

Model-Based Segmentation of Medical Imagery by Matching Distributions

Daniel Freedman, *Member, IEEE*, Richard J. Radke*, *Member, IEEE*, Tao Zhang, *Student Member, IEEE*, Yongwon Jeong, *Student Member, IEEE*, D. Michael Lovelock, and George T. Y. Chen

Abstract—The segmentation of deformable objects from three-dimensional (3-D) images is an important and challenging problem, especially in the context of medical imagery. We present a new segmentation algorithm based on matching probability distributions of photometric variables that incorporates learned shape and appearance models for the objects of interest. The main innovation over similar approaches is that there is no need to compute a pixelwise correspondence between the model and the image. This allows for a fast, principled algorithm. We present promising results on difficult imagery for 3-D computed tomography images of the male pelvis for the purpose of image-guided radiotherapy of the prostate.

Index Terms—Deformable segmentation, image-guided therapy, medical image segmentation, prostate cancer, prostate segmentation, shape and appearance model.

I. INTRODUCTION

THE segmentation of three-dimensional (3-D) deformable objects is both an important and interesting problem in computer vision. It is important because of its natural application in the medical arena; for example, segmentation of tumors from computed tomography (CT) or magnetic resonance imagery (MRI) images can be critical in the treatment of cancer. On the other hand, it is interesting because of the algorithmic challenges inherent in extracting deformable objects from real-world 3-D images. In the context of medical imagery, the key segmentation-related challenges are the following.

- Challenge 1: The objects of interest are often diffuse and lack strong edges.
- Challenge 2: There are often many objects, both of interest and not of interest, within a small volume.

Manuscript received June 23, 2004; revised November 8, 2004. This work was supported in part by the National Science Foundation (NSF) under award IIS-0133144, and the NSF Center for Subsurface Sensing and Imaging Systems (CenSSIS) under award EEC-9986821. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was A. Frangi. *Asterisk indicates corresponding author.*

D. Freedman and T. Zhang are with the Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: freedman@cs.rpi.edu; zhangt3@cs.rpi.edu).

*R. J. Radke is with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180 USA (e-mail: rjradke@ecse.rpi.edu).

Y. Jeong is with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: jeongy@rpi.edu).

D. M. Lovelock is with the Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY 10021 USA (e-mail: lovelock@mskcc.org).

G. T. Y. Chen is with the Department of Radiation Oncology and Radiation Therapy, Massachusetts General Hospital, Boston, MA 02114 USA (e-mail: gchen@partners.org).

Digital Object Identifier 10.1109/TMI.2004.841228

- Challenge 3: Many objects have fairly similar intensity profiles. Typically, this effect cannot be removed by simple pre-processing such as histogram equalization.
- Challenge 4: Many of the objects are of roughly the same shape. For example, the prostate and bladder are both “somewhat deformed” spheres.

The algorithm presented in this paper uses learned models for both the shape and appearance of objects to achieve segmentation; learning both types of information is the only reasonable way to deal with all four challenges. Our algorithm is certainly not the first algorithm to combine shape and appearance. However, existing algorithms that use both shape and appearance models (such as [6]) require a pixelwise correspondence between the model and the image; this correspondence problem can sometimes be difficult to pose rigorously and solve efficiently. Instead, our algorithm characterizes a model object by 1) its shape and 2) a probability distribution of the intensities (or colors, textures) of the pixels within its interior. As a result, comparing a particular model object to the image is as simple as comparing two probability distributions. The algorithm allows the shape to evolve until the optimal match is found.

The remainder of the paper is organized as follows. Section II reviews the existing literature on segmentation of 3-D deformable objects. Section III is the technical heart of the paper and derives the equations that comprise the multiple-object segmentation algorithm. In Section IV, we present the results of our segmentation algorithm applied to an important real-world problem in medical imaging: rapid, automatic contouring of the prostate and other structures from volumetric CT datasets for the purpose of image-guided radiation therapy. This experiment is meant to approximate the clinical application of interest: quickly fitting a patient-specific organ model to a CT image acquired just prior to treatment. Finally, Section V concludes. A shorter version of this work appeared in [10].

II. PRIOR WORK ON SEGMENTATION OF 3-D DEFORMABLE OBJECTS

The automatic segmentation of 3-D deformable objects is an active area of research in computer vision. A useful taxonomy of existing algorithms can be based on the type of learned information they use.

1) *No Learned Information*: The main exemplar of this type of approach is the traditional active contour or “snake” [20]. More recent work (e.g., [1]) has focused on geometric curve evolution, combined with level sets [23], to allow for both topological changes to the object and greater numerical stability. Standard active contour methods that seek edges tend to have

difficulty when the objects to be segmented are blurry or not sharply delineated from the background (such as the prostate in Section IV below). Extensions such as [3] try to segment on the basis of appearance by evolving the surface based on simple criteria related to the intensities (or colors) of the pixels in its interior; these methods achieve greater accuracy and robustness at the cost of a major reduction in speed. In general, algorithms that do not use learned information are constrained in terms of what they can segment; they will have difficulties with each of the four challenges posed above.

2) *Learned Shape Models*: Some researchers have augmented a level-set active contour segmentation algorithm with a term that biases the curve evolution toward shapes that are judged to be more likely based on the training set, based on principal component analysis (PCA) [22], [28]. These methods have been extended to allow for simple (nonlearned) models of appearance [33], [34]; for example, the intensities within the segmented areas may be forced to have highly differing means or variances. Segmentation of 3-D medical images has also been accomplished by the coarse-to-fine deformation of a shape-based medial representation model, or “m-rep” [25], [31]. Algorithms that possess only a learned shape model can fall prey to challenges 3 and 4.

3) *Learned Appearance Models*: A classic example of this type of technique is brain segmentation based on classification of the intensities of individual voxels, such as [37]. Another, rather different version of this sort of segmentation involves the nonparametric warping of a target surface to a deformable atlas [19]. Contours from the atlas can then be transferred onto the target volume. In either case, when no shape information is incorporated for the domain of interest, a low cost can be assigned to segmentations that could never occur in a real situation. Hence, these methods may have trouble with challenges 2, 3, and 4. Speed can also be a drawback when iterative warpings between entire image volumes are involved. Some active contour models [24] assume that one has a probabilistic characterization of appearance that is learned beforehand.

4) *Learned Shape and Appearance Models*: There are a variety of methods that model the shape and appearance of an object using PCA. The standard-bearer for such methods is the “active shape and appearance model” of Cootes *et al.* [6], which has been successfully applied to the 3-D segmentation of medical volumes, including magnetic resonance images of the brain, heart, and articular cartilage [15], [21], [38]. The main drawback of the active shape and appearance model has already been mentioned: it requires the computation of a pixelwise correspondence between the model and the image. We will say more about this in Section III-B.

III. THE SEGMENTATION ALGORITHM

In this section, we describe the heart of the proposed algorithm: the procedure for fitting a combined shape-appearance model to an image. The optimal fitting of the model results in the segmentation of the image. A key feature of the segmentation algorithm is that it can work with any parametric shape-appearance model that is C^1 in its parameters, as opposed to methods where the fitting algorithm is customized to a particular representation of shape-appearance.

The basic idea is as follows. The shape is given by a description of the surface, or multiple surfaces in the case of multi-object segmentation. The appearance is described by a probability distribution of some photometric variable inside the object of interest, or multiple distributions in the case of multi-object segmentation. A shape-appearance pair is then given by (surface, distribution), and this pair is considered sufficient to characterize an object for the sake of segmentation. The learned model is a low-dimensional manifold in the space of such pairs. To verify how well any particular shape-appearance pair matches the image, we compute the empirical distribution of the photometric variable inside the shape within the image; this distribution is then compared to the appearance model. We therefore evolve the shape of an object (or multiple objects) until the empirical distribution(s) best matches the model distribution(s). In the remainder of this section, we flesh out these ideas.

