SEGMENTATION OF 3D DEFORMABLE OBJECTS WITH LEVEL SET BASED PRIOR MODELS

Jing Yang *Hemant D. Tagare* *Lawrence H. Staib* *James S. Duncan*

Departments of Electrical Engineering and Diagnostic Radiology, Yale University
P.O. Box 208042, New Haven CT 06520-8042, USA
Email: j.yang@yale.edu, hemant.tagare@yale.edu, lawrence.staib@yale.edu, james.duncan@yale.edu

ABSTRACT

We propose a level set based deformable model for the segmentation of multiple objects from 3D medical images using shape prior constraints. As an extension to the level set distribution model of object shape presented in [1][2][3], this paper evaluates the performance of the level set representation of the object shape by comparing it with the point distribution model(PDM)[4] using the Chi-square test. We define a Maximum A Posteriori(MAP) estimation model using level set based prior information to realize the segmentation of the multiple objects. To achieve this , only one level set function is employed as the representation of the multiple objects of interest within the image. We then define the probability distribution over the variations of objects contained in a set of training images. We found the algorithm to be computationally efficent, robust to noise, able to handle multidimensional data, and avoids the need for explicit point correspondences during the training phase. Results and validation from various experiments on 2D/3D medical images are demonstrated.

1. INTRODUCTION

Segmenting the structures or objects in an image is of great importance in a variety of applications including medical image processing, computer vision and pattern recognition. Image gray level based methods often face difficult challenges such as poor image contrast, noise, and missing or diffuse boundaries. Prior model based algorithms can often solve this problem. We present such a model in this paper.

Our work shares the common observation that prior information about the expected shape can significantly aid in image segmentation as well as to reduce the complexity of the segmentation process. 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. Leventon et al. [1] incorporate shape information into an active contour evolution process. Tsai et al. [2] use mutual information to segment multiple objects simultaneously. Also in our previous work[3], we propose a neighbor-constrained probabilistic estimation framework for the segmentation of multiple objects from 3D medical images. In this paper, we also employ a level set representation for deformable objects and level set distribution models of object shapes.

Our approach to multiple objects segmentation is based on a MAP estimation framework using level set based prior information of the objects in the image. We evaluate this level set distribution model by comparing it with the traditional point distribution model[4] using the Chi-square test. A single level set function is used to represent multiple objects of interest. We then define the probability distribution over the variations of objects contained in a set of training images. By estimating the MAP shapes of the objects using this shape-based deformable model, multiple objects can be segmented simultaneously.

2. PROBABILITY DISTRIBUTION ON SHAPES

2.1. Level Set Distribution Models(LSDM)

Consider a set of n training images $\{I_1, I_2, ..., I_n\}$ with a shape of interest in each image, for example the left ventricle of the heart as seen in Fig. 1. In order to compare the structures from different training images, the images must be aligned in the same way with respect to a set of axes. We achieve this by rigidly aligning the training images so that they match as close as possible. To model the object shape, we choose a level set representation of the shape over the training set[1][2][3]. The surfaces of the structure in the training set are embedded as the zero level set of n separate higher dimensional level sets Ψ with negative distances inside and positive distances outside the object. The training set, τ, consists of a set of level sets $\tau = \{\Psi_1, \Psi_2, ..., \Psi_n\}$. Our goal is to build the shape model over this distribution of level set functions.

The mean and variance of the level sets can be computed using Principal Component Analysis(PCA).[4] The mean $\bar{\Psi} = \frac{1}{n}\sum_{i=1}^{n}\Psi_i$ is subtracted from each Ψ_i to create the deviation from the mean. Each such deviation $d\Psi_i$ is placed as

0-7803-8388-5/04/$20.00

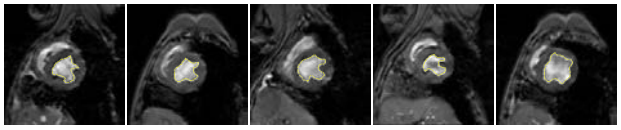

Fig. 1. Outlines of left ventricle in 5 out of 16 MR training images.

a column vector in a $N^d \times n$ dimensional matrix Q.(Where d is the number of dimensions, 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 level set variation and Σ is a diagonal matrix of corresponding singular values. An estimate of the level set Ψ can be approximated by k principal components in a k dimensional vector of coefficients, α[4]: $\tilde{\Psi} = \bar{\Psi} + U_k\alpha$. Thus, an estimate of the shape can be approximated by the zero level set of the estimated level set, i.e., $\tilde{S} = \{(x, y, z)|\tilde{\Psi}(x, y, z) = 0\}$.

