A Shape-Based Approach to the Segmentation of Medical Imagery Using Level Sets

Andy Tsai*, Anthony Yezzi, Jr., William Wells, Clare Tempany, Dewey Tucker, Ayres Fan, W. Eric Grimson, and Alan Willsky

***Abstract*—We propose a shape-based approach to curve evolution for the segmentation of medical images containing known object types. In particular, motivated by the work of Leventon, Grimson, and Faugeras [15], we derive a parametric model for an implicit representation of the segmenting curve by applying principal component analysis to a collection of signed distance representations of the training data. The parameters of this representation are then manipulated to minimize an objective function for segmentation. The resulting algorithm is able to handle multidimensional data, can deal with topological changes of the curve, is robust to noise and initial contour placements, and is computationally efficient. At the same time, it avoids the need for point correspondences during the training phase of the algorithm. We demonstrate this technique by applying it to two medical applications; two-dimensional segmentation of cardiac magnetic resonance imaging (MRI) and three-dimensional segmentation of prostate MRI.**

***Index Terms*—Active contours, binary image alignment, cardiac MRI segmentation, curve evolution, deformable model, distance transforms, eigenshapes, implicit shape representation, medical image segmentation, parametric shape model, principal component analysis, prostate segmentation, shape prior, statistical shape model.**

I. Introduction

MEDICAL image segmentation algorithms often face difficult challenges such as poor image contrast, noise, and missing or diffuse boundaries. For example, tissue boundaries in medical images may be smeared (due to patient movements), missing (due to low SNR of the acquisition apparatus), or nonexistence (when blended with similar surrounding tissues). Under such conditions, without a prior model to constrain the segmentation, most algorithms (including intensity- and curve-based techniques) fail-mostly due to the under-determined nature of the segmentation process. Similar problems arise in other imaging applications as well and they also hinder the segmentation of the image. These image segmentation problems demand the incorporation of as much prior information as possible to help the segmentation algorithms extract the tissue of interest. We propose such an algorithm in this paper. In particular, we derive a model-based, implicit parametric representation of the segmenting curve and calculate the parameters of this representation via gradient descent to minimize an energy functional for medical image segmentation.[1]

Manuscript received September 19, 2001; revised October 11, 2002. This work was supported in part by the Office of Naval Research (ONR) under Grant N00014-00-1-0089, in part by the Air Force Office of Scientific Research (AFOSR) under Grant F49620-98-1-0349, in part by the National Science Foundation (NSF) under an ERC Grant through Johns Hopkins Agreement 8810274, in part by the National Institutes of Health (NIH) under Grant 1P41RR13218 and NIH R01 Grant AG 19513-01. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was J. Duncan. *Asterisk indicates corresponding author.*

*A. Tsai is with the Massachusetts Institute of Technology, Laboratory for Information and Decision Systems, Department of Electrical Engineering, Room 35-427, 127 Massachusetts Ave., Cambridge, MA 02139 USA (e-mail: atsai@mit.edu).

A. Yezzi, Jr. is with the School of Electrical and Computer Engineering; Georgia Institute of Technology, Atlanta, GA 30332 USA.

W. Wells is with Brigham and Women's Hospital/ Harvard Medical School, Boston, MA 02115 USA, and the Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

C. Tempany is with Brigham and Women's Hospital/Harvard Medical School, Boston, MA 02115 USA.

D. Tucker, A. Fan and A. Willsky are with the Laboratory for Information and Decision Systems; Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

W. E. Grimson is with the Artificial Intelligence Laboratory; Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Digital Object Identifier 10.1109/TMI.2002.808355

A. Relationship to Prior Work

Our work shares common aspects with a number of model-based image segmentation algorithms in the literature. Chen *et al.* [6] employed an "average shape" to serve as the shape prior term in their geometric active contour model. Cootes *et al.* [10] 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. The shape and pose parameters of this point distribution model are determined to match the points to strong image gradients. Pentland and Sclaroff [21] later described a variant of this approach. Staib and Duncan [23] introduced a parametric point model based on an elliptic Fourier decomposition of the landmark points. The parameters of their curve are calculated to optimize the match between the segmenting curve and the gradient of the image. Chakraborty *et al.* [4] extended this approach to a hybrid segmentation model that incorporates both gradient and region-homogeneity information. More recently, Wang and Staib [30] developed a statistical point model for the segmenting curve by applying principal component analysis (PCA) to the covariance matrices that capture the statistical variations of the landmark points. They formulated their edge-detection and correspondence-determination problem in a maximum *a posteriori* Bayesian framework. Image gradient is used within that framework to calculate the pose and shape parameters that describes their segmenting curve. Leventon *et al.* [15] proposed a less restrictive model-based segmenter. They incorporated shape information as a prior model to restrict the flow of the geodesic active contour [3], [32]. Their prior parametric shape model is

[1] A preliminary conference paper based on this work can be found in [26].

derived by performing PCA on a collection of signed distance maps of the training shape. The segmenting curve then evolves according to two competing forces: 1) the gradient force of the image, and 2) the force exerted by the estimated shape where the parameters of the shape are calculated based on the image gradients and the current position of the curve.

Our work is also closely related to region-based active contour models [5], [20], [22], [34]. In general, these region-based models enjoy a number of attractive properties over gradient-based techniques for segmentation, including greater robustness to noise (by avoiding derivatives of the image intensity) and initial contour placement (by being less local than most edge-based approaches).

B. Contributions of Our Work

In our algorithm, we adopt the implicit representation of the segmenting curve proposed in [15] and calculate the parameters of this implicit model to minimize the region-based energy functionals proposed in [5] and [34] for image segmentation. The resulting algorithm is found to be computationally efficient and robust to noise (since the evolving curve has limited degrees of freedom), has an extended capture range (because the segmentation functional is region-based instead of edge-based), and does not require point correspondences (due to an Eulerian representation of the curve). Though in this paper, we only show the development of our technique for two-dimensional (2-D) data, this algorithm can easily be generalized to handle multidimensional data. We demonstrate a three–dimensional (3-D) application of our technique in Section VI. Also, in this paper, we focus on using the region-based models presented in [5] and [34] . However, it is important to point out that other region-based models are equally applicable in this framework.

The rest of the paper is organized as follows. Section II describes a gradient-based approach to align all the training shapes in the database to eliminate variations in pose. Based on this aligned training set, we show in Section III the development of an implicit parametric representation of the segmenting curve. Section IV describes the use of this implicit curve representation in various region-based models for image segmentation. Section V provides a brief overview to illustrate how the various components mentioned above fit within the scope of our algorithmic framework. In Section VI, we show the application of this technique to two medical applications; the segmentation of the left ventricle from 2-D cardiac MRI and prostate gland segmentation from 3-D pelvic MRI. We conclude in Section VII with a summary and some possible future research directions of this work.

II. Shape Alignment

We begin our shape modeling process with the alignment of training shapes.[2] There have been a number of works dealing with the alignment of images [6], [8], [11], [17], [28], [29]. For our application, we are interested in aligning binary images since that is how we encode the training shapes. This greatly simplifies the alignment task, which we approach from a variational perspective.

A. Alignment Model

Let the training set $\mathcal{T}$ consist of a set of n binary images $\{I^1, I^2, \ldots, I^n\}$, each with values of one inside and zero outside the object. The goal is to calculate the set of pose parameters $\{\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^n\}$ used to jointly align the n binary images, and hence remove any variations in shape due to pose differences. We focus on using similarity transformations to align these binary images to each other. That is, in two dimensions, $\mathbf{p} = [a \;\; b \;\; h \;\; \theta]^T$ with a, b, h, and θ corresponding to x, y-translation, scale, and rotation, respectively. The transformed image of I, based on the pose parameter $\mathbf{p}$, is denoted by $\tilde{I}$, and is defined as

$$\tilde{I}(\tilde{x}, \tilde{y}) = I(x, y)$$

where

$$\begin{aligned}\begin{bmatrix}\tilde{x}\\ \tilde{y}\\ 1\end{bmatrix} &= T[\mathbf{p}]\begin{bmatrix}x\\ y\\ 1\end{bmatrix}\\ &= \underbrace{\begin{bmatrix}1 & 0 & a\\ 0 & 1 & b\\ 0 & 0 & 1\end{bmatrix}}_{M(a,b)}\underbrace{\begin{bmatrix}h & 0 & 0\\ 0 & h & 0\\ 0 & 0 & 1\end{bmatrix}}_{H(h)}\\ &\quad\times\underbrace{\begin{bmatrix}\cos(\theta) & -\sin(\theta) & 0\\ \sin(\theta) & \cos(\theta) & 0\\ 0 & 0 & 1\end{bmatrix}}_{R(\theta)}\begin{bmatrix}x\\ y\\ 1\end{bmatrix}.\end{aligned} \tag{1}$$

The transformation matrix $T[\mathbf{p}]$ is the product of three matrices: a translation matrix $M(a, b)$, a scaling matrix $H(h)$, and an in-plane rotation matrix $R(\theta)$. This transformation matrix $T[\mathbf{p}]$ maps the coordinates $(x, y) \in \mathbb{R}^2$ into coordinates $(\tilde{x}, \tilde{y}) \in \mathbb{R}^2$.

An effective strategy to jointly align the n binary images is to use gradient descent to minimize the following energy functional:

$$E_{\text{align}} = \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left\{\frac{\iint_\Omega (\tilde{I}^i - \tilde{I}^j)^2 dA}{\iint_\Omega (\tilde{I}^i + \tilde{I}^j)^2 dA}\right\} \tag{2}$$

where Ω denotes the image domain. Minimizing (2) is equivalent to simultaneously minimizing the difference between any pair of binary images in the training database. The area normalization term in the denominator of (2) is employed to prevent all the images from shrinking to improve the cost function.

The gradient of E_{align}, taken with respect to $\mathbf{p}^i$ for any i, is given by

$$\begin{aligned}\nabla_{\mathbf{p}^i} E_{\text{align}} = \sum_{\substack{j=1\\ j\neq i}}^{n}\Bigg\{&\frac{2\iint_\Omega (\tilde{I}^i - \tilde{I}^j)\nabla_{\mathbf{p}^i}\tilde{I}^i dA}{\iint_\Omega (\tilde{I}^i + \tilde{I}^j)^2 dA}\\ &- \frac{2\iint_\Omega (\tilde{I}^i - \tilde{I}^j)^2 dA \iint_\Omega (\tilde{I}^i + \tilde{I}^j)\nabla_{\mathbf{p}^i}\tilde{I}^i dA}{\left(\iint_\Omega (\tilde{I}^i + \tilde{I}^j)^2 dA\right)^2}\Bigg\}\end{aligned} \tag{3}$$

[2]Our method can take advantage of any alignment technique. We need to employ an alignment technique as a preprocessing step to allow us to capture shape variations in our database without interference from pose variations.

Fig. 1. Training data: 12 2-D binary shape models of the fighter jet.

Fig. 2. Alignment results of the above 12 2-D shape models of the fighter jet.

where $\nabla_{\mathbf{p}^i} \tilde{I}^i$ is the gradient of the transformed image $\tilde{I}^i$ taken with respect to the pose parameter $\mathbf{p}^i$. Using the chain rule, the lth component of $\nabla_{\mathbf{p}^i} \tilde{I}^i$ is given by

$$\nabla_{\mathbf{p}_l^i} \tilde{I}^i(\tilde{x},\tilde{y}) = \begin{bmatrix} \frac{\partial \tilde{I}^i(\tilde{x},\tilde{y})}{\partial \tilde{x}} & \frac{\partial \tilde{I}^i(\tilde{x},\tilde{y})}{\partial \tilde{y}} & 0 \end{bmatrix} \frac{\partial T[\mathbf{p}^i]}{\partial \mathbf{p}_l^i} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

where

$$\frac{\partial T[\mathbf{p}^i]}{\partial \mathbf{p}_1^i} = \frac{\partial T[\mathbf{p}^i]}{\partial a^i} = \frac{\partial M(a^i,b^i)}{\partial a^i} H(h^i) R(\theta^i) \tag{4a}$$

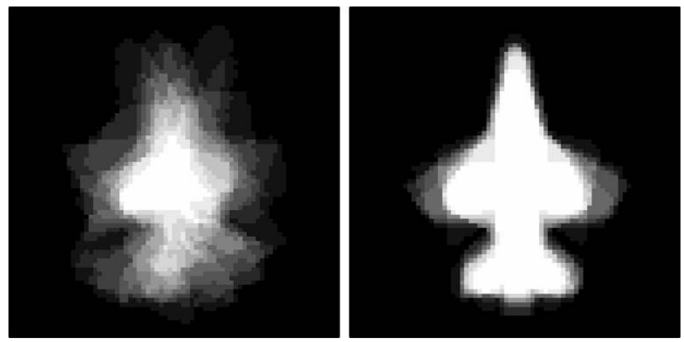
$$\frac{\partial T[\mathbf{p}^i]}{\partial \mathbf{p}_2^i} = \frac{\partial T[\mathbf{p}^i]}{\partial b^i} = \frac{\partial M(a^i,b^i)}{\partial b^i} H(h^i) R(\theta^i) \tag{4b}$$

$$\frac{\partial T[\mathbf{p}^i]}{\partial \mathbf{p}_3^i} = \frac{\partial T[\mathbf{p}^i]}{\partial h^i} = M(a^i,b^i) \frac{\partial H(h^i)}{\partial h^i} R(\theta^i) \tag{4c}$$

$$\frac{\partial T[\mathbf{p}^i]}{\partial \mathbf{p}_4^i} = \frac{\partial T[\mathbf{p}^i]}{\partial \theta^i} = M(a^i,b^i) H(h^i) \frac{\partial R(\theta^i)}{\partial \theta^i}. \tag{4d}$$

The matrix derivatives in (4) are taken componentwise. Since the solution of this alignment problem is under-determined, we regularize the problem by keeping the initial pose of one of the shapes fixed and calculating the pose parameters for the remaining shapes using the above approach. The initial poses of the training shapes in $\mathcal{T}$ are employed as the starting point for the alignment process and gradient descent is performed until convergence.