A. Terminology

In this section, we describe some of the notation needed to define the problem rigorously. In the case of single-object segmentation, a model-instance is described by a (surface, distribution) pair. The distribution is taken over some photometric variable; in the experiments we perform, this variable is grayscale intensity, though it may also be color or texture. Given that the image is discrete-valued, we will assume a probability *mass* function over the intensity. However, all of the analysis below can easily be transferred to the case of probability density functions (which might be more relevant in the case of textures). In the case of multi-object segmentation, a model-instance will be specified by J (surface, distribution) pairs, one for each object.

We assume each surface is a topological sphere and may therefore be written $S : \mathcal{S}^2 \rightarrow \mathbb{R}^3$. For convenience, we will denote a point on the surface using a parametrization of \mathcal{S}^2 as $S(u)$; however, the particular parametrization chosen is unimportant. Let us denote the image by $I : \mathbb{R}^3 \rightarrow \{1, \dots, n\}$; the image is piecewise constant, where the “pieces” correspond to voxels (which have positive volume). We denote a probability distribution by $\mathbf{q} = (q_1, \dots, q_n)$, where $q_i = \text{prob}(I(x) = i)$; of course $q_i \geq 0$ and $\sum_i q_i = 1$. Thus, a model-instance is given by $(S(\cdot), \mathbf{q})$. The shape-appearance model is a low-dimensional manifold in the space of such model-instances; a d -dimensional model is parametrized by $\beta \in \mathbb{R}^d$, and we will write $(S(\cdot; \beta), \mathbf{q}(\beta))$ (sometimes condensing $S(\cdot; \beta)$ to $S(\beta)$). One particular form for the shape and appearance model $(S(\beta), \mathbf{q}(\beta))$ will be discussed in Section IV; in the subsequent derivation, the particular form is unimportant.

The goal is to find the particular model-instance, i.e., the particular β , for which the model best matches the image. In the following section, we describe a natural criterion for scoring such matches.

B. Segmentation Criterion

Given a surface S , let \mathbf{p}^S be the distribution (probability mass function) of intensities lying inside the surface S . This can be formally defined as follows. Let V be the volume inside of S ; that is, let $S = \partial V$. In this case

$$p_i^S = \frac{\int_{x \in V} \delta(I(x), i) dx}{\int_{x \in V} dx} \quad (1)$$

where $\delta(i, j) = 1$ if $i = j$ and 0 otherwise. The numerator in this equation is simply a frequency count of the number of voxels with intensity i ; the denominator is the volume of V , which normalizes p_i^S so that it sums to 1. We will refer to \mathbf{p}^S as the *empirical* distribution corresponding the surface S .

The goal of segmentation is to find a region in the image that is most like the model. That is, we would like to find a model shape $S(\beta)$ whose empirical distribution $\mathbf{p}^{S(\beta)}$ most closely matches its model distribution $\mathbf{q}(\beta)$. In other words, the segmentation can be posed as

$$\min_{\beta} K(\mathbf{p}^{S(\beta)}, \mathbf{q}(\beta))$$

where K is some suitable measure of dissimilarity between probability distributions. Before turning to the important issue of the functional form of K , let us discuss, in general terms, the advantages of such an algorithm.

The key feature of this segmentation algorithm, compared to other joint shape-appearance model-based algorithms, is that there is no need to find correspondence between the pixels of the model and those of the image. Algorithms to find pixelwise correspondences are problematic for two reasons: first, they are relatively slow; second, it is often not obvious how to rigorously pose the problem of finding such a correspondence, and so heuristics are needed. For example, in the canonical Active Appearance Model [5], in order to directly compare the texture (appearance) of two shapes, they must first be warped into the same “shape-free” or “mean-shaped” coordinate system using a triangulation-based interpolation method. While modern graphical processing units (GPUs) may have made this warping process faster, it is not clear that interior “correspondences” induced far from the shape boundaries are very reliable; nor does the volume-to-volume warping satisfy any rigorous optimality conditions.

In the proposed algorithm, by contrast, pixels are not compared directly; instead, distributions are compared. Such a comparison can be performed very quickly; furthermore, as we shall see below, there are natural mathematical measures for comparing distributions. Thus, we circumvent the two problems associated with methods based on pixelwise correspondences mentioned above. It must be acknowledged that some information is obviously lost in performing distribution comparisons instead of pixelwise comparisons. However, we show that in relevant experiments this loss of information does not adversely affect performance. It is also natural to consider modeling the intensity distribution of the exterior of an object as well as the intensity distribution of its interior, to guard against underfitting of homogeneous regions; in our application this seems unnecessary and may even be disadvantageous (see the discussion of Fig. 9 in Section IV-C). In any case, it would be straightforward to incorporate such a term in the derivation.

There are several obvious candidates for the metric on distributions K from information theory. For example, we might choose the Kullback–Leibler divergence

$$K(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i}.$$

Other possibilities include one minus the Bhattacharyya coefficient [4] and the chi-squared measure. However, such information theoretic distances are known to be quite problematic [27].

The reason for this is that information theory is unconcerned with the *values* of the random variables in question; only the distributions over these variables matter. To see this, imagine three images: I_1 is all white— $I_1(x) = 255$ for all x ; $I_2(x)$ is nearly white— $I_2(x) = 254$ for all x ; and I_3 is all black— $I_3(x) = 0$ for all x . Let us now form empirical distributions $\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3$ of these images by histogramming their pixel values; all three have distributions that are δ -functions, peaked at 255, 254, and 0, respectively. Thus, if K is the Kullback–Leibler divergence, then $K(\mathbf{p}^1, \mathbf{p}^2) = K(\mathbf{p}^1, \mathbf{p}^3) = \infty$; that is, \mathbf{p}^1 is as close to \mathbf{p}^2 as it is to \mathbf{p}^3 . However, in an intuitive visual sense, the two images I_1 (white) and I_2 (almost white) are clearly much closer to each other than I_1 (white) is to I_3 (black). In order to take this into account, other metrics are needed.

Rubner *et al.* developed one such metric known as the Earth Mover’s Distance [27], which takes into account the values of the random variables. While this is an excellent candidate for K in theory, it can only be computed by solving an optimization problem, and the distance is continuous, but not everywhere differentiable, in the arguments \mathbf{p} and \mathbf{q} . Since we will need to take derivatives, the Earth Mover’s Distance is unsuitable. Instead, we use another distance based on cumulative distribution functions (cdfs). In particular, let

$$P_i = \sum_{j=1}^i p_j, \quad Q_i = \sum_{j=1}^i q_j$$

be the cdfs for \mathbf{p} and \mathbf{q} . Then we may use the *cdf distance*

$$K(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^n |P_i - Q_i|^\alpha \quad (2)$$

where $\alpha > 1$. This is simply an “unnormalized” L_α distance between the two cdfs, taken as vectors. (Note: we exclude the possibility of $\alpha = 1$ since this renders the function K nondifferentiable when $P_i = Q_i$ for any i .) This distance was first proposed in [30].

To see why the cdf distance is more effective at capturing similarity than the information theoretic distances, let us return to the example of the three images. Based on the pdfs \mathbf{p}^i , we may compute the cdfs \mathbf{P}^i . A simple calculation shows that using the cdf distance, the distance between the white and almost white image is $K(\mathbf{p}^1, \mathbf{p}^2) = 1$, while the distance between the white and the black image is $K(\mathbf{p}^1, \mathbf{p}^3) = 255$ (and $K(\mathbf{p}^2, \mathbf{p}^3) = 254$). One can imagine more complex cases that illustrate the same phenomenon. Note that the cdf distance is only sensible when the variable itself is one-dimensional; however, as we will be using grayscale intensities in our application, this does not pose a problem.

When we wish to segment multiple objects at once, our model is given by J object descriptors $\{(S_j(\cdot; \beta), \mathbf{q}_j(\beta)), j = 1, \dots, J\}$, and the goal is then to solve

$$\min_{\beta} \sum_{j=1}^J K(\mathbf{p}^{S_j(\beta)}, \mathbf{q}_j(\beta)).$$

Note that there is a single parameter vector β that controls all of the objects; this captures the notion that the objects’ shapes and appearances may be interrelated. Although a more general version of this criterion might be a weighted sum of K -functions, we have found the unweighted criterion works well in practice.

