even circumvent this intersection problem statistically. Segmenting the two structures took several minutes.

We also tested our method in 3-D images. We generated a training set of 12 synthetic images of two uniform ellipsoids with added Gaussian noise. Fig. 10 illustrates several steps in the segmentation of the two ellipsoids. Fig. 11 shows initial, middle, and final steps in the segmentation of the left and right amygdalae and hippocampi in a 3-D MR brain image using training set model shown in Figs. 1, 2, and 3. Three orthogonal slices and the 3-D surfaces are shown for each step. Segmenting these 3-D images (with size $172 \times 148 \times 124$) took approximately a couple of hours.

To validate the segmentation results, we tested our model on 12 different images for each of the above five cases (as shown in Figs. 7–11) respectively; the tested images are not in their training sets. We then computed the undirected average distance between the boundary of the computed segmentation and the boundary of the manual segmentation using (23) and (24). We also computed the standard deviation of the computed distances for each case. Table III shows the mean and standard deviation of the computed results for each of the cases: 1) 2-D cardiac images; 2) 2-D brain images with eight subcortical structures; 3) 2-D brain images with right amygdala and hippocampus; 4) 3-D synthetic images; and 5) 3-D brain images with left and right amygdalae and hippocampi. Virtually all the boundary points lie within one or two voxels of the manual segmentation when using the neighbor prior.

We also tested the robustness of our algorithm to noise as well as to the location of the initial seeds. First, we added Gaussian noise with different variances to the synthetic image (as in Fig. 10, the mean intensities of objects/background: 45/65), and ran our algorithm to segment the two ellipsoids, where we set the initial seeds at the centers of the objects. Fig. 12 shows the segmentation error in three cases: with no prior, with shape prior, and with neighbor prior. When the variance of the noise is small, the errors are also small for all the three cases. As the variance of the noise increases, the error for no prior increases rapidly since the objects are too noisy to be recognized using only gray level information. However, for shape prior and neighbor prior methods, the segmentation errors are much lower and are

locked in a very small range even when the variance of the noise is very large. We also notice that our neighbor prior model achieves the least error among all the cases.

Next, we fix the standard deviation of the noise to be 40, but vary the location of the initial seed inside the right ellipsoid and run the segmentation for the same three cases again. The segmentation error for different seed locations with each method is shown in Fig. 13. As the initial seed moves far from the center of the ellipsoid, the errors are kept in a small range for all cases because the models are based on level sets. Still, the method with the neighbor prior achieves the smallest error.

In our study, we have focused on balancing the weights of the image data (λ_i) and neighbor prior (ω_i) with all the other parameters fixed. We tested the sensitivity of our method to the balance of the prior weight and image data weight. For all five applications presented in the paper, the weights of the image data and prior information can be varied by around 30% with the corresponding segmentation errors changing by no more than 5%. Thus, our method is not sensitive to the balance of the weights.

IV. Discussion and Conclusion

We have demonstrated that statistical neighbor prior information can be used in automated image segmentation of difficult objects. We define a MAP estimation framework combining both the neighbor prior information provided by neighboring objects as well as the image gray level information to segment multiple objects simultaneously, where the key point is to design the prior term and the image term so that they can work accurately and effectively. In this paper, we have focused on formulating local neighbor prior models as described in Section II-B. This method is designed for situations with limited interobject information as opposed to full global atlases. More general neighbor prior models (as discussed in Section II-B) can also be used to incorporate more neighbor priors. Recently, we have proposed such a model [13].

The image term we proposed here may be enhanced. For instance, it would be interesting to extend the image gray level based model to a more general case where additional information is included besides the image intensity average, as has been discussed [3].

As a powerful approach to deformable models, the MAP framework can combine prior information and given image information to realize image segmentation. While we have not pursued it in this paper, a more unified model could be used to incorporate shape priors, neighbor constraints as well as image gray level priors [14].