To illustrate this alignment process, a training set, consisting of 12 binary representations of fighter jets, is shown in Fig. 1. In this example, the pose parameter of the fighter jet at the far left side of the figure is chosen to be fixed, i.e., $\mathbf{p} = [0\ \ 0\ \ 1\ \ 0]^T$. The aligned version of this data set is shown in Fig. 2. Note that all the aligned fighter jets share roughly the same center, are pointing in the same direction, and are approximately equal in size. One way to judge the effectiveness of this alignment process is to assess the amount of overlap between the shapes within the database before and after the alignment process. The prealignment overlap image, shown in Fig. 3(a), is generated by stacking together all the binary fighter jets within the database prior to alignment (i.e., the fighter jets shown in Fig. 1), and adding them together in a pixelwise fashion. The postalignment overlap image, shown in Fig. 3(b), is generated in a similar fashion except that the binary fighter jets used to calculate the overlap image have already been aligned. Specifically, the fighter jets used in this case are the ones shown in Fig. 2. By comparing the two overlap images, there is a dramatic increase in the amount of overlap between the shapes after the alignment process suggesting that this method is an effective alignment technique.



(a) (b)

Fig. 3. Comparison of the amount of shape overlap in the "fighter" database (a) before alignment and (b) after alignment.

B. Multiresolution Alignment

The nature of the gradient descent approach we just described allows for only infinitesimal updates of the pose parameters, thus giving rise to slow convergence properties and increased sensitivity to local minima. These unattractive features are especially evident when trying to align large and complicated objects. One standard extension to enhance alignment algorithms is to utilize a multiresolution approach. The basic idea behind this approach is to employ a coarsened representation of the training set to obtain a good initial estimate of the pose parameters. We then progressively refine these pose estimates as the resolution of the objects is increased.

Specifically, given a set of training objects, we repeatedly subsample all the objects within the training set by a factor of two in each axis direction to obtain a collection of training sets with varying resolutions. Initial alignment is performed on the coarsest resolution set of objects to obtain a good initial estimate of the pose parameters. Operating at such a coarse scale, we reduce the number of updates required for alignment (since the domain of the image is reduced) and the sensitivity of the algorithm to local minima (by allowing the parameter search to be less local). More importantly though, the computational burden of alignment at each gradient step is substantially reduced, mostly due to the decreased computational cost associated with calculating (3) on a coarser grid. The pose parameters estimated on this coarsened set of training objects are appropriately scaled to serve as the starting pose estimates for the next

Fig. 4. Training data: 12 2-D binary shape models of the number four with size of 200 × 200 pixels.

Fig. 5. Lowest resolution representation of the above training data with size of 50 × 50 pixels.

Fig. 6. Alignment results of the above 50 × 50 shape models of the number four.

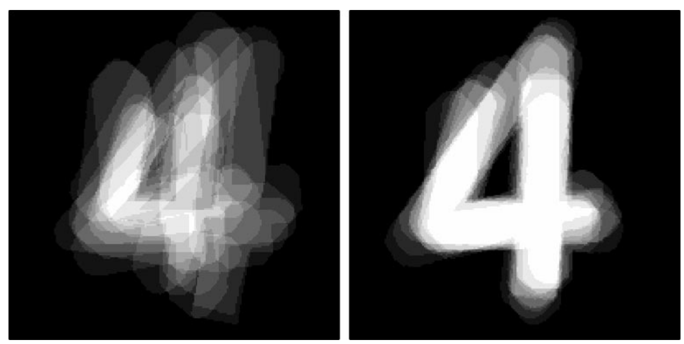
Fig. 7. Coarse-to-fine multiresolution refinement results of the 200 × 200 shape models of the number four.

higher resolution set of objects.[3] By providing a good starting estimate of the pose parameters at this new scale, only a small number of updates are required for convergence. This process of using the pose estimate at one resolution as the starting pose for the next finer resolution is repeated until the finest resolution set of objects is reached. To illustrate this multiresolution approach, we show in Fig. 4 a set of 12 binary representations of the number four. The fours are difficult objects to align due to the complicated structure of these objects. Fig. 5 shows this same data set with each shape down sampled by a factor of four in each direction. Initially, we align the fours in this reduced image domain. The results of this alignment are shown in Fig. 6. Next, we appropriately scale the pose parameters to serve as the starting pose for the next higher resolution. We continue this process until the finest resolution training set is reached. The final alignment results are shown in Fig. 7. Fig. 8 shows the prealignment and postalignment overlap images of the number four to visually demonstrate the effectiveness of this alignment process.

III. Implicit Parametric Shape Model

As mentioned earlier, a popular and natural approach to represent shapes is via point models where a set of marker points is used to describe the boundaries of the shape. This approach suffers from problems such as numerical instability, inability to accurately capture high curvature locations, difficulty in handling topological changes, and the need for point correspondences. To overcome these problem, we utilize an Eulerian approach to shape representation based on the level set methods of Osher and Sethian [19].

[3]Only the translational components of the pose are scaled up. The scaling and rotational components of the pose remain fixed.



(a) (b)

Fig. 8. Comparison of the amount of shape overlap in the "four" database (a) before alignment and (b) after alignment.

A. Shape Parameters

Following the lead of [15] and [19], we choose the signed distance function[4] as our representation for shape. In particular, the boundaries of each of the n aligned shapes in the database[5] are embedded as the zero level set of n separate signed distance functions $\{\Psi_1, \Psi_2, \ldots, \Psi_n\}$ with negative distances assigned to the inside and positive distances assigned to the outside of the object. Using the technique developed in [15], we compute $\bar{\Phi}$, the mean level set function of the shape database, as the average of these n signed distance functions, $\bar{\Phi} = (1/n) \sum \Psi_i$. To extract the shape variabilities, $\bar{\Phi}$ is subtracted from each of the n signed distance functions to create n mean-offset functions $\{\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_n\}$. These mean-offset functions are then used to capture the variabilities of the training shapes.

[4]The signed distance $\Psi(p)$ from an arbitrary point p to a known surface $\mathcal{Z}$ is the distance between p and the closest point z in $\mathcal{Z}$, multiplied by 1 or -1, depending on which side of the surface p lies in [1].

[5]The shapes in the database are aligned by employing the method presented in Section II.

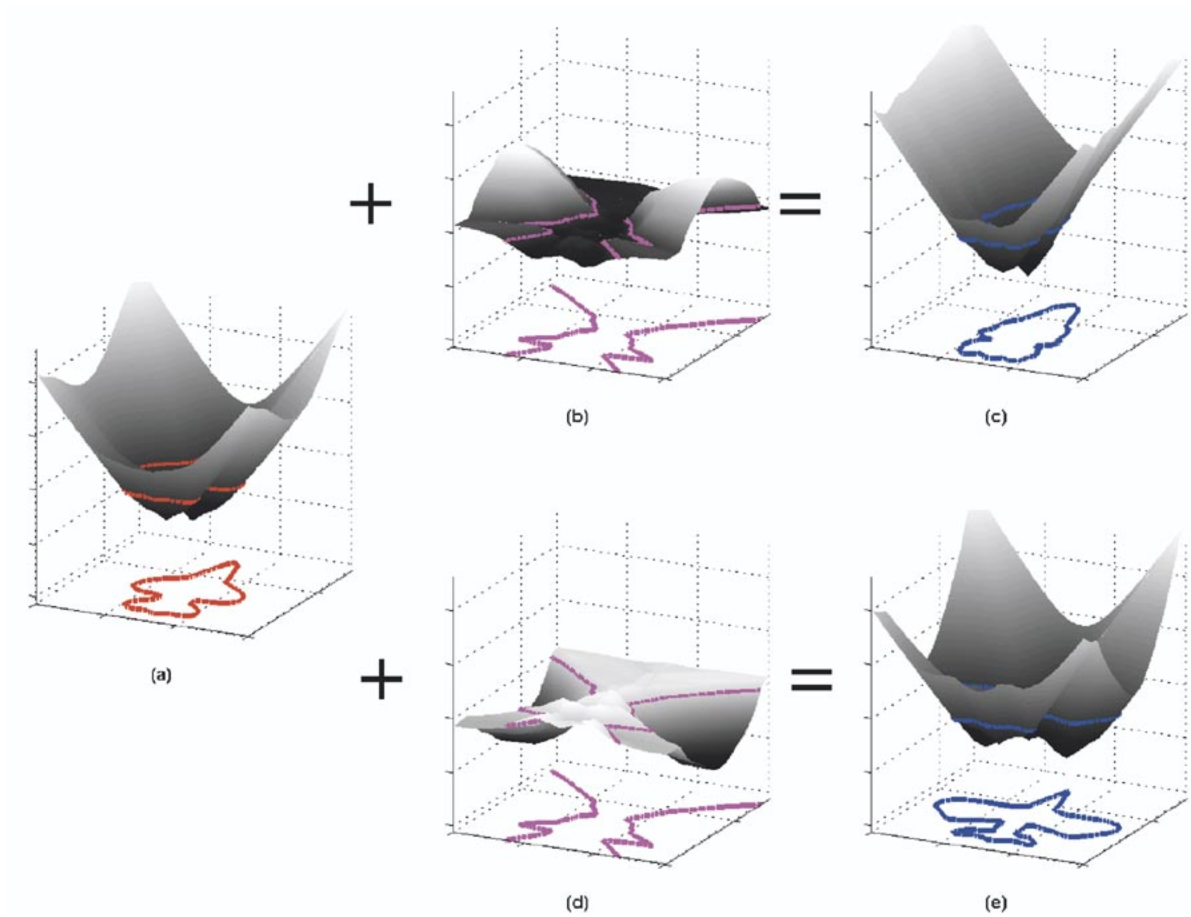


Fig. 9. Three-dimensional visualization of the fighter jet shape variability. (a) The mean level set function $\bar{\Phi}$. (b) Three-dimensional illustration of $+1\sigma_1\Phi_1$. (c) Level set of $+1\sigma$ variation of the first principal mode. (d) Three-dimensional illustration of $-1\sigma_1\Phi_1$. (e) Level set of -1σ variation of the first principal mode.

Specifically, we form n column vectors, $\tilde{\psi}_i$, consisting of N samples of each $\tilde{\Psi}_i$ (using identical sample locations for each function). The most natural sampling strategy is to utilize the $N_1 \times N_2$ rectangular grid of the training images to generate $N = N_1 \times N_2$ lexicographically ordered samples (where the columns of the image grid are sequentially stacked on top of one other to form one large column). Next, define the shape-variability matrix $\mathcal{S}$ as

$$\mathcal{S} = [\ \tilde{\psi}_1 \quad \tilde{\psi}_2 \quad \ldots \quad \tilde{\psi}_n\ ].$$

An eigenvalue decomposition is employed to factor $(1/n)\mathcal{S}\mathcal{S}^T$ as

$$\frac{1}{n}\mathcal{S}\mathcal{S}^T = U\Sigma U^T \tag{5}$$

where U is an $N \times n$ matrix whose columns represent the n orthogonal modes of variation in the shape and Σ is an $n \times n$ diagonal matrix whose diagonal elements represent the corresponding nonzero eigenvalues. The N elements of the ith column of U, denoted by U_i, are arranged back into the structure of the $N_1 \times N_2$ rectangular image grid (by undoing the earlier lexicographical concatenation of the grid columns) to yield Φ_i, the ith principal mode or eigenshape. Based on this approach, a maximum of n different eigenshapes $\{\Phi_1, \Phi_2, \ldots, \Phi_n\}$ are generated.

Note that in most cases, the dimension of the matrix $(1/n)\mathcal{S}\mathcal{S}^T$ is large $(N \times N)$ so the calculation of the eigenvectors and eigenvalues of this matrix is computationally expensive. In practice, the eigenvectors and eigenvalues of $(1/n)\mathcal{S}\mathcal{S}^T$ can be efficiently computed from a much smaller $n \times n$ matrix $\mathcal{W}$ given by

$$\mathcal{W} = \frac{1}{n}\mathcal{S}^T\mathcal{S}.$$

It is straightforward to show that if $\mathbf{d}$ is an eigenvector of $\mathcal{W}$ with corresponding eigenvalue λ, then $\mathcal{S}\mathbf{d}$ is an eigenvector of $(1/n)\mathcal{S}\mathcal{S}^T$ with eigenvalue λ (see [14] for a proof).

Let $k \leq n$, which is selected prior to segmentation, be the number of modes to consider. Choosing the appropriate k in our model is difficult and beyond the scope of this paper. Suffice it to say that k should be chosen large enough to be able to capture the prominent shape variations present in the training set, but not too large that the model begins to capture intricate details particular to a certain training shape.[6] In all of our examples, we chose k empirically. We now introduce a new level set function

$$\Phi[\mathbf{w}] = \bar{\Phi} + \sum_{i=1}^{k} w_i \Phi_i \tag{6}$$

[6]One way to choose the value of k is by examining the eigenvalues of the corresponding eigenvectors. In some sense, the size of each eigenvalue indicates the amount of influence or importance its corresponding eigenvector has in determining the shape. Perhapes by looking at a historgram of the eigenvalues, one can determine the threshold for determining the value of k. However, this approach would be difficult to implement as the threshold value for k varies for each application. In any case, there is no universal k that can be set.

where $\mathbf{w} = \{w_1, w_2, \ldots, w_k\}$ are the weights for the k eigenshapes with the variances of these weights $\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2\}$ given by the eigenvalues calculated earlier. We propose to use this newly constructed level set function Φ as our implicit representation of shape. Specifically, the zero level set of Φ describes the shape with the shape's variability directly linked to the variability of the level set function. Therefore, by varying $\mathbf{w}$, we vary Φ which indirectly varies the shape. Note that the shape variability we allow in this representation is restricted to the variability given by the k eigenshapes.