A related density matching approach has been explored previously by the authors, though not in the context of medical imagery. In [11], no shape information is used, resulting in a partial differential equation that aims simply to maximize the density matching without regard to shape. Other researchers have employed a related approach [18]. In [39], a shape term is added to the density matching functional, leading to a soft shape constraint.

C. Optimization of the Criterion

We wish to minimize

$$K(\beta) \equiv K(\mathbf{p}^{S(\beta)}, \mathbf{q}(\beta))$$

in the case of single object segmentation. (We will only deal with the single object case in this section; the multi-object case follows straightforwardly.) We will solve for a local minimum of the criterion via gradient descent, i.e.,

$$\frac{d\beta}{dt} = -\frac{\partial K}{\partial \beta}. \quad (3)$$

The computation of the derivative of K with respect to the parameters β is complicated by the reliance of the surface itself on the parameters. In the following, we will assume that K is a general distribution distance; we will specialize to the case of the cdf distance at the conclusion of the derivation.

Expanding (3) gives

$$\frac{d\beta}{dt} = -\sum_{i=1}^n \left[\frac{\partial K}{\partial p_i} \frac{\partial p_i}{\partial \beta} + \frac{\partial K}{\partial q_i} \frac{\partial q_i}{\partial \beta} \right] \quad (4)$$

where we have shortened $p_i^{S(\beta)}$ to p_i and $q_i(\beta)$ to q_i . But (1) yields

$$p_i^{S(\beta)} = \frac{\int_{x \in V(\beta)} \delta(I(x), i) dx}{\int_{x \in V(\beta)} dx} \equiv \frac{N_i^{S(\beta)}}{|V(\beta)|}$$

where $|V(\beta)|$ is the actual measure of the volume $V(\beta)$; thus

$$\frac{\partial p_i}{\partial \beta} = \frac{1}{|V(\beta)|} \left(\frac{\partial N_i}{\partial \beta} - p_i \frac{\partial |V|}{\partial \beta} \right). \quad (5)$$

In order to compute $\partial N_i / \partial \beta$ and $\partial |V| / \partial \beta$, we need to be able to determine derivatives of the form $\partial \psi / \partial \beta$, where $\psi = \int_{x \in V(\beta)} \nu(x) dx$. The variational derivative of ψ with respect to the surface S (where $S = \partial V$) is given by $\delta \psi / \delta S = \nu(u) n(u)$, where $n(u)$ is the normal to the surface at the point $S(u)$ (see, for example, [2]). It can be then be shown by a sort of generalized chain rule that

$$\frac{\partial \psi}{\partial \beta} = \int_{u \in S^2} \nu(u) \frac{\partial S}{\partial \beta}(u; \beta) n(u; \beta) du$$

where $\partial S / \partial \beta$ is a $d \times 3$ matrix ($d = \dim(\beta)$). To simplify future computations, we introduce the notation

$$\Gamma(u; \beta) = \frac{1}{|V(\beta)|} \frac{\partial S}{\partial \beta}(u; \beta) n(u; \beta)$$

so that

$$\frac{\partial \psi}{\partial \beta} = \int_{u \in S^2} |V(\beta)| \Gamma(u; \beta) \nu(u) du.$$

We have therefore that

$$\begin{aligned} \frac{1}{|V(\beta)|} \frac{\partial N_i}{\partial \beta} &= \int_{u \in S^2} \Gamma(u; \beta) \delta(I(S(u; \beta)), i) du \\ \frac{1}{|V(\beta)|} \frac{\partial |V|}{\partial \beta} &= \int_{u \in S^2} \Gamma(u; \beta) du. \end{aligned}$$

Plugging these results into (5) yields

$$\frac{\partial p_i}{\partial \beta} = \int_{u \in S^2} [(\delta(I(S(u; \beta)), i) - p_i) \Gamma(u; \beta)] du.$$

Combining this in turn with (4) gives

$$\begin{aligned} \frac{d\beta}{dt} &= -\sum_{i=1}^n \left\{ \frac{\partial K}{\partial q_i} \frac{\partial q_i}{\partial \beta} \right. \\ &\quad \left. + \frac{\partial K}{\partial p_i} \left[\int_{u \in S^2} [(\delta(I(S(u; \beta)), i) - p_i) \Gamma(u; \beta)] du \right] \right\} \\ &= \int_{u \in S^2} \left[\sum_{i=1}^n p_i \frac{\partial K}{\partial p_i} - \sum_{i=1}^n \frac{\partial K}{\partial p_i} \delta(I(S(u; \beta)), i) \right] \\ &\quad \times \Gamma(u; \beta) du - \sum_{i=1}^n \frac{\partial K}{\partial q_i} \frac{\partial q_i}{\partial \beta} \end{aligned}$$

yielding finally

$$\begin{aligned} \frac{d\beta}{dt} &= \int_{u \in S^2} \left[\sum_{i=1}^n p_i \frac{\partial K}{\partial p_i} - \frac{\partial K}{\partial p_j} \Big|_{j=I(S(u; \beta))} \right] \Gamma(u; \beta) du \\ &\quad - \sum_{i=1}^n \frac{\partial K}{\partial q_i} \frac{\partial q_i}{\partial \beta}. \quad (6) \end{aligned}$$

Equation (6) contains the situation for the case of a general distance function K between distributions; we now specialize to the case where K is the cdf distance given in (2). In this case, it can be shown that

$$\frac{\partial K}{\partial p_i} = \alpha \sum_{j=i}^n \text{sign}(P_j - Q_j) |P_j - Q_j|^{\alpha-1}$$

and $\partial K / \partial q_i = -\partial K / \partial p_i$. This can be plugged directly into (6) to get the final parameter flow, which in turn will yield the minimum of the distance K .

In the next section, we discuss some implementation issues that arise in the computation of $d\beta/dt$.

D. Implementation Issues

In general, we cannot compute the integral in (6) analytically; we must resort to a finite element method. This is relatively straightforward, given that the surface representation we use in our application is a mesh (simplicial complex). For any triangle of the mesh, the normal is fixed; furthermore, the triangles are chosen to be small enough so that neither $\partial S / \partial \beta$ nor I varies

much over the triangle. As a result, we can approximate the integral in (6) by

$$\begin{aligned} \int_{u \in S^2} \left[\sum_{i=1}^n p_i \frac{\partial K}{\partial p_i} - \frac{\partial K}{\partial p_j} \Big|_{j=I(S(u;\beta))} \right] \Gamma(u; \beta) du \\ \approx \sum_{t \in T} \left[\sum_{i=1}^n p_i \frac{\partial K}{\partial p_i} - \frac{\partial K}{\partial p_j} \Big|_{j=I(S(u_t;\beta))} \right] \Gamma(u_t; \beta) a_t \end{aligned}$$

where T is the set of triangles in the mesh, u_t is a representative point on the triangle t (typically the centroid), and a_t is the area of the triangle t . While $\partial S / \partial \beta$ is only explicitly given at the vertices of the mesh (this is the way it is learned), one can easily interpolate $\partial S / \partial \beta$ to any other point using the piecewise linear structure of the mesh.

In principle, the empirical distribution p_i could be computed exactly, using the fact that the surface S encloses a polyhedral volume, and the image is piecewise constant. However, this computation would be rather expensive. Instead, we use the following approximation that has proven to be very accurate. Each slice of the image is intersected with the mesh representing the surface S ; this yields a series of polygons. There are standard, fast algorithms for determining which pixels in a two-dimensional (2-D) image lie within a polygon; for each slice, one such algorithm is used to compute the histograms for that slice's polygon. These histograms are then added together, and normalized. The resulting empirical distribution is extremely close to the true empirical distribution, and can be computed very quickly.