In this paper, we introduced a level set representation for the object shapes, and defined distributions over the variations of local neighboring positions and shapes in a set of training images. We evaluated the LSDM of object shape by comparing it with PDM. For the endocardial boundary representations shown in Section III-A, it is clear that either representation would be acceptable even though they are slightly different. Both models can be used as approximations to the distribution of object shape. However, both LSDM and PDM have not been analyzed extensively to their utility across all problems. Further study on this is ongoing work and might be difficult and challenging. Level set representation of shape provides tolerance to slight misalignment of object shape, in the attempt to avoid having to solve the general correspondence problem. In practice, the variations captured by the principal components in LSDM (U_k and P_{ik}) in this paper are based on rigid alignment of the training data and may contain undesired residuals due to misalignment. We are looking for better alignment method that can reduce such residuals and the topology problem.

We have shown that level set based neighbor prior models can be used to improve image segmentation. We estimate the MAP shapes of the objects using evolving level sets based on the associated Euler-Lagrange equations. The contours evolve both according to the neighbor prior information and the image gray level information. Multiple objects in an image can be automatically detected simultaneously.

Acknowledgments

The authors would like to thank L. Win and R. T. Schultz for help with the visualization and manual tracing of the MR data. They also thank H. Tagare for the many thoughtful discussions and comments. Finally, they wish to thank the editors and reviewers for their general and detailed comments and suggestions, which very much helped to improve the presentation of the paper.

This work was supported by the National Institutes of Health (NIH) under Grant NINDS NS035193.

References

1. Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. *Int J Comput Vis.* 1987; 1:321–331.
2. Osher S, Sethian JA. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulation. *J Comp Phys.* 1988; 79:12–49.
3. Chan T, Vese L. Active contours without edges. *IEEE Trans Image Processing.* Feb.2001 10:266–277.
4. Cootes TF, Hill A, Taylor CJ, Haslam J. Use of active shape models for locating structures in medical images. *Image Vis Computing.* July; 1994 12(6):355–365.
5. Staib L, Duncan J. Boundary finding with parametrically deformable models. *IEEE Trans Pattern Anal Machine Intell.* Nov.1992 14:1061–1075.
6. Zeng, X.; Staib, LH.; Schultz, RT.; Duncan, JS. Volumetric layer segmentation using coupled surfaces propagation. *Proc. IEEE Conf. Comp. Vision Pattern Recognition;* June 1998; p. 708-715.
7. Leventon M, Grimson E, Faugeras O. Statistical shape influence in geodesic active contours. *Proc IEEE Conf Comp Vision Pattern Recognition.* 2000; 1:316–323.
8. Caselles V, Kimmel R, Sapiro G. Geodesic active contours. *Int J Comput Vis.* 1997; 22(1):61–79.
9. Tsai, A.; Yezzi, A.; Wells, W., III; Tempany, C.; Tucker, D.; Fan, A.; Grimson, WE.; Willsky, A. Model-based curve evolution technique for image segmentation. *Proc. IEEE Conf. Comp. Vision Pattern Recognition;* 2001. p. 463-468.
10. Tu Z, Zhu S. Image segmentation by data-driven Markov Chain Monte Carlo. *IEEE Trans Pattern Ana Machine Intell.* May.2002 24:657–673.
11. Yang, J.; Staib, L.; Duncan, J. Statistical neighbor distance influence in active contours. *Proc. Medical Image Computing and Computer-Assisted Intervention Conf. (MICCAI);* 2002. p. 588-596.
12. Yang, J.; Staib, L.; Duncan, J. Neighbor-constrained segmentation with 3-D deformable models. In: Taylor, CJ.; Noble, JA., editors. *Lecture Notes in Computer Science; Proc. 18th Int. Conf. Information Processing in Medical Imaging (IPMI 2003);* 2003. p. 198-209.
13. Yang, J.; Duncan, J. Joint prior models of neighboring objects for 3-D image segmentation. presented at the *IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR);* Washington, DC. 2004;
14. Yang, J.; Duncan, J. 3-D image segmentation of deformable objects with shape-appearance joint prior models. *Proc. Medical Image Computing and Computer-Assisted Intervention Conf. (MICCAI);* 2003. p. 573-580.
15. Davies, RhH; Twining, CJ.; Cootes, TF.; Waterson, JC.; Taylor, CJ. A minimum description length approach to statistical shape modeling. *IEEE Trans Med Imag.* May.2002 21:525–537.

Fig. 1.
Training set: outlines of four subcortical structures—left and right amygdalae and hippocampi in 12 3-D MR brain images.

Fig. 2. The three primary modes of variance of the left amygdala, showing the mean and $\pm$ standard deviation (σ).