Fig. 9 provides some intuition as to how the level set representation of (6) captures shape variability. The set of 12 fighter jets shown in Fig. 2 is used as the shape training set to obtain $\{\bar{\Phi}, \Phi_1, \Phi_2, \ldots, \Phi_{12}\}$ and $\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_{12}^2\}$. Fig. 9(a) shows the mean level set function $\bar{\Phi}$ with the red curve outlining the zero level set of $\bar{\Phi}$. Fig. 9(b) shows the function $+1\sigma_1\Phi_1$ with the magenta curve outlining the zero crossings of this function. Notice that most of the spatial variations associated with this function lie in the area corresponding to the wings of the fighter jet. Specifically, a large rising "hump" can be seen in those areas. When this function is added to $\bar{\Phi}$, a new level set representation of the fighter jet is obtained. This new level set function is shown in Fig. 9(c) with the blue curve outlining the zero level set. As expected, adding $+1\sigma_1\Phi_1$ to $\bar{\Phi}$ causes the wing size to shrink, thus yielding a new fighter jet with a much smaller wing span. In Fig. 9(d), we show the function $-1\sigma_1\Phi_1$ with the magenta curve outlining the zero crossings of this function. This is simply the negative of Fig. 9(b) and hence adding this function to $\bar{\Phi}$ causes the wing span of the fighter jet to increase. This resulting level set function is illustrated in Fig. 9(e) with the blue curve outlining the zero level set. To further illustrate the parametric shape encoding scheme of (6), we show in Fig. 10 the mean shape of the fighter jet as well as its shape variations based on varying its first three principal modes by $\pm 1\sigma$. As another demonstration, we employ the set of training shapes shown in Fig. 7 to obtain an implicit parametric representation of the number four. Fig. 11 shows the mean shape of the number four as well as its shape variations based on varying its first three principal modes by $\pm 1\sigma$. Notice that by varying the first two principal modes, the shape of the number four changes topology going from two curves to one curve. This is an additional advantage of using the Eulerian framework for shape representation as it can handle topological changes in a seamless fashion. This ability is of value for biomedical applications. One such application is the tracking of changes in multiple sclerosis lesions over time (as they shrink, migrate, split, disappear, etc.). Another is in the segmentation of the pancreas which often presents as one solid organ. But at times, the pancreas does not fuse in utero and hence presents as two separate lobes which may require segmentation algorithms that can deal with topology changes. Another application might be in segmenting skin lesions. Some skin pathologies can present both as one confluent lesion or as an island of lesions.

B. Pose Parameters

At this point, our implicit representation of shape cannot accommodate shape variabilities due to differences in pose. To have the flexibility of handling pose variations, $\mathbf{p}$ is added as another parameter to the level set function of (6). With this new addition, the implicit description of shape is given by the zero level set of the following function:

$$\Phi[\mathbf{w}, \mathbf{p}](x, y) = \bar{\Phi}(\tilde{x}, \tilde{y}) + \sum_{i=1}^{k} w_i \Phi_i(\tilde{x}, \tilde{y}) \tag{7}$$

where

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ 1 \end{bmatrix} = T[\mathbf{p}] \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}.$$

with $T[\mathbf{p}]$ defined earlier in (1). The addition of $\mathbf{p}$ to our parametric shape model enables us to accomodate a larger class of objects. In particular, the model can now handle object shapes that may differ from each other in terms of scale, orientation, or center location. In Section IV, we describe how $\mathbf{w}$ and $\mathbf{p}$ are optimized, via coordinate descent, for image segmentation.

IV. Region-Based Models for Segmentation

In region-based segmentation models [5], [20], [22], [34], the evolution of the segmenting curve depends upon the pixel intensities within entire regions. That is, region-based models regard an image as the composition of a finite number of regions and rely on regional statistics for segmentation. The statistics of entire regions (such as sample mean and variance) are used to direct the movement of the curve toward the boundaries of the image. This is in sharp contrast to edge-based segmentation models [2], [3], [9], [12], [13], [16], [24], [25], [32] where the evolution of the curve depends strictly on nearby pixel intensities (i.e., gradient information). As a result, region-based models are more global than edge-based models. Furthermore, because of the global nature of region-based models, these models do not require the use of inflationary terms commonly employed by edge-based techniques to drive the curve toward image boundaries. Region-based models are also more robust to noise since they do not employ gradient operators, which are inherently sensitive to noise, to explicitly detect the location of edges. In this section, we present three recently developed region-based models for segmentation and describe how these models fit within the scope of our shape-based curve evolution framework. Specifically, in this section, we present the Chan-Vese model, the binary mean model, and the binary variance model for image segmentation. However, instead of deriving the evolution equation for the curves used to segment the image (which is the original design of these models), we derive gradient descent equations used to optimize the shape and pose parameters that indirectly describe the segmenting curve.

A. Description of the Models

We begin with a simple synthetic example to present how region-based segmentation models are incorporated into our model-based algorithm. Assume that the domain of the observed image I is formed by two regions distinguishable by some region statistic (e.g., sample mean or variance). We would like to segment this image via the curve $\vec{C}$, which in our framework, is represented by the zero level set of Φ, i.e.,

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

Moreover, as a result of this implicit parametric representation of $\vec{C}$, the regions inside and outside the curve, denoted, respectively, by R^u and R^v, are given by

$$R^u = \{(x,y) \in \mathbb{R}^2 : \Phi(x,y) < 0\}$$
$$R^v = \{(x,y) \in \mathbb{R}^2 : \Phi(x,y) > 0\}.$$

In our algorithmic framework, we calculate the parameters of $\Phi[\mathbf{w},\mathbf{p}]$ to vary $\vec{C}$ and hence segment the image I. These parameters, $\mathbf{w}$ and $\mathbf{p}$, are obtained by minimizing region-based energy functionals that are constructed using various image statistics. Some useful image statistics, written in terms of $\Phi[\mathbf{w},\mathbf{p}]$, are

$$\text{area in } R^u: A_u = \iint_\Omega \mathcal{H}(-\Phi[\mathbf{w},\mathbf{p}])dA$$
$$\text{area in } R^v: A_v = \iint_\Omega \mathcal{H}(\Phi[\mathbf{w},\mathbf{p}])dA$$
$$\text{sum intensity in } R^u: S_u = \iint_\Omega I\mathcal{H}(-\Phi[\mathbf{w},\mathbf{p}])dA$$
$$\text{sum intensity in } R^v: S_v = \iint_\Omega I\mathcal{H}(\Phi[\mathbf{w},\mathbf{p}])dA$$
$$\text{sum of squared intensity in } R^u: Q_u = \iint_\Omega I^2\mathcal{H}(-\Phi[\mathbf{w},\mathbf{p}])dA$$
$$\text{sum of squared intensity in } R^v: Q_v = \iint_\Omega I^2\mathcal{H}(\Phi[\mathbf{w},\mathbf{p}])dA$$
$$\text{average intensity in } R^u: \mu = \frac{S_u}{A_u}$$
$$\text{average intensity in } R^v: \nu = \frac{S_v}{A_v}$$
$$\text{sample variance in } R^u: \sigma_u^2 = \frac{Q_u}{A_u} - \mu^2$$
$$\text{sample variance in } R^v: \sigma_v^2 = \frac{Q_v}{A_v} - \nu^2$$

where the Heaviside function $\mathcal{H}$ is given by

$$\mathcal{H}(\Phi[\mathbf{w},\mathbf{p}]) = \begin{cases} 1, & \text{if } \Phi[\mathbf{w},\mathbf{p}] \geq 0 \\ 0, & \text{if } \Phi[\mathbf{w},\mathbf{p}] < 0. \end{cases}$$

Chan and Vese in [5], and Yezzi *et al.* in [34] proposed pure region-based models to segment I using these region statistics. Below, we provide descriptions of their models, describe the role of $\mathbf{w}$ and $\mathbf{p}$ in these models, and detail the optimization of these models with respect to $\mathbf{w}$ and $\mathbf{p}$ (instead of $\vec{C}$) for image segmentation. As detailed in Section III, by calculating the parameters $\mathbf{w}$ and $\mathbf{p}$ that optimize the segmentation energy functionals, we have implicitly determined the segmenting curve $\vec{C}$. Thus, our segmentation approach can be considered as a parameter optimization technique.

1) The Chan-Vese Model: Chan and Vese in [5] proposed the following energy functional for segmenting I:

$$E_{\text{cv}} = \int_{R^u} (I-\mu)^2 dA + \int_{R^v} (I-\nu)^2 dA$$

which is equivalent, (up to a term which does not depend upon the evolving curve), to the energy functional below

$$E_{\text{cv}} = -(\mu^2 A_u + \nu^2 A_v) = -\left(\frac{S_u^2}{A_u} + \frac{S_v^2}{A_v}\right). \quad (8)$$

The Chan-Vese energy functional E_{cv} can be viewed as a piecewise constant generalization of the Mumford-Shah functional [18]. Gradient descent is employed to search for the parameters $\mathbf{w}$ and $\mathbf{p}$ that minimize E_{cv} to implicitly determine the segmenting curve. The gradients of E_{cv}, taken with respect to $\mathbf{w}$ and $\mathbf{p}$, are given by

$$\nabla_\mathbf{w} E_{\text{cv}} = -2(\mu\nabla_\mathbf{w} S_u + \nu\nabla_\mathbf{w} S_v) + (\mu^2\nabla_\mathbf{w} A_u + \nu^2\nabla_\mathbf{w} A_v) \quad (9a)$$
$$\nabla_\mathbf{p} E_{\text{cv}} = -2(\mu\nabla_\mathbf{p} S_u + \nu\nabla_\mathbf{p} S_v) + (\mu^2\nabla_\mathbf{p} A_u + \nu^2\nabla_\mathbf{p} A_v). \quad (9b)$$

2) The Binary Mean Model: A different strategy was proposed by Yezzi *et al.* in [34] to segment I. They propose to evolve $\vec{C}$ so as to maximize the distance between μ and ν. A natural cost functional they employed is to minimize the following:

$$E_{\text{binary}} = -\frac{1}{2}(\mu-\nu)^2 = -\frac{1}{2}\left(\frac{S_u}{A_u} - \frac{S_v}{A_v}\right)^2. \quad (10)$$

The authors in [34] called this the *binary model* (since it is initially designed to segment images consisting of two distinct but constant intensity regions). Once again, gradient descent is employed to calculate the parameters $\mathbf{w}$ and $\mathbf{p}$ that minimize E_{binary} to implicitly determine the segmenting curve. The gradients of E_{binary}, taken with respect to $\mathbf{w}$ and $\mathbf{p}$, are given by

$$\nabla_\mathbf{w} E_{\text{binary}} = (\nu-\mu) \times \left(\frac{\nabla_\mathbf{w} S_u - \mu\nabla_\mathbf{w} A_u}{A_u} - \frac{\nabla_\mathbf{w} S_v - \nu\nabla_\mathbf{w} A_v}{A_v}\right) \quad (11a)$$
$$\nabla_\mathbf{p} E_{\text{binary}} = (\nu-\mu) \times \left(\frac{\nabla_\mathbf{p} S_u - \mu\nabla_\mathbf{p} A_u}{A_u} - \frac{\nabla_\mathbf{p} S_v - \nu\nabla_\mathbf{p} A_v}{A_v}\right). \quad (11b)$$

3) The Binary Variance Model: So far, we have focused on using the mean as the image statistic in differentiating the two regions in I. Other image statistics can also be used in a region-based segmentation model. For example, Yezzi *et al.* in [34] proposed a segmentation model based on image variances. Consider the following energy functional for segmentation:

$$E_{\text{variance}} = -\frac{1}{2}\left(\sigma_u^2 - \sigma_v^2\right)^2 = -\frac{1}{2}\left(\left(\frac{Q_u}{A_u} - \frac{Q_v}{A_v}\right) - (\mu^2 - \nu^2)\right)^2. \quad (12)$$

The design of this model is to partition an image into two regions, one of low variance and one of high variance, by maximally separating the sample variances inside and outside the curve. The gradients of E_{variance}, taken with respect to $\mathbf{w}$ and $\mathbf{p}$, are given by

$$\nabla_\mathbf{w} E_{\text{variance}} = \left(\sigma_v^2 - \sigma_u^2\right)\left(\nabla_\mathbf{w}\sigma_u^2 - \nabla_\mathbf{w}\sigma_v^2\right) \quad (13a)$$
$$\nabla_\mathbf{p} E_{\text{variance}} = \left(\sigma_v^2 - \sigma_u^2\right)\left(\nabla_\mathbf{p}\sigma_u^2 - \nabla_\mathbf{p}\sigma_v^2\right) \quad (13b)$$

where

$$\nabla_\mathbf{w}\sigma_u^2 = \frac{\left(\mu^2 - \sigma_u^2\right)\nabla_\mathbf{w} A_u - 2\mu\nabla_\mathbf{w} S_u + \nabla_\mathbf{w} Q_u}{A_u}$$
$$\nabla_\mathbf{w}\sigma_v^2 = \frac{\left(\nu^2 - \sigma_v^2\right)\nabla_\mathbf{w} A_v - 2\nu\nabla_\mathbf{w} S_v + \nabla_\mathbf{w} Q_v}{A_v}$$
$$\nabla_\mathbf{p}\sigma_u^2 = \frac{\left(\mu^2 - \sigma_u^2\right)\nabla_\mathbf{p} A_u - 2\mu\nabla_\mathbf{p} S_u + \nabla_\mathbf{p} Q_u}{A_u}$$
$$\nabla_\mathbf{p}\sigma_v^2 = \frac{\left(\nu^2 - \sigma_v^2\right)\nabla_\mathbf{p} A_v - 2\nu\nabla_\mathbf{p} S_v + \nabla_\mathbf{p} Q_v}{A_v}.$$