IV. APPLICATION TO IMAGE-GUIDED RADIOTHERAPY OF THE PROSTATE

Now we demonstrate the application of our segmentation algorithm to our motivating problem: the segmentation of the prostate, as well as adjacent radiation-sensitive organs (e.g., the bladder and rectum) from 3-D CT imagery, for the purpose of improving radiation therapy. Typically, a patient is treated on thirty to forty separate occasions, and the goal is to reliably irradiate the same tissue at each session by localizing the organs of interest immediately before treatment. This is a difficult problem, since the position, shape, and appearance of the bladder and rectum can change quite dramatically from day to day, and these organs in turn press on the prostate, causing it to deform. Manual segmentation of the organs is out of the question; a radiation oncologist can take 20 minutes to outline the prostate alone, while treatment sessions are typically scheduled in half-hour slots. Hence, an efficient and accurate computer vision method for automatically estimating the organ outlines in the daily CT images, using a patient-specific model, would be of enormous benefit in the process. While methods based on semi-automatic translation [17] or rigid motion [36] of a fixed prostate model have been considered, these are generally not sufficient to capture the true deformations of all objects of interest. Using a nonoptimized implementation on a modest PC, our algorithm is able to perform the nonrigid multi-object contouring procedure in well under a minute.

Several vision approaches have been presented for prostate segmentation from 2-D transrectal ultrasound (TRUS) images; see [14], [29] for reviews. For example, Gong *et al.* [14]

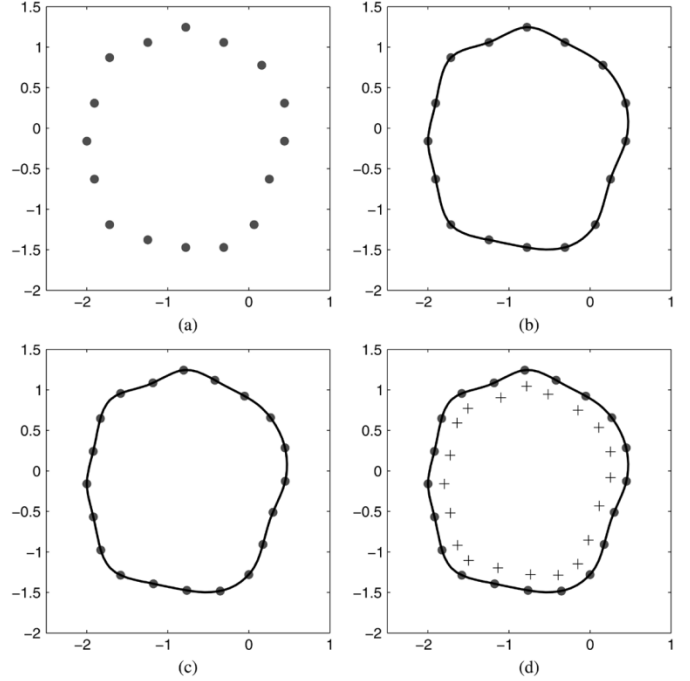


Fig. 1. The interpolation and resampling process. (a) Original points on organ boundary drawn by radiation oncologist. (b) Spline fit interpolated through original points. (c) Evenly spaced points resampled around spline contour. (d) Evenly spaced points, with interior constraints indicated by '+'s.

described a method based on 2-D deformable superellipses, while Shen, Zhan, and Davatzikos [29] proposed an active-contour-based method involving Gabor filter banks. Snake-based algorithms with manual initialization for 3-D ultrasound images were described by Hu *et al.* [16] and Ghanei *et al.* [13]. The work most comparable to the algorithm described here was recently proposed by Tsai *et al.* [33], [34] and applied to 3-D MRI images of the prostate, using PCA on level sets combined with a region-based segmentation algorithm.

A. Learning the Shape Model

In order to implement (6), we must have a shape-appearance model, $(S(\beta), \mathbf{q}(\beta))$; this model is to be learned from training data. Here we describe one algorithm for learning a statistical shape model from this data. We emphasize that if necessary, this simple model could be replaced with a more sophisticated shape model like those discussed in Section II, but the segmentation equations would remain the same. The treatment of the appearance model is left for the next section.

The training images consist of 3-D clinical data sets in which the organs of interest in each 2-D axial image have been contoured. In particular, a radiation oncologist typically contours organs in a planning scan by marking several boundary points in each of 20–40 slices of a CT volume. There are generally different numbers of points per slice, and the points are not usually equally spaced around the object contours [see Fig. 1(a)]. Furthermore, the number of slices per organ varies within the training data set. Our goal is to build a 3-D deformable shape model of the objects of interest represented by the training data.

Our approach is to construct a surface that interpolates all of the original contour points for each training data set, and then

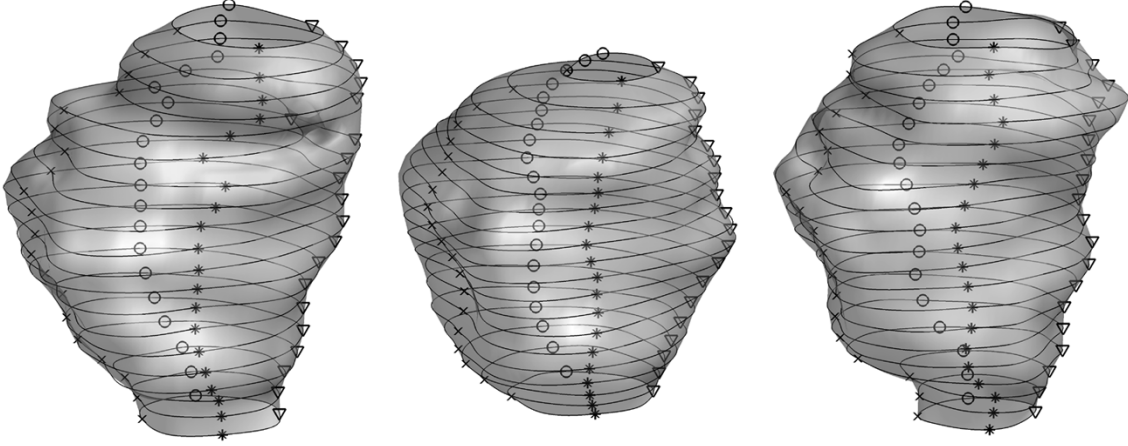


Fig. 2. Three 3-D prostates recovered by the variational implicit surface method, which interpolates the original contour points. The symbols on each contour indicate the locations of several corresponding points.

to resample this surface in the same way for each training data set in order to obtain a set of roughly corresponding points. The first step is to fit a spline through the unevenly spaced points for each contour [Fig. 1(b)]. This spline is then resampled to obtain a given number of evenly spaced points around the contour [Fig. 1(c)]. We imagine that these resampled contour points (for all slices) are on the 0-level-set of a function f defined over \mathbb{R}^3 . We would like to find the entire 0-level-set, and hence, a surface in 3-D that interpolates the original contour points. The framework of variational implicit surfaces proposed by Turk and O'Brien [35] does exactly this. Briefly, we solve an optimization problem that looks for the smoothest function whose 0-level-set exactly contains the sample points. It turns out that this optimization leads to an analytical solution for the level-set function

$$f(x) = \sum_{j=1}^n w_j \phi(x - x_j) + P(x)$$

where the x_j 's are contour points, $\phi(x) = |x|^2 \log |x|$, and $P(x)$ is a linear polynomial in x . In order to compute the weights w_j and the coefficients inherent in $P(x)$, we use the fact that we already have several samples of this function $f(x_i) = v_i$, where x_i is a contour point, and $v_i = 0$, since each contour point is on the 0-level-set. We can rewrite these samples as

$$v_i = \sum_{j=1}^n w_j \phi(x_i - x_j) + P(x_i)$$

which is a linear least-squares problem in the weights w_j and the coefficients of the polynomial $P(x)$. For stability, Turk and O'Brien suggested specifying an additional set of sample points, slightly inset from the contour sample points, where the functional value is equal to 1. These additional points are indicated by + 's in Fig. 1(d), and are obtained simply by travelling a small fixed distance inwards and normal to the spline at each contour point.