Under the assumption of a Gaussian distribution of level set represented by α, we can compute the probability of a level set corresponding to a certain shape:

$$p(\alpha) = \frac{1}{\sqrt{(2\pi)^k|\Sigma_k|}} \exp[-\frac{1}{2}\alpha^T\Sigma_k^{-1}\alpha] \qquad (1)$$

By using PCA of the level sets of the left ventricles in Fig. 1, we can build a model of the shape of left ventricles. Fig. 2 left 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 reasonable representative shapes of the class of object being learned.

2.2. Evaluation of LSDMs

The LSDM presented above can be used as an approximation of the distribution of the object 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 PDM, each shape is represented by a set of points on the boundary. The labelling of the points is important. Each point represents a particular part of the object boundary, so it must be located in the same way on each of the training images, which is particularly difficult for 3D shapes.

Let $X_i = (x_{i1}, y_{i1}, x_{i2}, y_{i2}, ..., x_{iL}, y_{iL})^T$ be a vector describing the set of L points of the i^{th} shape in the set. (x_{ij}, y_{ij}) is the j^{th} point of the i^{th} shape. Table 1 shows the PCA processes of getting the LSDM and PDM over a set of n aligned training images.

Fig. 2 right shows the three PDM-based primary modes of variance of the left ventricle using the same training set as in Fig. 1. Compared to the PDM, the zero level set tends

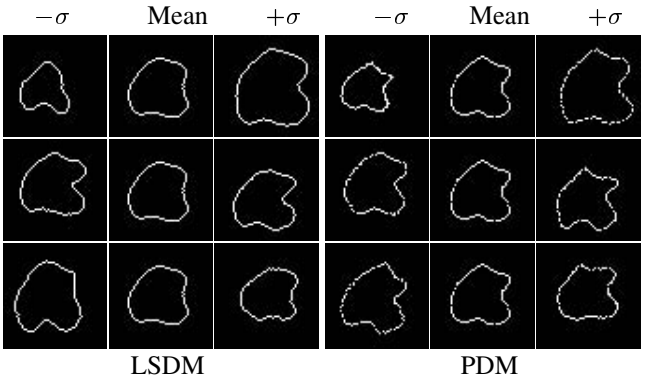

Fig. 2. The 3 primary modes of variance of the left ventricle using LSDM(left columns) and PDM(right columns).

Table 1. LSDM and PDM

	LSDM	PDM
Shape Representation	$\{\Psi_1, \Psi_2, ..., \Psi_n\}$ $\Psi_i : N^d$ elements	$\{X_1, X_2, ..., X_n\}$ $X_i : dL$ 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	$Q_\Psi = U_\Psi\Sigma_\Psi V_\Psi^T$	$Q_X = U_X\Sigma_X V_X^T$
Parametric Model	$\tilde{\Psi} = \bar{\Psi} + U_{\Psi,k}\alpha_\Psi$	$\tilde{X} = \bar{X} + U_{X,k}\alpha_X$
Probability of Shape	$p(\alpha_\Psi) \sim Gauss.$	$p(\alpha_X) \sim Gauss.$

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 1, the 6^{th} row. Fig. 3 shows the two estimations of the left ventricles(with 5 primary modes). Most parts of the two curves overlap well. To compare the two estimations, we compute the undirected distance between the zero level set $A(N_A$ points) and the sample points B: $H(A, B) = \max(h(A, B), h(B, A))$, $h(A, B) = \frac{1}{N_A}\sum_{a\in A}\min_{b\in B}\|a-b\|$. The distances of the estimations for the 5 cases in Fig. 3 are(unit:pixel): 2.1, 1.8, 1.3, 0.9, and 0.7. Virtually all the points in the zero level set lie within one or two pixels of the PDM estimation.