B. Gradients of Region Statistics

As shown in (9), (11), and (13), to update the shape and pose parameters via gradient descent, the gradients of region statistics A_u, A_v, S_u, S_v, Q_u, and Q_v, taken with respect to $\mathbf{w}$ and $\mathbf{p}$, are required. Defining the one-dimensional Dirac measure δ concentrated at zero by

$$\delta(z) = \frac{d}{dz}\mathcal{H}(z)$$

we can now express the ith component of each of the gradient terms in (9), (11), and (13) as line integrals along $\vec{C}$

$$\begin{aligned}
\nabla_{w_i} A_u &= -\nabla_{w_i} A_v = -\oint_{\vec{C}} \Phi_i ds \\
\nabla_{p_i} A_u &= -\nabla_{p_i} A_v = -\oint_{\vec{C}} \nabla_{p_i} \Phi ds \\
\nabla_{w_i} S_u &= -\nabla_{w_i} S_v = -\oint_{\vec{C}} I\Phi_i ds \\
\nabla_{p_i} S_u &= -\nabla_{p_i} S_v = -\oint_{\vec{C}} I\nabla_{p_i} \Phi ds \\
\nabla_{w_i} Q_u &= -\nabla_{w_i} Q_v = -\oint_{\vec{C}} I^2\Phi_i ds \\
\nabla_{p_i} Q_u &= -\nabla_{p_i} Q_v = -\oint_{\vec{C}} I^2\nabla_{p_i} \Phi ds
\end{aligned}$$

where

$$\nabla_{p_i}\Phi = \nabla_{p_i}\Phi(\tilde{x},\tilde{y}) = \begin{bmatrix} \frac{\partial\Phi(\tilde{x},\tilde{y})}{\partial\tilde{x}} & \frac{\partial\Phi(\tilde{x},\tilde{y})}{\partial\tilde{y}} & 0 \end{bmatrix} \frac{\partial T[p_i]}{\partial p_i} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

with $\partial T[p_i]/\partial p_i$ previously defined in (4).

C. Parameter Optimization Via Gradient Descent

The gradients of the various energy functionals taken with respect to $\mathbf{w}$ and $\mathbf{p}$ are given by (9), (11), and (13). For conciseness of notation, denote $\nabla_{\mathbf{w}}E$ and $\nabla_{\mathbf{p}}E$ as the gradients of any of the above energy functionals taken with respect to $\mathbf{w}$ and $\mathbf{p}$, respectively. With this introduction, the update equations for the shape and pose parameters in our gradient descent approach are given by

$$\begin{aligned}
\mathbf{w}^{(t+1)} &= \mathbf{w}^{(t)} - \alpha_{\mathbf{w}}\nabla\mathbf{w}E \\
\mathbf{p}^{(t+1)} &= \mathbf{p}^{(t)} - \alpha_{\mathbf{p}}\nabla\mathbf{p}E
\end{aligned}$$

where $\alpha_{\mathbf{w}}$ and $\alpha_{\mathbf{p}}$ are positive step-size paramters, and $\mathbf{w}^{(t)}$ and $\mathbf{p}^{(t)}$ denote the values of $\mathbf{w}$ and $\mathbf{p}$ at the tth iteration, respectively. The updated shape and pose parameters are then used to implicitly determine the updated location of the segmenting curve.

It is important to note that no special numerics were required in our proposed technique as it does not involve any partial differential equations. This results in fast and simple implementation of our methodology. In fact, this is one of the main departure between our model and the earlier one put forth by Leventon *et al.* [15]

D. Extension to Three Dimensions

The generalization of this algorithm to three dimensions is straightforward. The pose parameter is expanded to consist of seven terms: x, y, and z translation; pitch; yaw; roll; and magnification. The shape alignment strategy is to jointly align the n binary volumetric data via gradient descent. Signed distance function is similarly employed to represent the 3-D shapes. In particular, the bounding surfaces of each shape is embedded as the zero level set of a signed distance function with negative distances assigned to the inside and positive distances assigned to the outside of the 3-D object. The 3-D shape parameters are derived in a similar fashion as the 2-D shape parameters. However, these 3-D shape parameters implicitly describe a 3-D segmenting surface rather than a 2-D segmenting curve. The region statistics used in the region-based models for segmentation are now calculated over an entire volume rather than over an entire region.

E. Illustration of the Models Using Synthetic Data

Figs. 12–14 show the use of E_{cv}, E_{binary}, and E_{variance} for segmentation. We show in Fig. 12(a) a fighter jet (that is not part of the fighter jet database of Fig. 1). Fig. 12(b) shows the same fighter jet surrounded by horizontal and vertical line clutter. The presence of these lines creates missing edges in the fighter jet which can cause problems in conventional segmentation algorithms that do not rely on prior shape information. Fig. 12(c) shows this line-cluttered fighter jet image contaminated by additive Gaussian noise. The goal is to segment the fighter jet from this noisy test image. Knowing *a priori* that the object in the image is a fighter jet, we employ the database shown in Fig. 2 to derive an implicit parametric curve model for the fighter jet [in the form of (7)]. In this example, we use $k = 6$. The zero level set of $\bar{\Phi}$ is employed as the starting curve which is illustrated in Fig. 12(d). The parameters of the segmenting curve, $\mathbf{w}$ and $\mathbf{p}$, are calculated to minimize E_{cv}. Fig. 12(e) shows the final shape and position of the segmenting curve. Notice that we are able to successfully find the boundaries of the fighter jet without being distracted by the line clutter. In Fig. 13, we show a slight variant of the experiment just described. Specifically, a new fighter jet (which is also not part of the database of Fig. 1) is employed as the object in the test image, and E_{binary} is employed as the segmentation functional. Using the same $\{\bar{\Phi}, \Phi_1, \Phi_2, \ldots, \Phi_6\}$ as before, we are able to successfully segment this new object.

Fig. 14 shows a different experiment. The object in this experiment is the number four which is shown in Fig. 14(a). Vertical and horizontal lines are again added to this image to create missing edges in the object. The resulting line-cluttered image is shown in Fig. 14(b). This binary mask is used to create the variance image shown in Fig. 14(c) which consists of two regions, each of identical means but of different variances. The goal is to segment the object from this noisy test image. Knowing *a priori* that the object in the image is a handwritten four, we employ the database of fours, shown in Fig. 7, to obtain the mean shape and the eigenshapes for our implicit representation of the object. As before, we use $k = 6$. The zero level set of $\bar{\Phi}$ is employed as the starting curves as illustrated in Fig. 14(d). Notice in this figure that two curves are used to describe the starting shape. Because the image statistic that characterizes the two regions in this test image is variance, the parameters of the segmenting curve, $\mathbf{w}$ and $\mathbf{p}$, are calculated to minimize E_{variance}. Fig. 14(e) shows the successful segmentation of the number four image. Notice that without any additional effort, the two starting curves merged to form one single segmenting curve at the end.

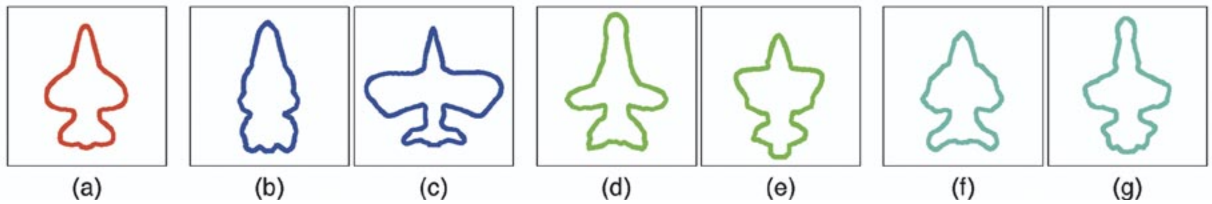


Fig. 10. Shape variability of the fighter jet. (a) The mean shape. (b) $+1\sigma$ variation of the first principal mode. (c) -1σ variation of the first principal mode. (d) $+1\sigma$ variation of the second principal mode. (e) -1σ variation of the second principal mode. (f) $+1\sigma$ variation of the third principal mode. (g) -1σ variation of the third principal mode. Grossly, the first three principal modes vary the shape and size of the wings as well as the length of the fighter jets.

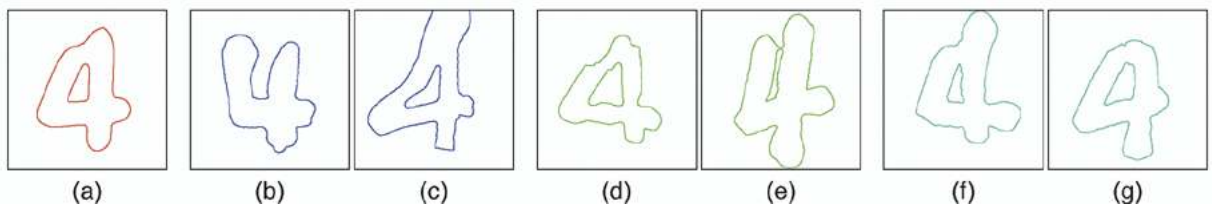


Fig. 11. Shape variability of the number four. (a) The mean shape. (b) $+1\sigma$ variation of the first principal mode. (c) -1σ variation of the first principal mode. (d) $+1\sigma$ variation of the second principal mode. (e) -1σ variation of the second principal mode. (f) $+1\sigma$ variation of the third principal mode. (g) -1σ variation of the third principal mode.

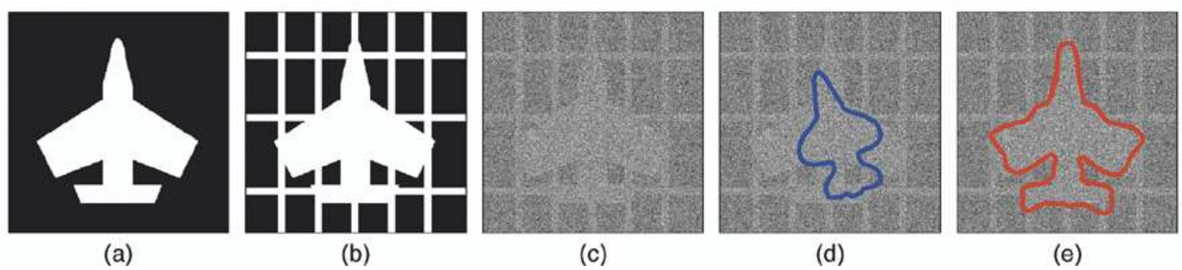


Fig. 12. Segmentation of a noisy fighter jet with missing edges using E_{cv}. (a) Original binary image. (b) Original binary image surrounded by line clutter. (c) Image in (b) with additive Gaussian noise. (d) Blue curve shows the initializing contour. (e) Red curve shows the final contour.

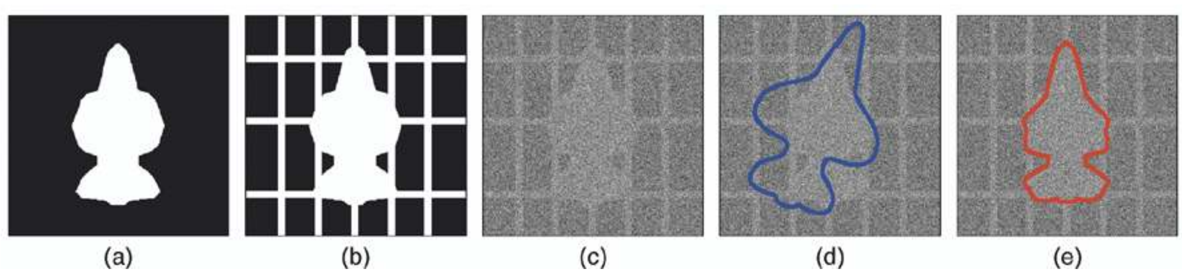


Fig. 13. Segmentation of a noisy fighter jet with missing edges using E_{binary}. (a) Original binary image. (b) Original binary image surrounded by line clutter. (c) Image in (b) with additive Gaussian noise. (d) Blue curve shows the initializing contour. (e) Red curve shows the final contour.

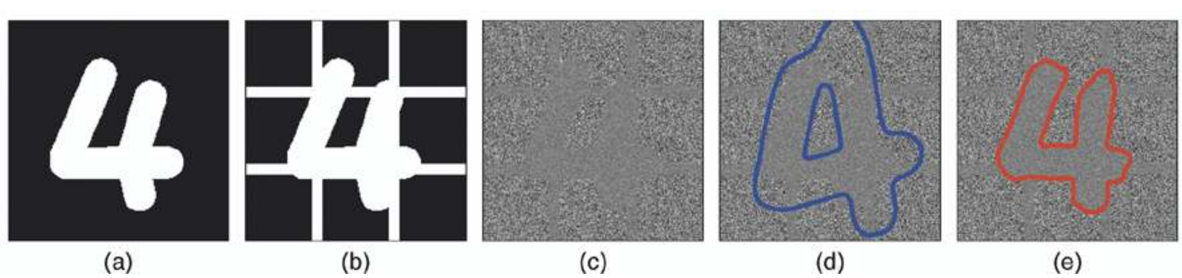


Fig. 14. Segmentation of a noisy number four with missing edges using E_{variance}. (a) Original binary image. (b) Original binary image surrounded by line clutter. (c) Image in (b) with additive Gaussian noise. (d) Blue curve shows the initializing contour. (e) Red curve shows the final contour.