At this point, we have obtained a complete level-set function f_k for a given organ in each training data set k . Now we resample each 0-level-set surface with the same number of equally

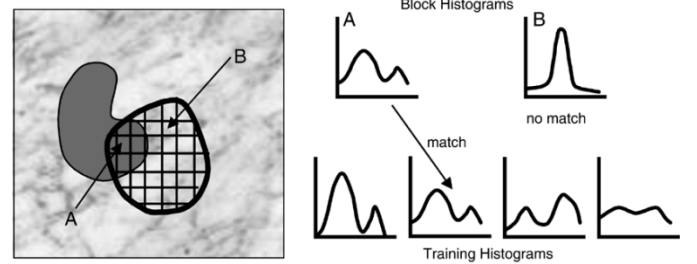


Fig. 3. The image is shown on the left; the object is textured, and the initial position is shown in bold. Note that the initial position only overlaps the object to a small extent. On the right, the histograms of two blocks, A and B are shown; the histogram of block A matches one of the training histograms well, while the histogram of block B does not.

spaced slices and the same number of points distributed around the contour in each slice. Points in the contours from matching slices in different training sets are put into correspondence based on angle and arc-length constraints. Fig. 2 illustrates the variational implicit surface and correspondence grid for three different training datasets.

Once the contours have been resampled and put into correspondence, we train a 3-D deformable model using a linear representation derived from PCA, a common technique in the model-based segmentation literature; see for example [6]. The general scheme is as follows: each training surface is represented by the mesh obtained above, and is thus specified by a vector whose elements are the x -, y -, and z -coordinates of each of the vertices. The vectors representing the three organs are then stacked into a combined vector. There is one such combined vector for each training image, and PCA is performed on these combined vectors. This results in a mean shape as well as a rank-ordered set of orthogonal modes of variation. An instance of the deformable model is specified by the weights on these orthogonal modes. Hence, the β parameters in the segmentation algorithm correspond to the weights on the principal components. We note that there has been some work on discovering an "optimal" correspondence between 3-D objects for the purpose of PCA [7], [8]; however, this process seems to be very time-consuming. We have found the algorithm described here to work well enough in practice.

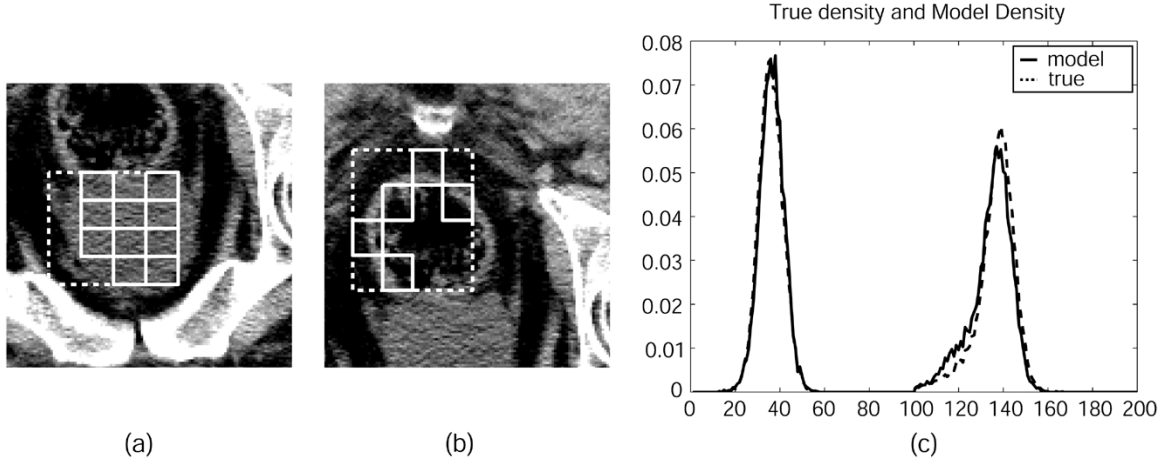


Fig. 4. (a) and (b) The blocks that contributed to the estimation of the joint model density in one slice of patient 6082 for the prostate and rectum, respectively. The search region is shown with a dotted boundary, and the selected blocks are outlined in white. (c) The true model density compared with the model density generated by the selected blocks (left, prostate; right, rectum). The match is quite good.

We may now turn to the actual shape models that are generated by running this algorithm on three patient datasets, each of which contains about $15\,512 \times 512 \times 35$ CT images of the male pelvis from the same patient taken several days apart (under prone, bladder-empty conditions). In each of the images, a radiation oncologist had outlined the prostate and outer rectal wall. We resampled each object to have 20 slices with 20 points on each contour, and built two PCA models for each scan: one model using the prostate alone, and one joint model for the prostate and outer rectal wall (since the rectum is a radiation-sensitive structure that should be detected and avoided during radiation delivery). All PCA models were leave-one-out, i.e., the shape model for a given scan was learned from all the remaining scans of the same patient. In these examples, we used 10 modes for both shape models. We note that this intra-patient problem is of clinical importance; we would like to segment each volume based not on a model of how the organs of the general population vary, but of how the organs of the specific patient under treatment are expected to vary.

At this point we re-emphasize that while the specific shape model we use in our application is built based on sparse pixelwise correspondences estimated from training data, the segmentation itself does not depend on estimating pixelwise correspondence between the shape model and the image. The segmentation algorithm could easily apply to a shape model built by “correspondence-less” means. In any case, this shape model is built on geometric correspondences on surfaces, rather than photometric correspondences on volumes.

B. Computing the Appearance Model

To form an appearance model $\mathbf{q}(\beta)$ from the training data, we could perform PCA on training histograms. In fact, the vectors representing the histograms could be appended to those representing the shapes, which would yield the desired joint model. However, there are two major problems with this approach. First, PCA on histograms does not preserve the property that histograms are positive and sum to 1. Second, a linear combination of training histograms often produces new histograms very unlike any training example.

Instead, we employ a different approach, based on the idea that there will be some overlap (perhaps small) between the initial guess of the object’s position and its true position. Our goal should be to extract that section of the initial object volume that overlaps with the true object, and to form our model density solely based on this. Of course, it is not obvious how to extract this overlapping volume. The following heuristic is extremely successful in practice, and is illustrated in Fig. 3.

For a given test image, the volume corresponding to the initial object position is divided into blocks; denote the set of blocks $B = \{b_j\}$. For each block, we compute the histogram, $\mathbf{h}(b_j)$; we then determine how similar a particular block is to the model by measuring its similarity to each of the training histograms, $\{\mathbf{q}_i^{\text{train}}\}$ (there is one such training histogram for each training image we receive). In particular, we compute

$$K_j = \min_i K(\mathbf{h}(b_j), \mathbf{q}_i^{\text{train}}).$$

If such a value is low, then we know that the cdf distance between the block’s histogram and at least one of the training histograms is small; as a result, the block is likely to belong to the true object. We can then rank order the blocks by their K_j values, and choose only the fraction α of the blocks with the lowest K_j values. These “good” blocks are then deemed to be part of the true object, and the model density \mathbf{q} can be computed as the histogram of the union of these blocks. Note that α must be chosen to be less than the fraction of the initial volume that overlaps the true volume; while this fraction is not known *a priori*, $\alpha = 0.25$ produced good results in practice. Note that this strategy relies on the relative constancy, across space, of the histogram of the relevant object. This assumption does not seem to be a problem in practice, as the method generally performs very well. In fact, the model density as computed using this algorithm is often almost indistinguishable from the density corresponding to the true position of the object; an example of this phenomenon is shown in Fig. 4.

While the actual images had a dynamic range from 0 to 4095, the image contrast is very low. Therefore, the appearance model and fitting algorithm were applied to the images using values from [864, 1264] to build 100-bin histograms.