Next, we evaluate the two distributions by doing 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 in the same space first in order to do the comparison. We first randomly generate a set of samples $\{X_1, X_2, ..., X_m, m = 10^5\}$ (each sample corresponds

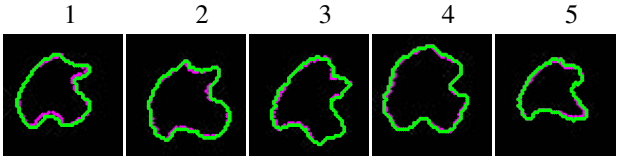

Fig. 3. Level set(green) and point model(red) based estimations of the left ventricles.

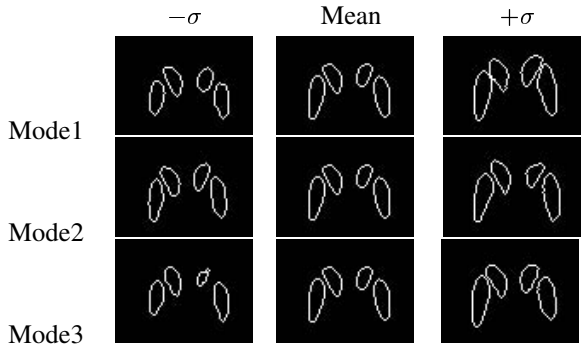

Fig. 4. The three primary modes of variance of 4 subcortical structures using level set model.

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 Y_i, the projection of X_i in the level set space. Then we compare the set of projections $\{Y_1, Y_2, ..., Y_m, m = 10^5\}$ with the LSDM shown in equation (1) using the Chi-square test, where we define the H_0 hypothesis to be that the projections have the same distribution defined by the LSDM. Consider the first 5 principal components in the LSDM, with 5 bins along each of the principal modes, the Chi-square statistic is given by:

$$\chi^2 = \sum (Ob. - Ex.)^2 / Ex. \tag{2}$$

where the sum is over all the bins considered, $Ob.$ is the observed frequency of Ys falling in each bin, $Ex.$ is the corresponding expected frequencies which can be calculated by integrating the probability function in equation (1) over the bin. For our test(1024 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.

We can also use LSDM for multiple shapes in each training image, where the surfaces of the shapes in each image can be embedded as the zero level set of one single level set Ψ. Fig. 4 shows zero level set corresponding to the mean and three primary modes of variance of the distribution of the level set of the 4 sub-cortical structures.

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

3. SEGMENTATION MODEL WITH MULTI-SHAPE PRIOR

3.1. MAP Framework

Consider an image I that has M shapes of interest: S_1, S_2,...,S_M; a MAP framework can be used to realize image segmentation combining shape prior information and image information[3] $(i = 1, 2, ..., M)$:

$$\begin{aligned} \hat{S}_i &= \arg\max_{S_i} p(S_1, S_2, ..., S_M / I) \\ &= \arg\max_{S_i} p(I / S_1, S_2, ..., S_M) p(S_1, S_2, ..., S_M) \end{aligned} \tag{3}$$

$p(I/S_1, S_2, ..., S_M)$ is the probability of producing an image I given S_1, S_2, ... , S_M. In 3D, assuming gray level homogeneity within all the M objects, we use the following imaging model[3]:

$$\begin{aligned} p(I/S_1, S_2, ..., S_M) &= \prod_{(p,q,r) inside(S_{all})} e^{-\frac{[I(p,q,r)-c_1]^2}{2\sigma_1^2}} \\ &\cdot \prod_{(p,q,r) outside(S_{all}), inside(\Omega')} e^{-\frac{[I(p,q,r)-c_2]^2}{2\sigma_2^2}} \end{aligned} \tag{4}$$

where $S_{all} = \{S_i | i = 1, 2, ..., M\}$, c_1 and σ_1 are the average and variance of I inside S_{all}, c_2 and σ_2 are the average and variance of I outside S_{all} but also inside a certain domain Ω' that contains S_{all}.