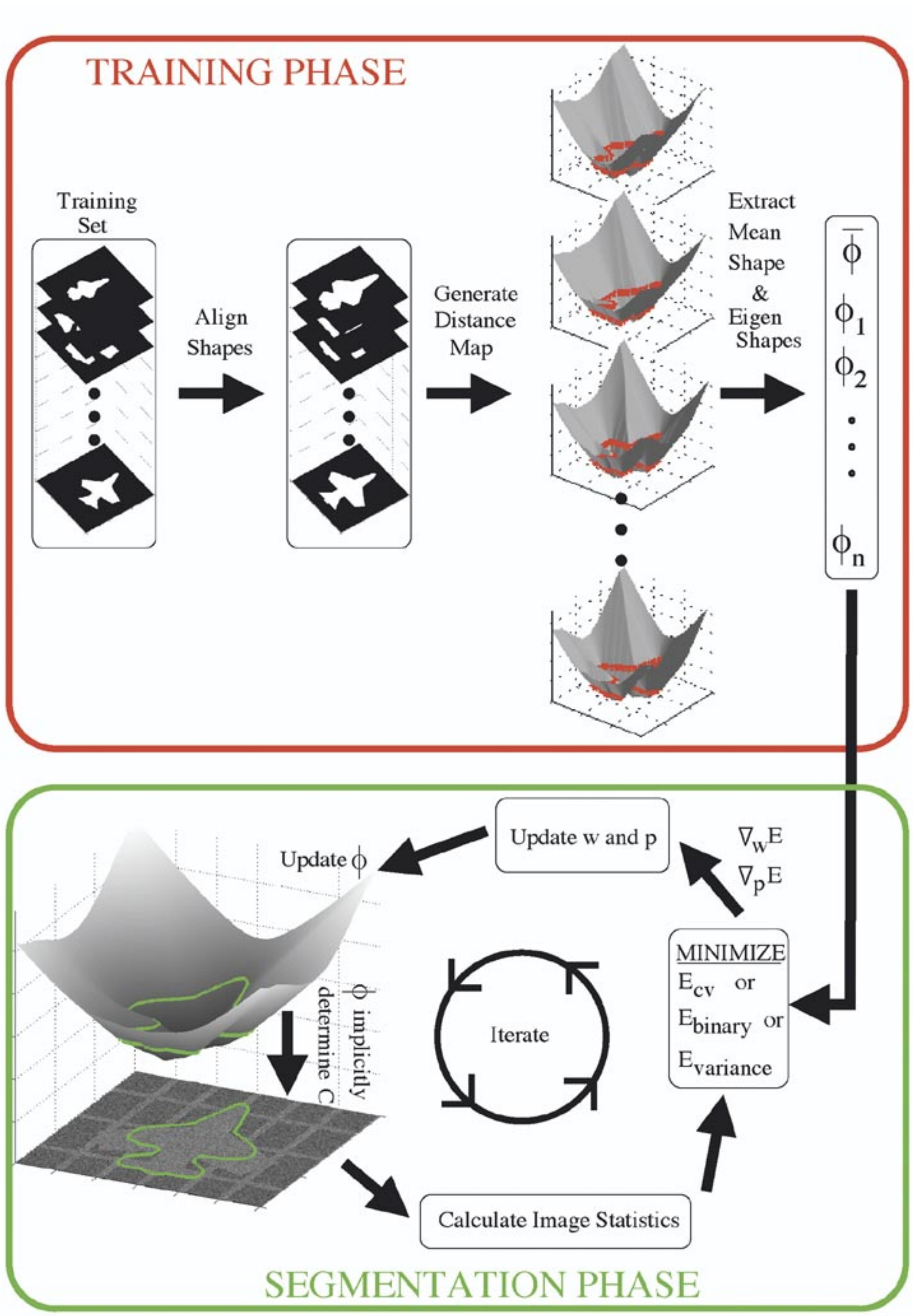


Fig. 15. A conceptual representation of our algorithmic framework. The top frame summarizes the training phase of our approach (Sections II and III). The bottom frame illustrates the segmentation phase of our algorithm (Section IV).

These figures demonstrate that our segmentation method is robust to the presence of clutter pixel-something that can not be said of many other segmentation algorithms. The reason for this is that our use of a finitely parameterized shape model makes the impact of such anomalous pixels much less significant than in other curve evolution or other segmentation methods.

V. Outline of the Algorithmic Framework

In this section, we provide a brief overview of our algorithmic framework. Fig. 15 shows a block diagram to illustrate how the different components described throughout this paper fit within the scope of our algorithmic framework. As illustrated in this diagram, our segmentation algorithm can be divided into two phases—a training phase and a segmentation phase. The training phase consists of shape alignment (described in Section II) and parametric shape modeling (described in Section III). Given a set of training shapes, gradient descent is employed to minimize the alignment model of (2) to jointly align them. Signed distance maps are generated to represent each of the shapes in the aligned database. By applying PCA to this collection of distance maps, we extract the mean shape

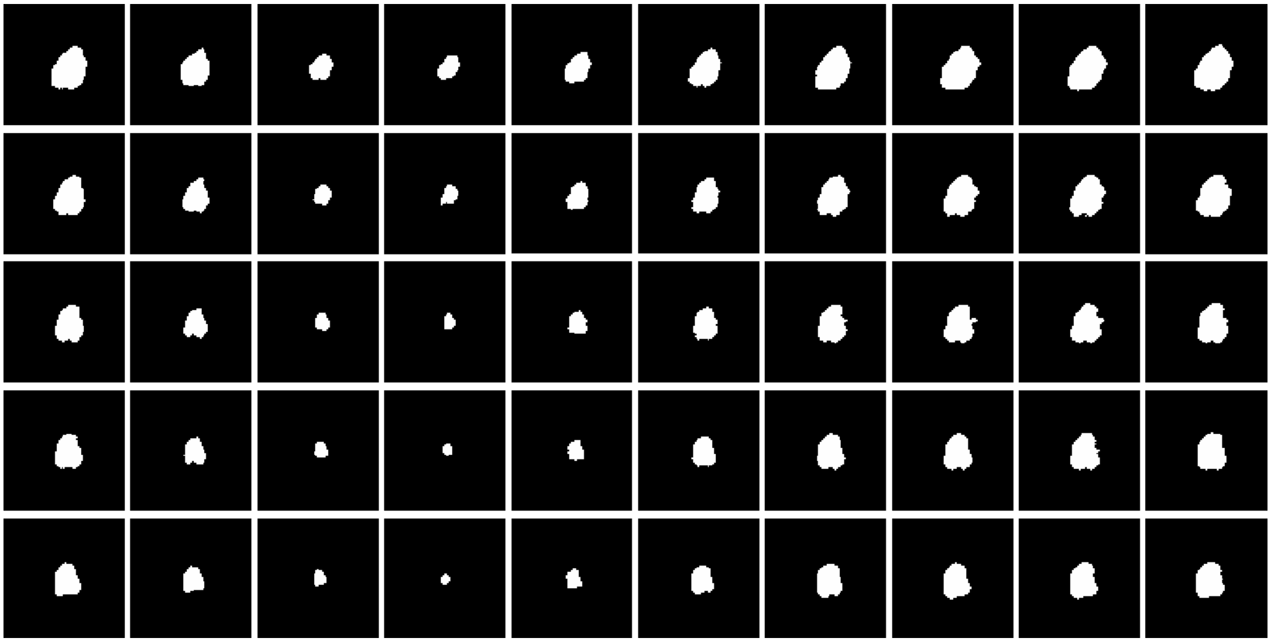


Fig. 16. Training data: 2-D binary shape models of the left ventricle based on human interactive segmentations of different spatial and temporal slices of a patient's cardiac MRI.

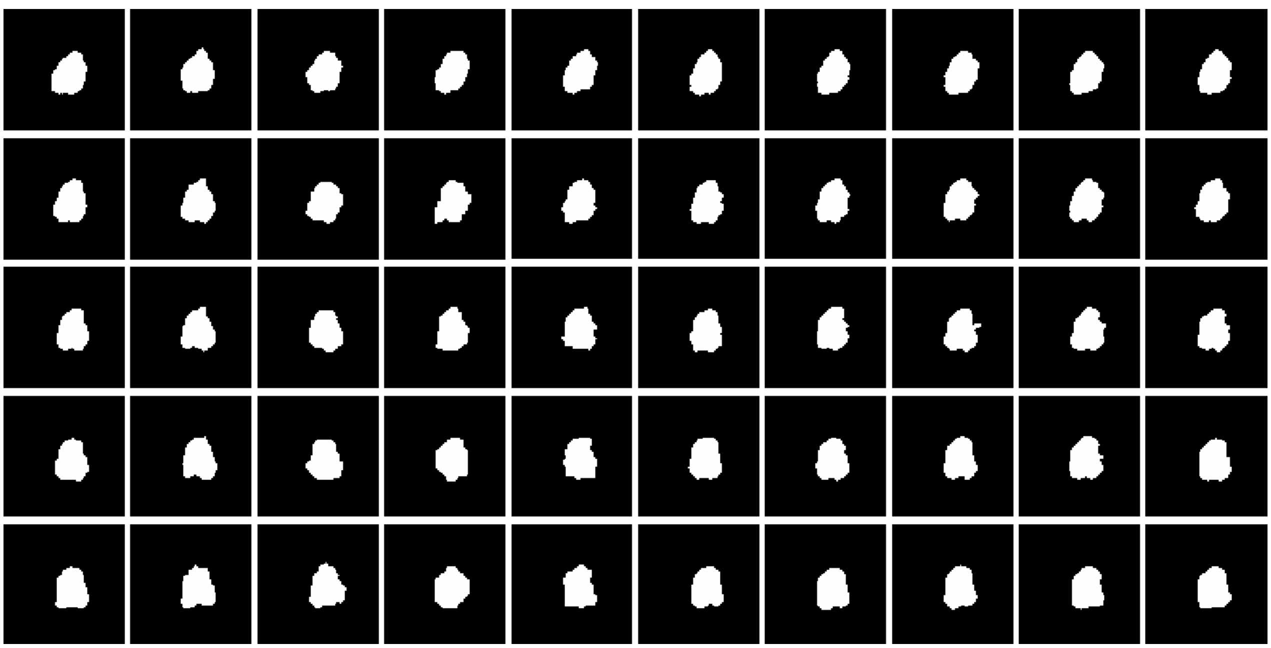
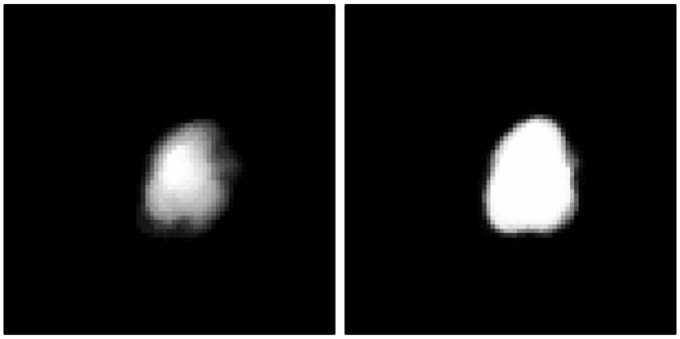


Fig. 17. Alignment results of the 50 2-D binary shape models of the left ventricle.

and the eigenshapes particular to this shape database. The mean shape and the eigenshapes are used to form the implicit parametric shape representation described in (7). The next part of our algorithm, the segmentation phase (described in Section IV), involves calculating $\mathbf{w}$ and $\mathbf{p}$, the parameters of our implicit shape representation, to minimize a segmentation functional. This minimization is performed as an iterative process using gradient descent. At each gradient step, $\mathbf{w}$ and $\mathbf{p}$ are updated to generate a new level set $\Phi[\mathbf{w}, \mathbf{p}]$. The segmenting curve $\vec{C}$ is implicitly determined by this new level set. Based on the new position and shape of $\vec{C}$, we recalculate the image statistic inside and outside the curve. These newly computed statistics are used in the segmentation functional to determine the update rules for $\mathbf{w}$ and $\mathbf{p}$. We continue this iterative scheme until convergence is reached for segmentation.

VI. Applications to Medical Imagery

We now apply the model-based curve evolution technique derived in this paper to two medical applications. Section VI-A illustrates a 2-D example (cardiac MRI segmentation), while Section VI-B illustrates a 3-D example (prostate gland segmentation from pelvic MRI).



(a) (b)

Fig. 18. Comparison of the amount of shape overlap in the cardiac database (a) before alignment and (b) after alignment.

A. A 2-D Example: Left Ventricle Segmentation of Cardiac MRI

Cardiac MRI is an important clinical tool used to provide four–dimensional (4-D) (temporal as well as spatial) information about the heart. Typically, one study generates 80–120 2-D images of a patient's heart. In a variety of clinical scenarios (such as assessing cardiac function and diagnosing cardiac diseases), it is important to extract the boundaries of the left ventricle from this data set. For example, the segmentation of the left ventricle is a prerequisite in calculating important physiological parameters such as ejection fraction and stroke volume. Manual tracing of the left ventricle from such a large data set is both tedious and time-consuming. A robust automated segmentation algorithm of the left ventricle would be preferred.