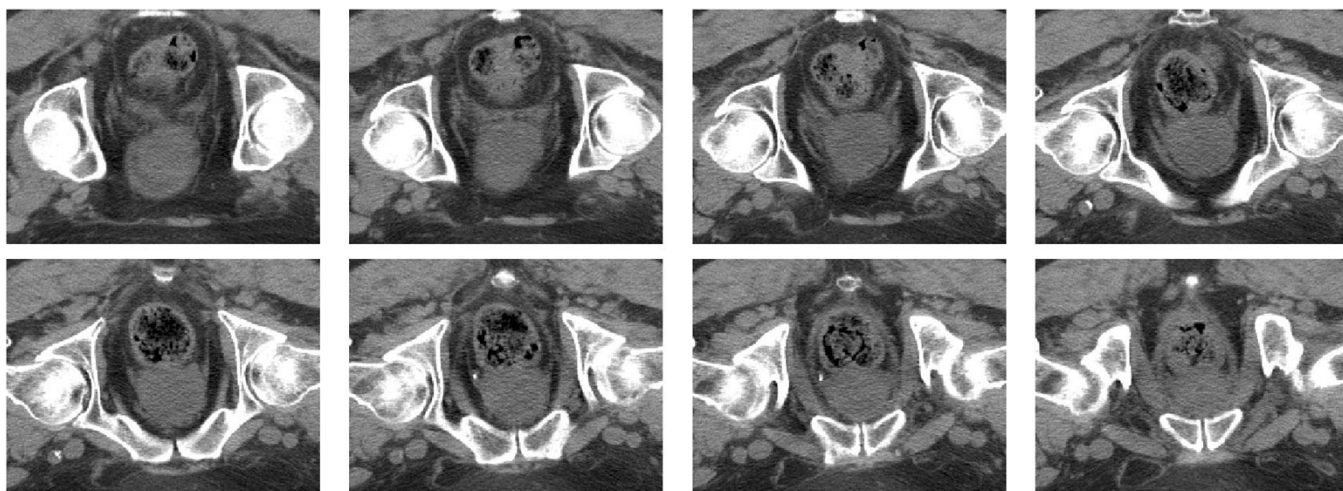


Fig. 5. Evenly spaced axial slices of scan 6042, without segmentation. Eight sequential slices of the CT volume are displayed (out of the 35 total slices input to the algorithm.)

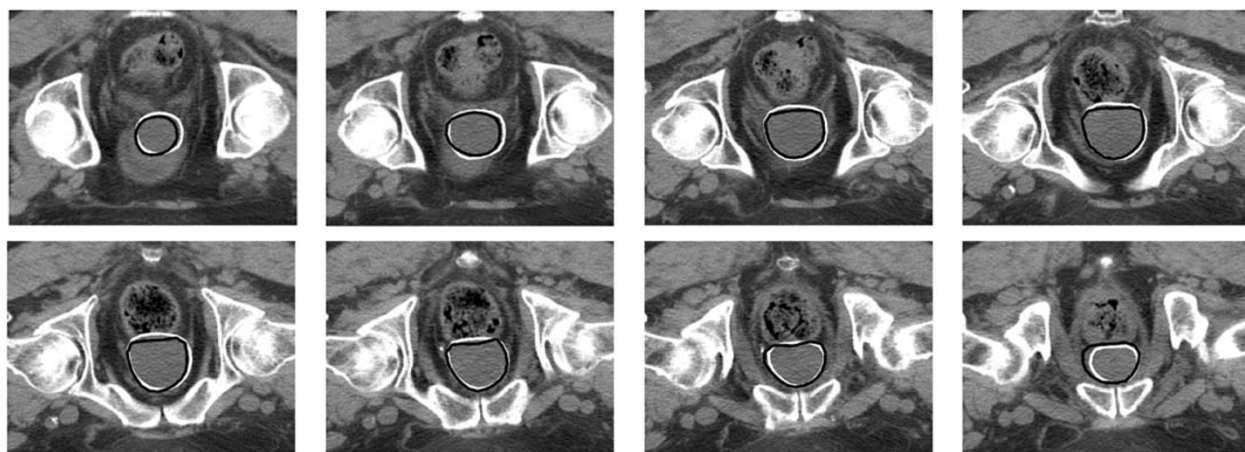


Fig. 6. Segmentation results for evenly spaced slices of scan 6042 using the prostate-only model (10 modes). The white contour shows the result at convergence. The black contour shows the hand-drawn ground-truth contours supplied by a radiation oncologist.

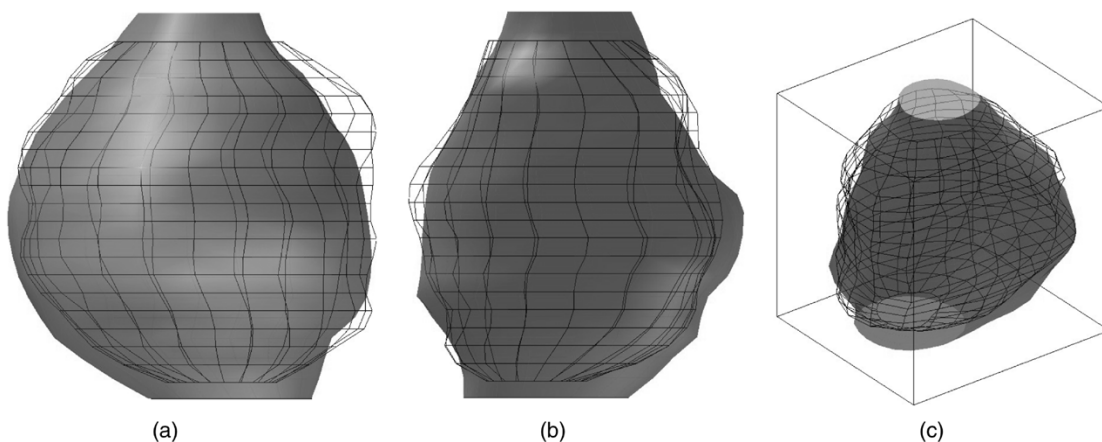


Fig. 7. (a) AP, (b) lateral, and (c) oblique 3-D visualizations of the prostate-only model (solid) versus ground truth (wireframe) for scan 6042.

C. Results

We initialized the model at the position of the mean shape of the PCA model (corresponding to $\beta = 0_{10 \times 1}$), and allowed it to converge. In our experiments, we used $\alpha = 1.2$ as the cdf exponent in (2). Illustrative results for two of the scans of patient 06000 (i.e., scans 6042 and 6082) are shown in the following

Figs. 5–14. The median results for 48 total scans from all three patients are summarized in Tables I and II.

Fig. 6 shows the segmentation results of the prostate-only model, along with ground-truth contours as drawn by a radiation oncologist, for sequential 2-D axial slices of the 3-D CT volume of patient 6042 (for visualization only; the model and segmentation algorithms are fully 3-D). The raw

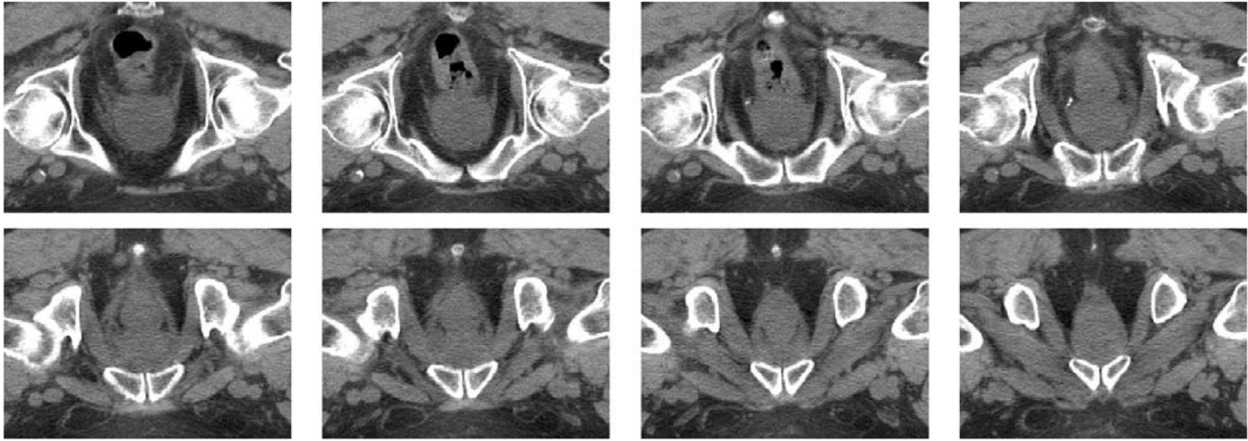


Fig. 8. Evenly spaced axial slices of scan 6082, without segmentation. Eight sequential slices of the CT volume are displayed (out of the 35 total slices input to the algorithm).