$p(S_1, S_2, ..., S_M)$ is the probability function of all the M objects, which can be approximated by equation (1) using the LSDM as discussed in section 2, where the surfaces of the M shapes in each training image can be embedded as the zero level set of one level set Ψ. In our active contour model, we also add a general smoothness Gibbs prior[3] for the region boundaries $p_B(S_{all}) = \prod_{i=1}^{M} e^{-\mu \oint_{S_i} ds}$. where μ is a scalar factor. Thus, $p(S_1, S_2, ..., S_M)$ can be approximated by a product of the following probabilities:

$$p(S_1, S_2, ..., S_M) \quad = \quad p(\alpha) \cdot p_B(S_{all}) \tag{5}$$

Combining equations (1) and (3)-(5), we introduce the energy functional E defined by

$$\begin{aligned} & E = -\ln p(S_1, S_2, ..., S_M / I) \\ & \propto \lambda_1 \int_{(p,q,r) inside(S_{all})} |I(p,q,r) - c_1|^2 dpdqdr \\ & + \lambda_2 \int_{(p,q,r) outside(S_{all}), inside(\Omega')} |I(p,q,r) - c_2|^2 dpdqdr \\ & + \mu \sum_{i=1}^{M} \oint_{S_i} ds + \tfrac{1}{2} \alpha^T \Sigma_k^{-1} \alpha \end{aligned} \tag{6}$$

The MAP estimation of the shapes in equation (3), $\hat{S}_i (i = 1, 2, ..., M)$, 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.

3.2. Level Set Formulation of the Model

In the level set method, S_{all} is the zero level set of a higher dimensional level set ψ corresponding to the M objects being segmented: $S_{all} = \{(x, y, z) | \psi(x, y, z) = 0\}$. For the level set formulation of our model, we replace S_{all} with ψ in (6) using regularized versions of the Heaviside function H and the Dirac function δ, denoted by H_ε and δ_ε[5]:

$$\begin{aligned} & E(c_1, c_2, \psi) = \{\mu \int_\Omega \delta_\varepsilon(\psi(x,y,z)) |\nabla \psi(x,y,z)| dxdydz \\ & + \lambda_1 \int_\Omega |I(x,y,z) - c_1|^2 (1 - H_\varepsilon(\psi(x,y,z))) dxdydz \\ & + \lambda_2 \int_\Omega |I(x,y,z) - c_2|^2 H_\varepsilon(\psi(x,y,z)) dxdydz\} \\ & + \tfrac{1}{2} [G(\psi - \bar{\Psi})]^T U_k \Sigma_k^{-1} U_k^T [G(\psi - \bar{\Psi})] \end{aligned}$$

where Ω denotes the image domain. $G(\cdot)$ is an operator to generate the vector representation(as in section 2) of a

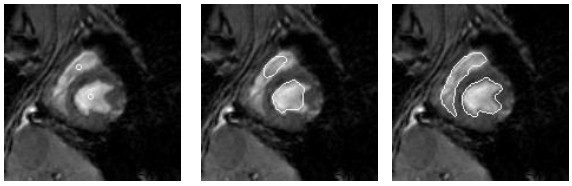

Fig. 5. 3 steps in the segmentation of 2 shapes in a 2D cardiac MR image. The training set consists of 16 images.

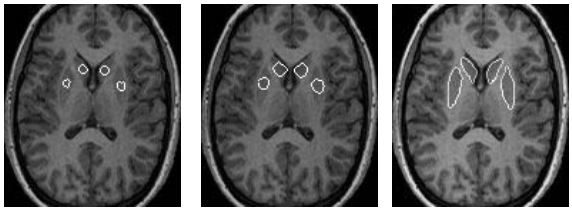

Fig. 6. Detection of 4 sub-cortical structures(heads of the caudate nucleus and putamina) in a MR brain image. The training set consists of 12 images.

matrix by column scanning. $g(\cdot)$ is the inverse operator of $G(\cdot)$. $H_\varepsilon(z) = \frac{1}{2}[1 + \frac{2}{\pi}\arctan(\frac{z}{\varepsilon})]$, $\delta_\varepsilon(z) = \frac{\varepsilon}{\pi(\varepsilon^2+z^2)}$.