Conventional automated segmentation techniques usually encounter difficulties in segmenting the left ventricle because 1) the intensity contrast between the ventricle and the myocardium is low (due to the smearing of the blood pool in the ventricle into the myocardium), and 2) the boundaries of the left ventricle are missing at certain locations due to the presence of protruding papillary muscles which have the same intensity profile as the myocardium.

In the experiment to illustrate our technique, we equally divided the 100 2-D images from a single patient's cardiac MRI into two sets: a training set and a test set. Fifty 4-D interactive segmentations of the left ventricle from the training set form the 2-D shape database shown in Fig. 16. This particular database is employed to allow our model to capture both the spatial and the temporal variations of the left ventricle. Fig. 17 shows the aligned version of this database. Fig. 18 compares the overlap images of the left ventricle database before and after alignment. Using the aligned database, we derived the mean level set and the eigenshapes to form the implicit shape model of the left ventricle using $k = 25$. Fig. 19 shows the mean shape of the left ventricle as well as its shape variations by varying the first three eigenshapes by $\pm 1\sigma$. The parameters of this implicit parametric representation are calculated to minimize E_{cv} using statistics calculated in the entire region both inside and outside the curve. Fig. 20 shows the segmentation result of the testing set by our algorithm (red curves). These results are comparable with the ones given by a 4-D interactive cardiac MRI segmenter [33] (green curves) which utilizes a 4-D conformal surface shrinking technqiue based upon the models outlined in [32].

B. A 3-D Example: Prostate Segmentation of Pelvic MRI Taken With Endorectal Coil

Pelvic MRI, when taken in conjunction with an endorectal coil (ERC) (a receive-only surface coil placed within the rectum) using T1 and T2 weighting, provides high-resolution images of the prostate with smaller field of view and thinner slice thickness than previously attainable. Because of the high-quality anatomical images obtainable by this technique, it may become the imaging modality of choice in the future for detection and staging of prostate cancer [7], [31]. For assignment of appropriate radiation therapy after cancer detection, the segmentation of the prostate gland from these pelvic MRI images is required. Manual outlining of sequential cross-sectional slices of the prostate images is currently used to identify the prostate gland and its substructures, but this process is difficult, time-consuming, and tedious. The idea of being able to automatically segment the prostate is very attractive.

Automatic segmentation of the prostate is difficult because the prostate is a small glandular structure buried deep within the pelvic region and surrounded by a variety of different tissues which show up as varying intensity levels on the MRI. This segmentation problem is further complicated by an artifact called the near-field effect which is caused by the use of the ERC. The near-field effect causes an intensity artifact to appear in the tissues surrounding the ERC. This can be seen as a white circular halo surrounding the rectum in each image slice of Figs. 27 and 30. The intensity artifact can bleach out the borders of the prostate near the rectum, making the prostate segmentation problem even more difficult.

We employ a 3-D version of our shape-based curve evolution technique to segment the prostate gland. By utilizing a surface (instead of a curve), the segmentation algorithm is able to utilize the full 3-D spatial information to extract the boundaries of the prostate gland. Fig. 21 shows the prostate training data we use which consists of eight 3-D binary shape models of the prostate gland-obtained by stacking together 2-D expert hand segmentations of eight patients' pelvic MRIs taken with an ERC. The alignment results of these 3-D models are shown in Fig. 21. To evaluate the alignment process, Fig. 23 shows 12 consecutive axial slice overlap images of the eight 3-D prostate gland models prior to alignment. And Fig. 24 shows the same 12 overlap images after alignment for comparison. Prior to shape training, these 3-D shape models are smoothed to remove the "step-like" artifact along the axial direction of the prostate. Based on these 3-D models, we derived the mean level set and the eigenshapes to form the implicit shape model of the prostate gland using $k = 5$. Fig. 25 shows the mean shape of the prostate gland as well as its shape variations based on varying the first three eigenshapes by $\pm 1\sigma$.

In this particular application, it is important to realize that despite the fact that the prostate gland is mostly deformed by its neighboring structures, the prostate shape parameters are still very important in describing its shape. In our method, by capturing how its surrounding structures deform the prostate gland, we obtain shape parameters that can effectively describe the deformations of the prostate gland. Specifically, we looked at a population of patients and learned the total net resultant

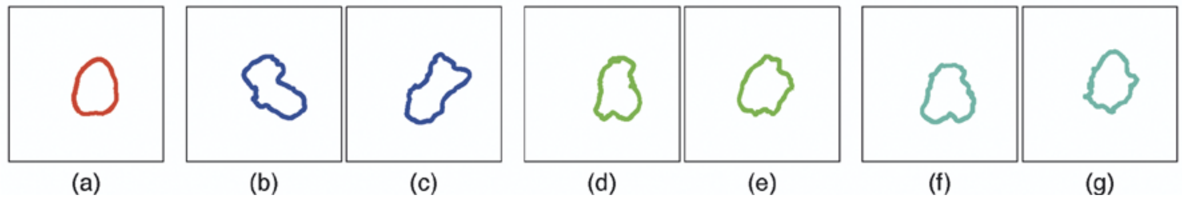


Fig. 19. Shape variability of the left ventricle. (a) The mean shape. (b) $+1\sigma$ variation of the first principal mode. (c) -1σ variation of the first principal mode. (d) $+1\sigma$ variation of the second principal mode. (e) -1σ variation of the second principal mode. (f) $+1\sigma$ variation of the third principal mode. (g) -1σ variation of the third principal mode.

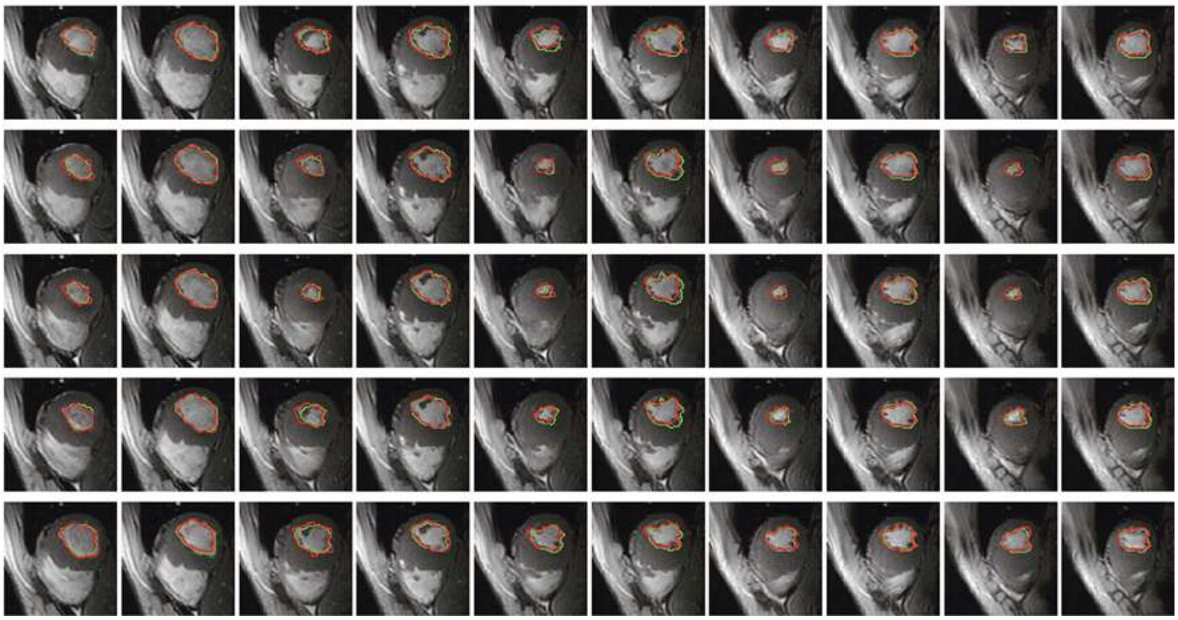


Fig. 20. Left ventricle segmentation of cardiac MRI. The segmentation by our algorithm (red curves) is compared to the segmentation by an interactive 4-D cardiac MRI segmenter (green curves).

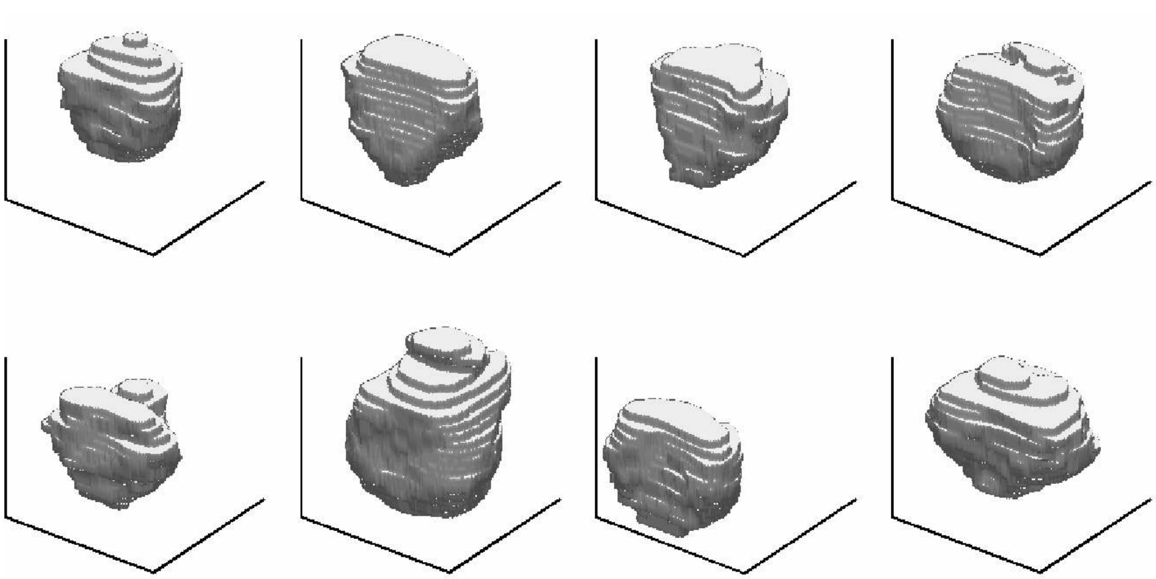


Fig. 21. Training data: eight 3-D shape models of the prostate gland obtained based on axially stacking together 2-D expert hand segmentations of the prostate.

effect of the surrounding structures in deforming the prostate gland, and incorporated this information within the prostate shape parameters. Thus, instead of looking at how the prostate gland deforms in a vacuum by itself, we have taken into account how the prostate deforms *in vivo* by the surrounding structures.

To accentuate the boundaries of the prostate gland as well as to minimize the intensity artifact caused by the ERC, the pelvic

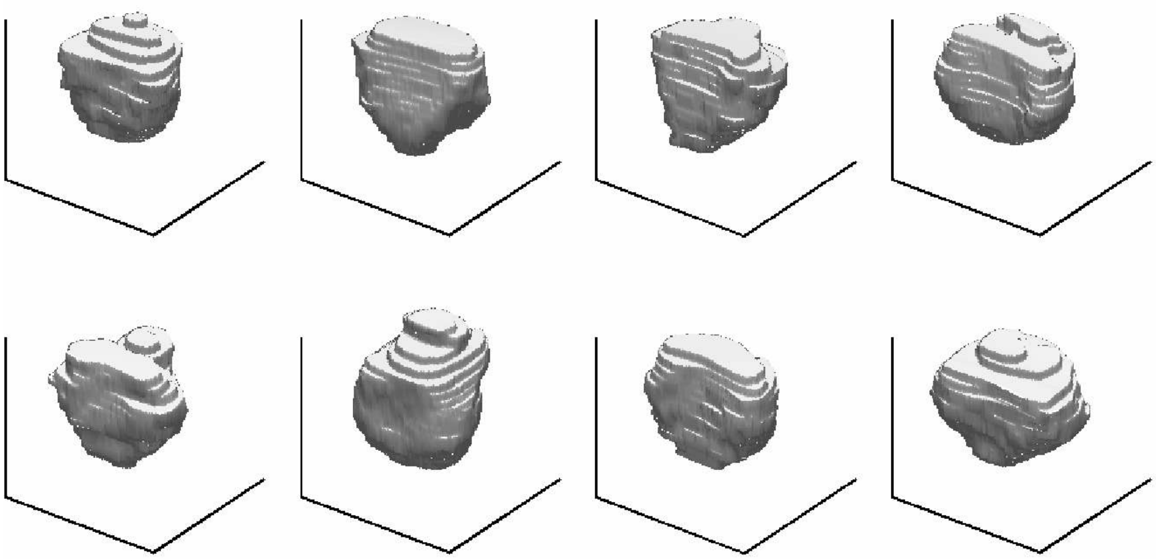


Fig. 22. Alignment results of the eight 3-D shape models of the prostate gland.