Fig. 3. The three primary modes of variance of the left hippocampus relative to the left amygdala (zero level set of $\tilde{\Delta}_{2,1} + \Psi_1$ for visualization).

Fig. 4. Outlines of left ventricles in 6 out of 16 MR training images gated and at a fixed point in the cardiac cycle.

Fig. 5. The three primary modes of variance of the left ventricle using level set (top rows) and point (bottom rows) model.

Fig. 6.
Level set (black) and point model (red) based estimates of the left ventricles.

Fig. 7.

Three steps in the segmentation of two shapes in a 2-D cardiac MR image (top) without and (bottom) with neighbor prior. The training set consists of 16 images. $\lambda_i = \omega_i = 0.5$, $i = 1, 2$.

Fig. 8. Detection of eight subcortical structures (the lateral ventricles ($\lambda = 0.8$, $\omega = 0.2$), heads of the caudate nucleus ($\lambda = 0.3$, $\omega = 0.7$), and putamina ($\lambda = 0.2$, $\omega = 0.8$)) in a MR brain image. (top) Results with no prior information comparing with manual segmentation. (bottom) Results with neighbor prior comparing with manual segmentation. The training set consists of 12 images.

Fig. 9.

Four steps in the segmentation of the right amygdala and hippocampus. (top) Results with no prior information. (middle) Results using individual shape priors. (bottom) Results using our neighbor prior model. The training set consists of 12 brain images. $\lambda_i = 0.1$, $\omega_i = 0.9$, $i = 1, 2$.

Fig. 10.

Initial, middle, and final steps (top to bottom) in the segmentation of two shapes in a synthetic image. Three orthogonal slices and the 3-D surfaces are shown for each step. $\lambda_i = 0.2$, $\omega_i = 0.8$, $i = 1, 2$.

Fig. 11.

Initial, middle, and final steps (top to bottom) in the segmentation of four shapes in a brain image. Three orthogonal slices and the 3-D surfaces are shown for each step. $\lambda_i = 0.1$, $\omega_i = 0.9$, $i = 1, 2, 3, 4$.

Fig. 12. Segmentation errors (unit: voxel) with different noise variances for synthetic image with mean intensities of objects/background: 45/65.

Fig. 13. Segmentation errors (unit:voxel) with different locations (unit:voxel) of initial seeds for synthetic image with mean intensities of objects/background: 45/65.

TABLE I

LSDM and PDM

	LSDM	PDM
Shape Representation	$\{\Psi_1, \Psi_2, \dots, \Psi_n\}$ Ψ_i : N^d elements	$\{X_1, X_2, \dots, X_n\}$ X_i : dL elements
Mean Shape	$\bar{\Psi} = \frac{1}{n} \sum_{i=1}^n \Psi_i$	$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
Mean Offset Matrix	Q_{Ψ}	Q_X
SVD	$Q_{\Psi} = U_{\Psi} \sum_{\Psi} V_{\Psi}^T$	$Q_X = U_X \sum_X V_X^T$
Parametric Model	$\Psi = \bar{\Psi} + U_{\Psi,k} \alpha_{\Psi}$	$\hat{X} = \bar{X} + U_{X,k} \alpha_X$
Probability of Shape	$p(\alpha_{\Psi}) \sim \text{Gauss.}$	$p(\alpha_X) \sim \text{Gauss.}$

TABLE II

Distance Between the LSDM and the PDM Estimates

Example	1	2	3	4	5	6
Distance (unit pixel)	2.1	1.8	1.3	0.9	0.7	1.1

TABLE III
Mean Distance and Standard Deviation Between the Computed and the Manual Boundaries

	1	2	3	4	5	
Without neighbor prior	Mean	4.1	9.8	6.6	11.2	7.6
	std.	4.8	12.5	10.1	12.8	20.6
With neighbor prior	Mean	1.7	2.0	0.9	1.1	1.8
	std.	1.1	1.3	0.8	1.0	1.2

std.: standard deviation, unit: pixel/voxel. 1: 2D cardiac images; 2: 2D brain images with eight subcortical structures; 3: 2D brain images with right amygdala and hippocampus; 4: 3D synthetic images; 5: 3D brain images with left and right amygdalae and hippocampi.