To compute the associated Euler-Lagrange equation for the unknown level set function ψ, we keep c_1 and c_2 fixed, and minimize E with respect to ψ. Parameterizing the descent direction by artificial time $t \geq 0$, the evolution equation in $\psi(t, x, y, z)$ is:

$$\begin{aligned}\frac{\partial \psi}{\partial t} = \quad & \delta_\varepsilon(\psi)[\mu \cdot div[\frac{\nabla\psi}{|\nabla\psi|}] + \lambda_1|I - c_1|^2 \\ & -\lambda_2|I - c_2|^2] - g\{U_k\Sigma_k^{-1}U_k^T[G(\psi - \bar{\Psi})]\}\end{aligned}$$

4. APPLICATIONS TO MEDICAL IMAGERY

We have used our model on various medical images, with multiple types of contours and shapes. In Fig. 5, we show the segmentation of the left and right ventricles using our level set prior 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. In Fig. 6, we show that our model can detect multiple objects of similar intensities and with blurred boundaries. We also tested our method in 3D images. Fig. 7 shows the segmentation of the left amygdala and hippocampus in a 3D MR brain image.

To validate the segmentation results, we compute the undirected distance (defined in Section 2.2) between the computed boundary and a manually traced boundary. For our experiments, the mean distances show improvement in all these cases comparing with/without the level set based prior: average left and right ventricles $1.6mm/5.8mm$, subcortical structures $1.8mm/9.2mm$, amygdala and hippocampus $2.3mm/11.2mm$. Virtually all the boundary points obtained using our model lie within one or two voxels of the manual segmentation.

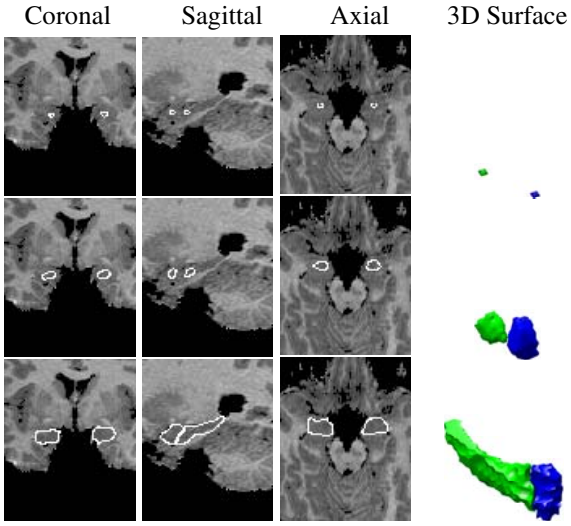

Fig. 7. Initial, middle, and final steps in the segmentation of 2 shapes in a brain MRI. Three orthogonal slices and the 3D surfaces are shown for each step. The training set consists of 12 MR images.

5. CONCLUSION

This work presents a level set based deformable model for the segmentation of multiple objects from 3D medical images using shape prior constraints. We evaluate this level set distribution model by comparing it with the traditional point distribution model. Our approach to multiple objects segmentation is based on a MAP estimation framework using level set based prior information of the objects in the image. Only one level set function is employed as the representation of the multiple objects of interest within the image. We define the probability distribution over the variations of objects contained in a set of training images. We found the algorithm to be computationally efficent, robust to noise, able to handle multidimensional data, and avoids the need for point correspondences during the training phase.

6. REFERENCES

[1] M. Leventon, E. Grimson, and O. Faugeras, *Statistical shape influence in geodesic active contours*, CVPR(2000) 316–323.

[2] A. Tsai, W. Wells, C. Tempany, E. Grimson, and A. Willsky, *Coupled Multi-shape Model and Mutual Information for Medical Image Segmentation*, IPMI(2003) 185–197.

[3] J. Yang, L. Staib and J. Duncan, *Neighbor-Constrained Segmentation with 3D Deformable Models*, IPMI(2003) 198–209.

[4] T.F. Cootes, A. Hill, C.J. Taylor, and J. Haslam, *Use of active shape models for locating structures in medical images*, Image and Vision Computing, 12(6):355-365, July 1994.

[5] T. Chan, L. Vese, *Active Contours Without Edges*, IEEE Trans. on Image Processing, vol.**10** No. **2** (2001) 266–277