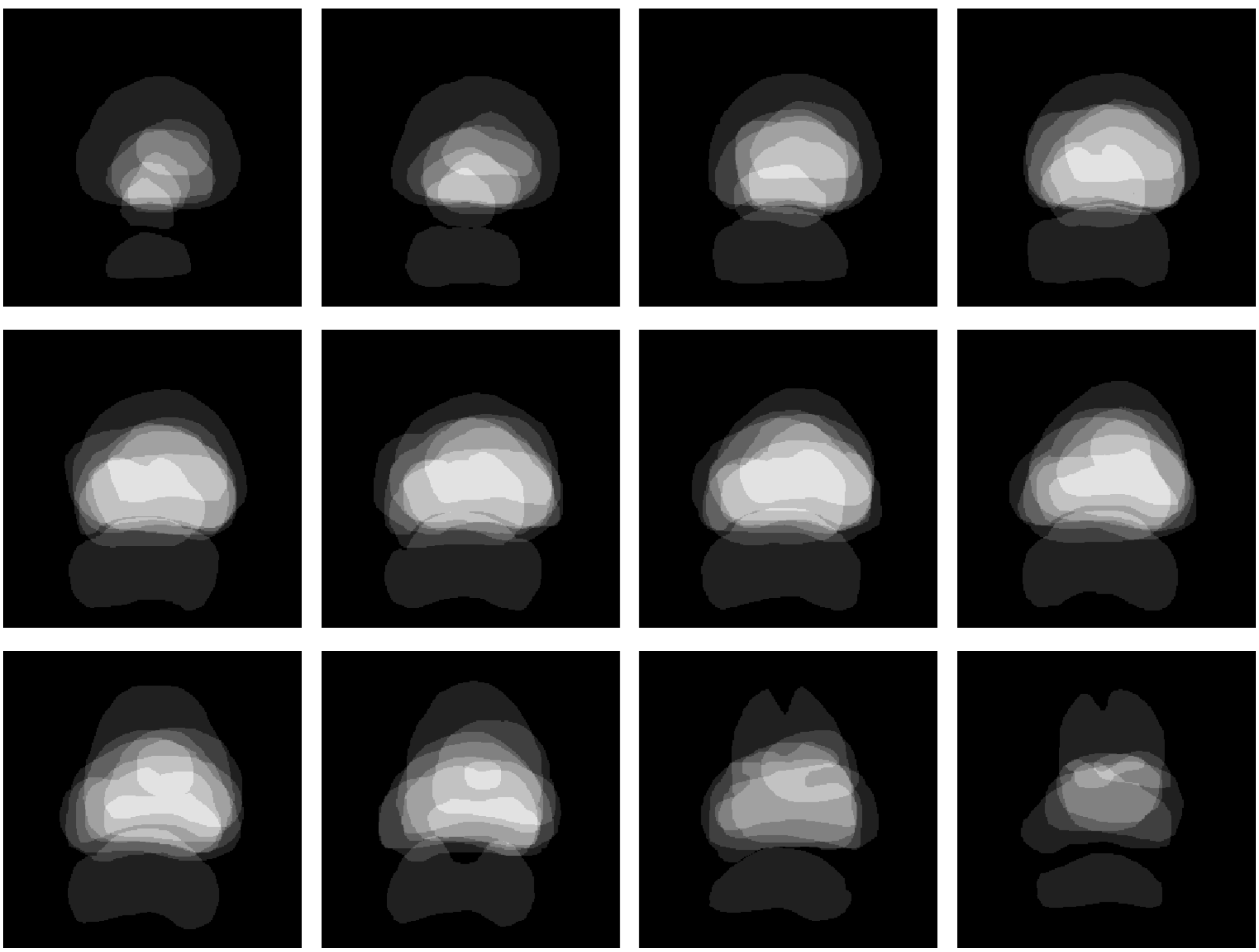


Fig. 23. Overlap images of consecutive axial slices of the eight 3-D prostate models prior to alignment.

MRI data set I_{MRI} is transformed to a bimodal data set I by applying the following map:

$$I = \|\nabla I_{\mathrm{MRI}}\|^2$$

where ∇ here denotes a 3-D gradient operator. This mapping was employed because: 1) the interior of the prostate is homogeneous in intensity, so with this mapping, the interior regions of the prostate are mapped to low values while the boundaries

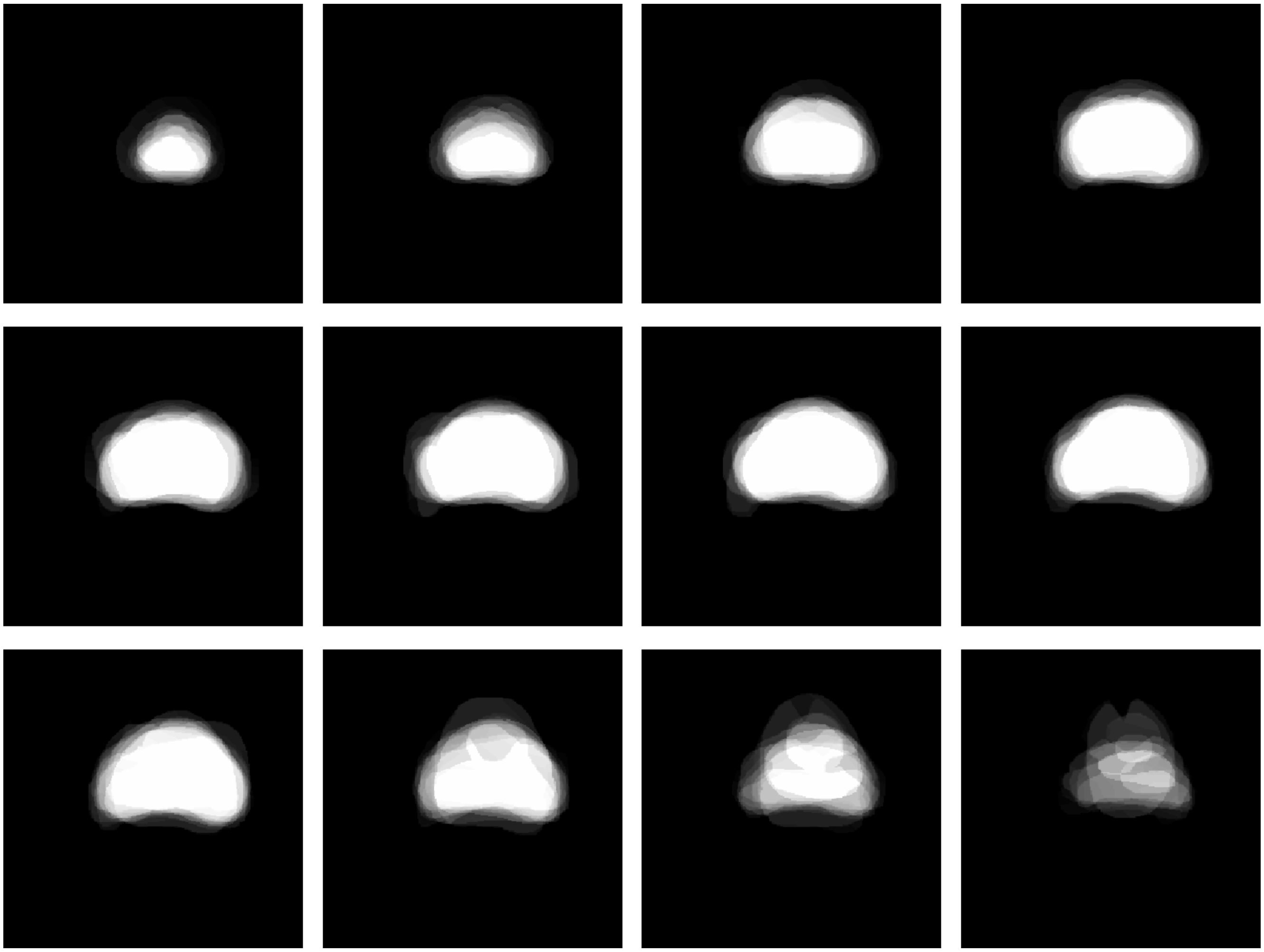


Fig. 24. Overlap images of consecutive axial slices of the eight 3-D prostate models after alignment.

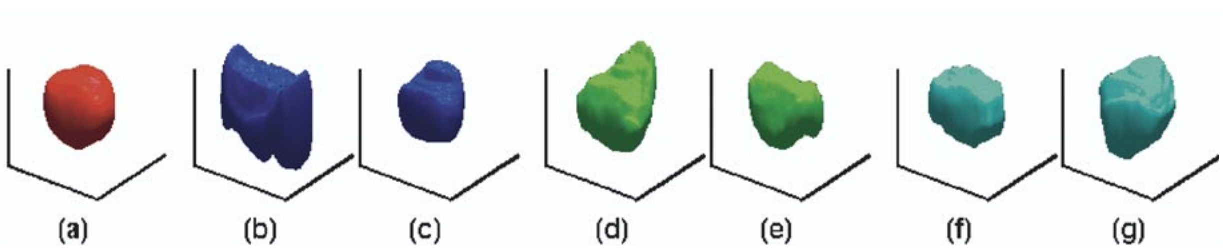


Fig. 25. Shape variability of the prostate. (a) The mean shape. (b) $+1\sigma$ variation of the first principal mode. (c) -1σ variation of the first principal mode. (d) $+1\sigma$ variation of the second principal mode. (e) -1σ variation of the second principal mode. (f) $+1\sigma$ variation of the third principal mode. (g) -1σ variation of the third principal mode.

of the prostate are mapped to high values; and 2) this mapping is robust to the smooth spatially varying intensity artifact cause by the ERC. We segment the prostate gland by minimizing $E_{\rm cv}$ using the transformed data set I. The statistics used in $E_{\rm cv}$ are calculated in the entire volumetric data both inside and outside the segmenting surface. The energy functional $E_{\rm cv}$ was employed in this application because we found it to be more robust empirically. We start by initializing the segmenting surface to be within the interior of the prostate gland so that the evolving surface does not get distracted by various other high gradient features surrounding the prostate (such as interfaces between various hard and soft tissue types). With each iteration, the segmenting surface moves outward to capture more and more of the low-valued region in the transformed data (which corresponds to the prostate gland). Eventually, the segmenting surface converges to a local minimum near the boundaries of the prostate (corresponding to high values in the transformed data).

Twelve contiguous axial slices of patient A's and B's MRI data set containing the prostate gland are displayed in Figs. 26 and 29, respectively. These two data sets are not part of the training database of Fig. 21. We show in Figs. 27 and 30 the prostate segmentation results of patient A's and B's MRI data set, respectively. In each of these figures, the MRI data set containing the prostate gland are displayed along with the segmentation by our algorithm (outlined in red), and the segmentation by a radiologist from Brigham and Women's Hospital (outlined

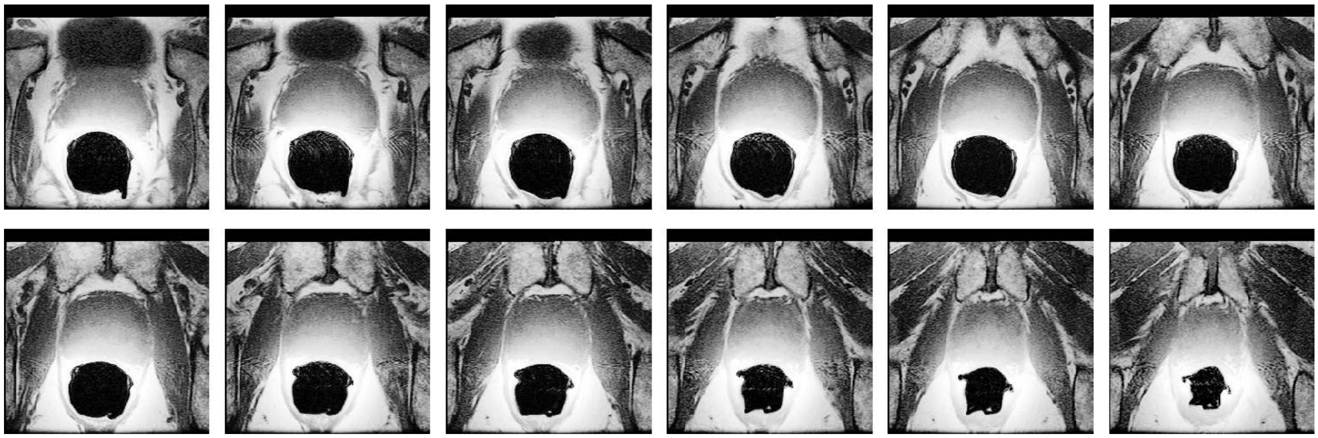


Fig. 26. Prostate images of patient A. These images represent consecutive axial slices of the prostate. Segmenting curves were not superimposed on the images for better visualization of the prostate organ.