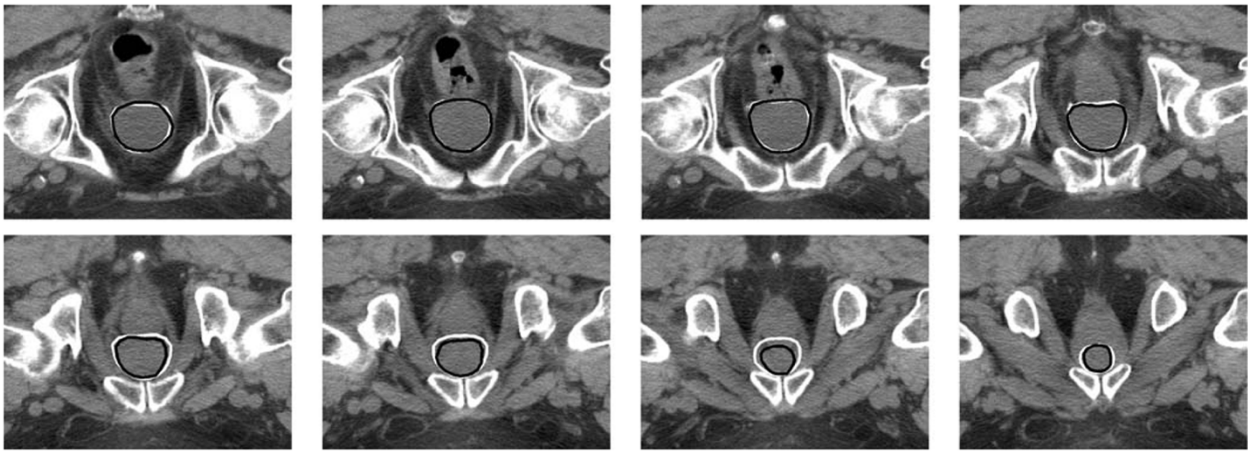


Fig. 9. Segmentation results for evenly spaced slices of scan 6082 using the prostate-only model (10 modes). The white contour shows the result at convergence. The black contour shows the hand-drawn ground-truth contours supplied by a radiation oncologist.

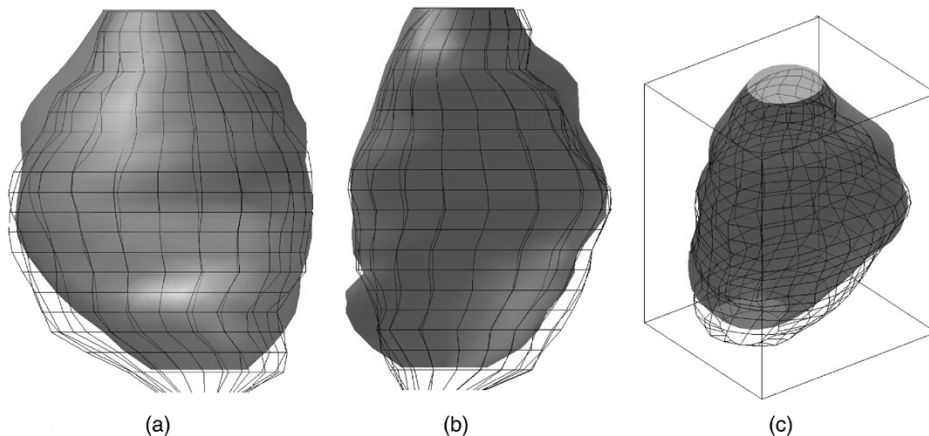


Fig. 10. (a) AP, (b) lateral, and (c) oblique 3-D visualizations of the prostate-only model (solid) versus ground truth (wireframe) for scan 6082.

input images are shown in Fig. 5 for comparison. The CT images have relatively low contrast in the area of interest; the images in these figures have been histogram-stretched for easier visibility (in medical imaging terminology, the window is 400 and the level is 40). Fig. 7 shows 3-D visualizations of the final prostate-only segmentation results for patient 6042 along with the ground truth. Figs. 8–10 show analogous results for a different scan (6082) of the same patient.

Fig. 11 shows the segmentation results of the joint prostate/rectum model, along with ground-truth contours as drawn by a radiation oncologist, for sequential 2-D axial slices of patient 6042. Fig. 12 shows 3-D visualizations of the final joint-model segmentation results for patient 6042 along with the ground truth. Figs. 13–14 show analogous results for a different scan (6082) of the same patient. There is definite intra-patient variability visible, especially in the rectum shape and appearance.

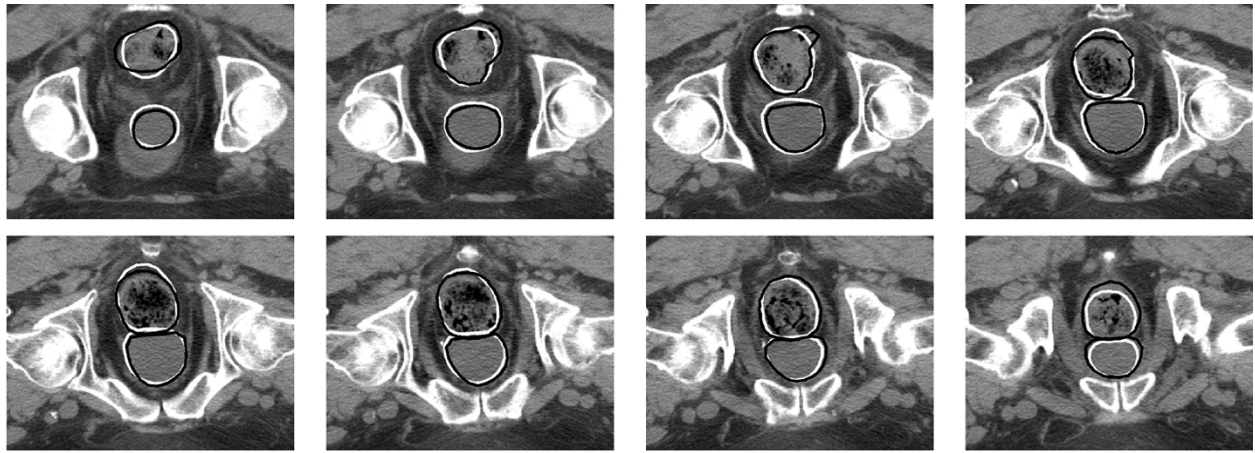


Fig. 11. Segmentation results for sequential slices of scan 6042 using the joint object model (10 modes). The white contour shows the segmentation result at convergence. The black contour shows the hand-drawn ground-truth contours supplied by a radiation oncologist. The segmented organs are, from top to bottom, the outer rectal wall and the prostate.

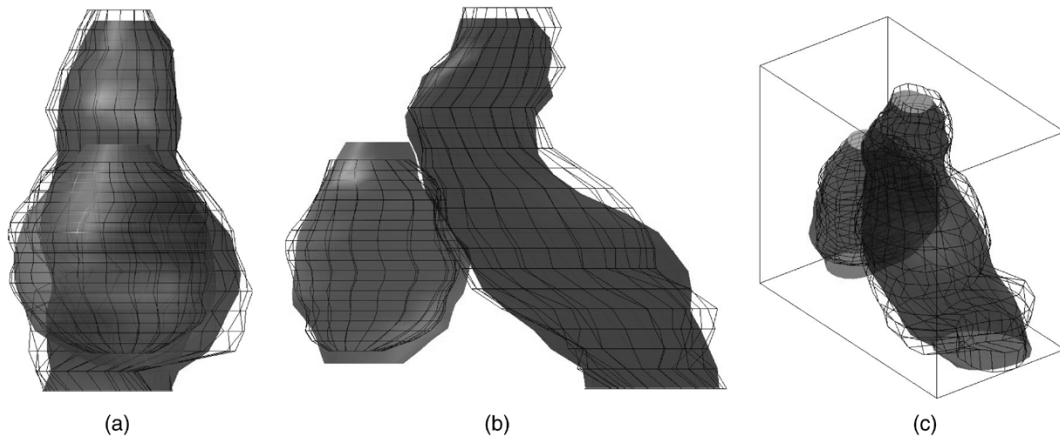


Fig. 12. (a) AP, (b) lateral, and (c) oblique 3-D visualizations of the joint model (solid) versus ground truth (wireframe) for scan 6042.