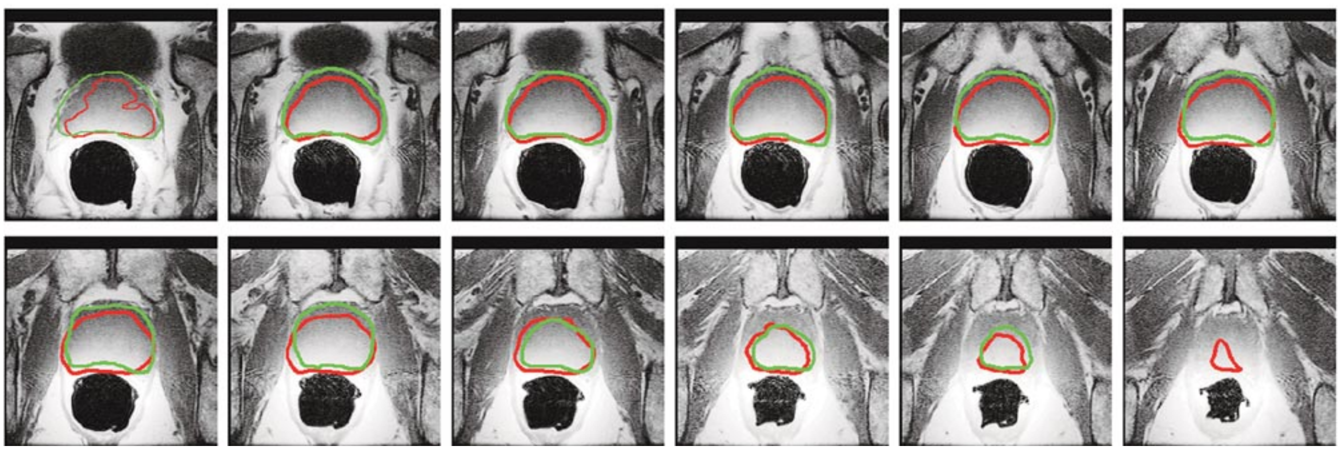
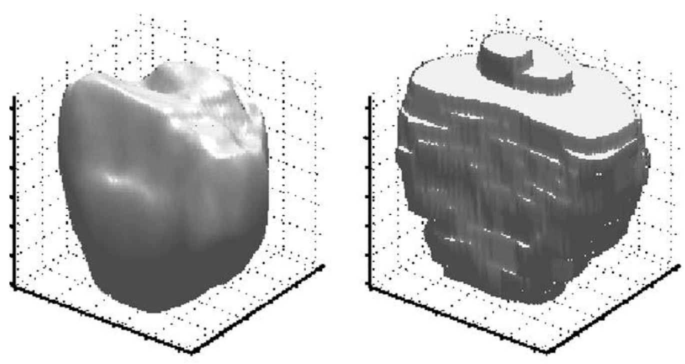


Fig. 27. Prostate segmentation of patient A. The segmentation by the radiologist (green curves) is compared to the segmentation by our algorithm (red curves).

in green). Another radiologist, also from Brigham and Women's Hospital, rated the first radiologist's segmentation of data set A to be slightly better than our algorithm's, and rated our algorithm's segmentation of data set B to be slightly better than the radiologist's. For visual comparison, Figs. 28 and 31 show the 3-D models of the prostate gland generated by our algorithm and by stacking together 2-D expert hand segmentations. Notice that by employing a surface to capture the prostate gland, our 3-D model does not display any of the "step-like" artifacts that mar the radiologist's 3-D rendition of the prostate gland. In addition, working in 3-D space allows our algorithm to utilize the full 3-D structural information of the prostate for segmentation (instead of just the information from neighboring slices which are typically used by the radiologists).



(a) (b)

Fig. 28. Three-dimensional models of patient A's prostate gland. (a) Based on our segmentation algorithm. (b) Based on the radiologist's segmentation.

VII. Conclusion and Future Research Directions

We have outlined a statistically robust and computationally efficient model-based segmentation algorithm using an implicit representation of the segmenting curve. Because this implicit representation is set in an Eulerian framework, it does not require point correspondences during the training phase of the algorithm and can be used to handle topological changes of the segmenting curve in a seamless fashion. This algorithmic framework is capable of segmenting images contaminated by heavy noise and delineate structures complicated by missing or diffuse edges. In addition, this framework is flexible, both in terms of its ability to model and segment complicated shapes (as long as the shape variations are consistent with the training data), as well as its ability to accommodate the segmentation of multidimensional data sets. Furthermore, by employing a region-based segmentation functional, our algorithm is more global, exhibits

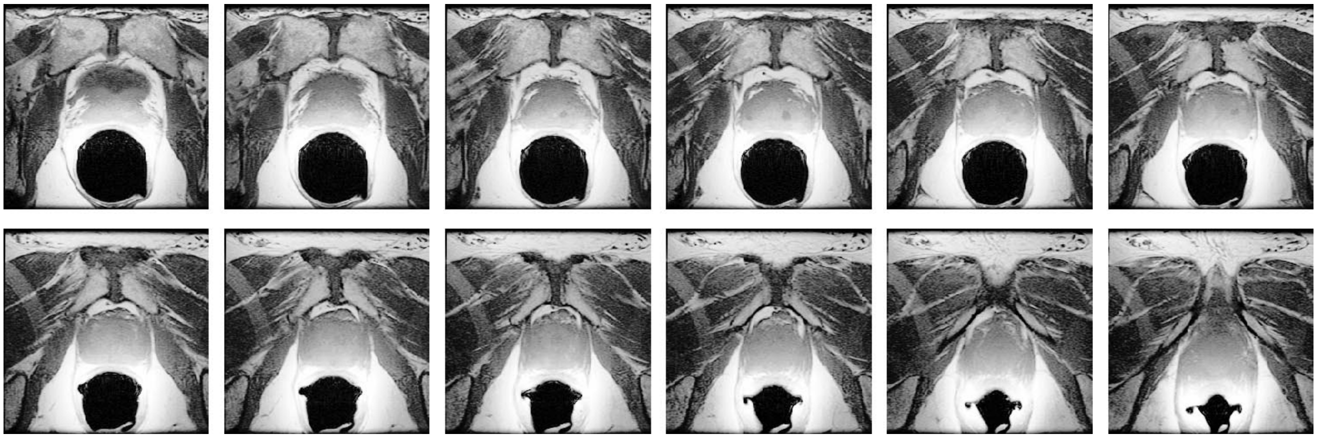


Fig. 29. Prostate images of patient B. These images represent consecutive axial slices of the prostate. Segmenting curves were not superimposed on the images for better visualization of the prostate organ.

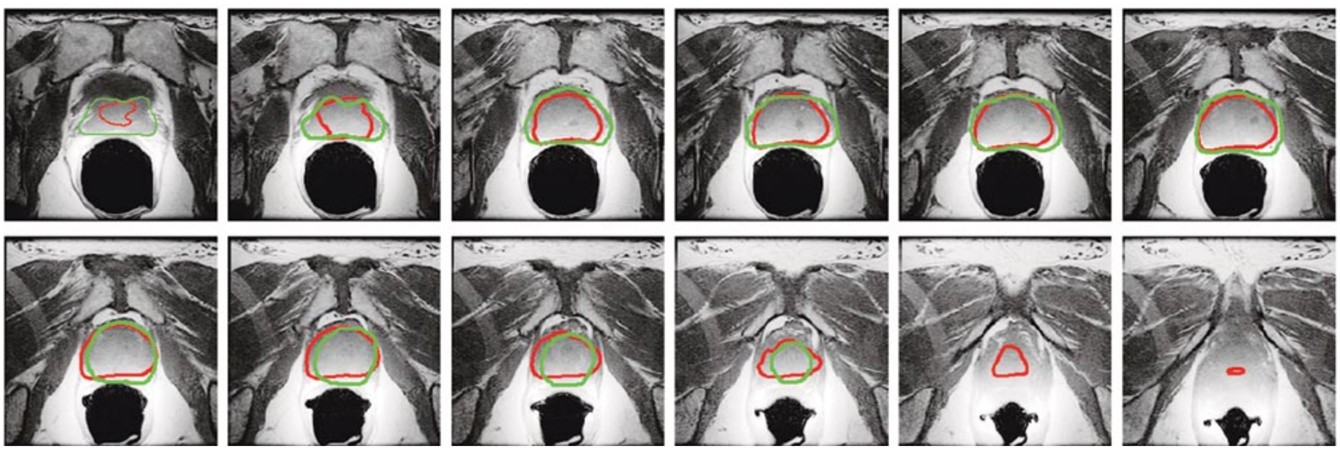
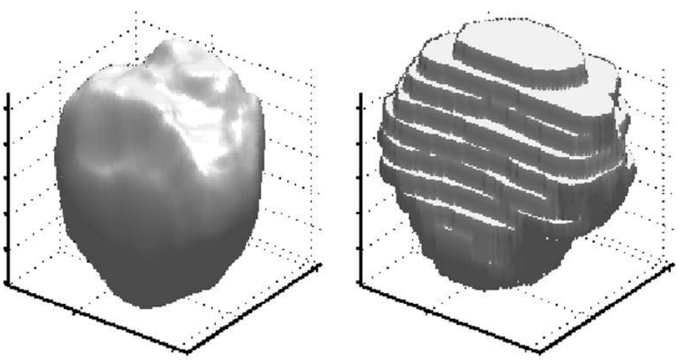


Fig. 30. Prostate segmentation of patient B. The segmentation by the radiologist (green curves) is compared to the segmentation by our algorithm (red curves).



(a) (b)

Fig. 31. Three–dimensional models of patient B's prostate gland. (a) Based on our segmentation algorithm. (b) Based on the radiologist's segmentation.

increased robustness to noise, displays extensive capture range, and is less sensitive to initial contour placements compared with other model-based segmentation algorithms.

The performance of our model-based curve evolution technique depends largely upon how well the chosen set of statistics is able to distinguish the various regions within a given image. In this paper, we detailed the use of means and variances as the discriminating statistics. However, this approach may be applied to any computed statistics. We are interested in extending our method by constructing different segmentation functionals based on first (and maybe higher) order statistics such as skewness, kurtosis, and entropy.

In this paper, we discussed the use of signed distance functions as a way to represent shapes. However, because distance functions are not closed under linear operations, the level set representation of our segmenting curve, based on the PCA approach described in Section III, is not a distance function. This gives rise to an inconsistent framework for shape modeling. This intellectual issue remains an important and challenging problem (indeed one on which we are now working ourselves), but the method developed in this paper stands on its performance in practice.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and thoughtful suggestions.

REFERENCES

[1] G. Borgefors, "Distance transformations in digital images," *CVGIP: Image Understanding*, vol. 34, pp. 344–371, 1986.

[2] V. Caselles, F. Catte, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," *Numerische Mathematik*, vol. 66, pp. 1–31, 1993.

[3] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic snakes," *Int. J. Comput. Vis.*, 1998.

[4] A. Chakraborty, L. Staib, and J. Duncan, "An integrated approach to boundary finding in medical images," in *Proc. IEEE Workshop Biomedical Image Analysis*, 1994, pp. 13–22.

[5] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Processing*, vol. 10, pp. 266–277, Feb. 2001.

[6] Y. Chen, S. Thiruenkadam, H. Tagare, F. Huang, D. Wilson, and E. Geiser, "On the incorporation of shape priors into geometric active contours," in *Proc. IEEE Workshop Variational and Level Set Methods*, 2001, pp. 145–152.

[7] D. Cheng and C. Tempany, "MR imaging of the prostate and bladder," *Seminars Ultrasound, CT, MRI*, vol. 19, no. 1, pp. 67–89, 1998.

[8] G. Christensen *et al.*, "Consistent linear-elastic transformation for image matching," in *Lecture Notes in Computer Science*, A. Kuba *et al.*, Eds., 1999, vol. 1613, Information Processing in Medical Imaging (Proc. 16th Int. Conf.), pp. 224–237.

[9] L. Cohen, "On active contour models and ballooms," *CVGIP: Image Understanding*, vol. 53, pp. 211–218, 1991.

[10] 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.

[11] B. Frey and N. Jojic, "Estimating mixture models of images and inferring spatial transformations using the EM algorithm," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, 1999, pp. 416–422.

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

[13] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Conformal curvature flows: From phase transitions to active vision," *Arch. Rational Mech. Anal.*, vol. 134, pp. 275–301, 1996.

[14] M. Leveton, "Statistical models in medical image analysis," Ph.D. dissertation, Massachusetts Inst. Technol, Dept. Elect. Eng., 2000.

[15] M. Leventon, E. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, 2000, pp. 316–323.

[16] R. Malladi, J. Sethian, and B. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 158–175, Feb. 1995.

[17] E. Miller, N. Matsakis, and P. Viola, "Learning from one example through shared densities on transforms," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, 2000, pp. 464–471.

[18] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.

[19] S. Osher and J. Sethian, "Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.*, vol. 79, pp. 12–49, 1988.

[20] N. Paragios and R. Deriche, "Geodesic Active Regions for Texture Segmentation," INRIA, Sophia Antipolis, France, Res. Rep. 3440, 1998.

[21] A. Pentland and S. Sclaroff, "Closed-form solutions for physically based shape modeling and recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 715–729, July 1991.

[22] R. Ronfard, "Region-based strategies for active contour models," *Int. J. Comput. Vis.*, vol. 13, pp. 229–251, 1994.

[23] L. Staib and J. Duncan, "Boundary finding with parametrically deformable contour models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 1061–1075, Nov. 1992.

[24] H. Tek and B. Kimia, "Image segmentation by reaction diffusion bubbles," in *Proc. Int. Conf. Computer Vision*, 1995, pp. 156–162.

[25] D. Terzopoulos and A. Witkin, "Constraints on deformable models: Recovering shape and nonrigid motion," *Artif. Intell.*, vol. 36, pp. 91–123, 1988.

[26] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, E. Grimson, and A. Willsky, "Model-based curve evolution technique for image segmentation," in *IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, 2001, pp. 463–468.

[27] A. Tsai, A. Yezzi, and A. Willsky, "A curve evolution approach to smoothing and segmentation using the mumford-shah functional," in *IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, 2000, pp. 1119–1124.

[28] T. Vetter, M. Jones, and T. Poggio, "A bootstrapping algorithm for learning linear models of object classes," in *IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, 1997, pp. 40–46.

[29] P. Viola and W. Wells, "Mutual information: An approach for the registration of object models and images," *Int. J. Comput. Vis.*, 1997.

[30] Y. Wang and L. Staib, "Boundary finding with correspondence using statistical shape models," in *IEEE Conf. Computer Vision and Pattern Recognition*, 1998, pp. 338–345.

[31] T. Wong, G. Silverman, J. Fielding, C. Tempany, K. Hynynen, and F. Jolesz, "Open-configuration MR imaging, intervention, and surgery of the urinary tract," *Urologic Clin. No. Amer.*, vol. 25, pp. 113–122, 1998.

[32] A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum, "A geometric snake model for segmentation of medical imagery," *IEEE Trans. Med. Imag.*, vol. 16, pp. 199–209, Apr. 1997.

[33] A. Yezzi and A. Tannenbaum, "4D active surfaces for cardiac segmentation," Med. Image Computing Comput. Assist. Intervention, pp. 667–673, 2002, submitted for publication.

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