One can appreciate the difficulty of the segmentation problem in this context. The prostate organ presents no distinct “edge” in the image itself (e.g., see the last two images in Fig. 9); an active contour algorithm would naturally expand to segment the entire region that has similar intensity to the prostate, including the pelvic floor muscles and seminal vesicles. Radiation oncologists know the pelvic anatomy and take the intensity similarities into account while contouring; this expert knowledge about the underlying shape is implicitly incorporated during training of the shape model.

In addition to the qualitative accuracy of the segmentation observed above, we also report several quantitative measures taken over the entire dataset (i.e., 48 total scans from three different patients). The measures we evaluated include the following:

- v_d , the probability of detection, calculated as the fraction of the ground truth organ that was contained by the estimated organ. For a good segmentation, v_d should be close to 1.
- v_{fa} , the probability of false alarm, calculated as the fraction of the estimated organ that lies outside the ground truth organ. For a good segmentation, v_{fa} should be close to 0.
- The centroid distance, calculated as the norm of the vector connecting the centroids of the ground truth and estimated

organs. The centroid in particular is of interest to a radiation oncologist when designing a treatment plan.

- The surface distance, calculated as the median distance between the surfaces of the ground truth and estimated organs, evaluated along 1000 rays randomly directed on the surface of a sphere centered at the ground truth centroid.

The medians of these statistics over all the scans for a given patient are reported in Tables I and II. In addition, we are currently working with radiation physicists to define an error measure that is more relevant to the regions of interest for image-guided therapy (e.g., the accuracy of the point where the anterior rectal wall is tangent to the prostate). Timings on a modest computing platform (uncompiled Matlab code, 1.67 GHz AMD machine with 448 MB RAM) and iteration counts are also provided.

Our algorithm generally returns a high-quality segmentation (i.e., v_d near 0.9, v_{fa} near 0.1, prostate surface distance about 1 pixel). As can be seen from the original axial slices, this is a difficult segmentation problem. We note in passing that even determining “ground truth” in such cases is subjective; there have been several studies on expert inter-observer variability in segmentation (e.g., [9]) that report similar variability to the above (though our algorithm, or indeed any automatic vision algorithm, will certainly make some mistakes that no expert would). In the future, we would be interested to compare our algorithm’s performance to more than one expert observer’s contours.

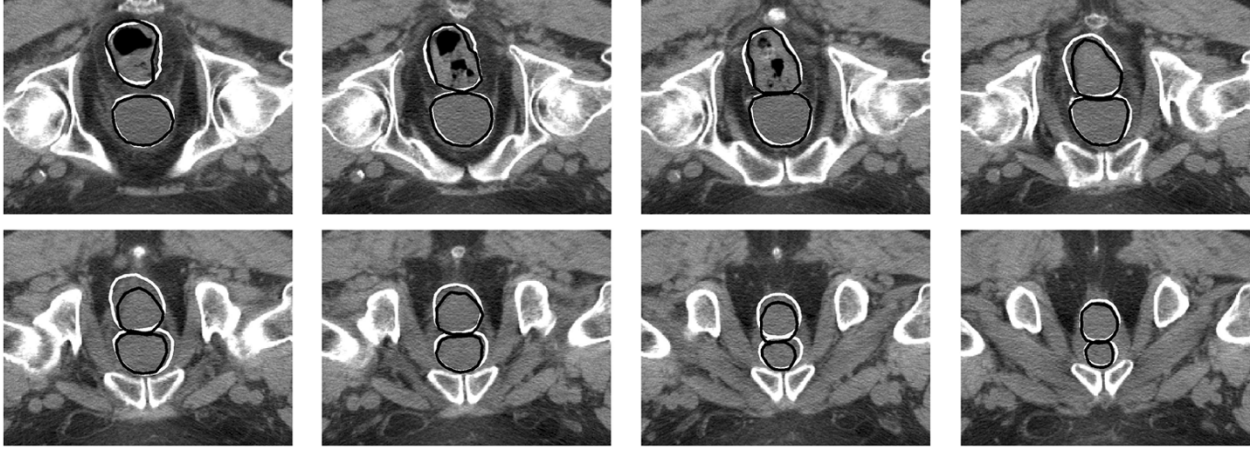


Fig. 13. Segmentation results for sequential slices of scan 6082 using the joint object model (10 modes). The white contour shows the segmentation result at convergence. The black contour shows the hand-drawn ground-truth contours supplied by a radiation oncologist. The segmented organs are, from top to bottom, the outer rectal wall and the prostate.

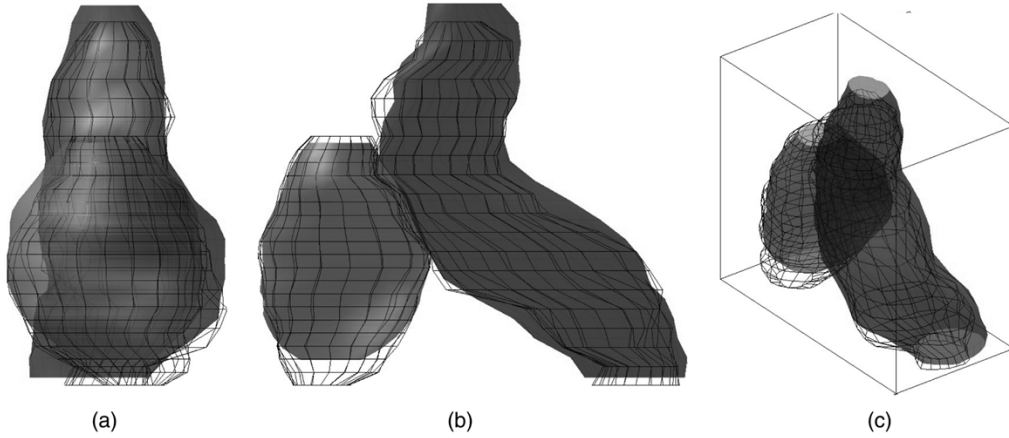


Fig. 14. (a) AP, (b) lateral, and (c) oblique 3-D visualizations of the joint model (solid) versus ground truth (wireframe) for scan 6082.

TABLE I

QUANTITATIVE RESULTS FOR THE PROSTATE, USING THE PROSTATE-ONLY MODEL APPLIED TO THREE DATASETS. THE CENTROID AND SURFACE DISTANCE MEASUREMENTS ARE IN PIXELS (1 PIXEL = 0.9375 MM)

Patient ID	06000	07000	08000
Number of scans	14	17	17
median v_d , prostate	0.894	0.811	0.857
median v_{fa} , prostate	0.102	0.182	0.118
Median of centroid distance, prostate (pixels)	3.17	5.29	4.07
Median surface distance, prostate (pixels)	0.57	1.02	0.83
Mean fitting time (sec)	14.7	10.2	17.3
Mean number of iterations	8.6	7.6	8.8

TABLE II

QUANTITATIVE RESULTS FOR THE PROSTATE AND RECTUM, USING THE JOINT MODEL APPLIED TO THREE DATASETS. THE CENTROID AND SURFACE DISTANCE MEASUREMENTS ARE IN PIXELS (1 PIXEL = 0.9375 MM)

Patient ID	06000	07000	08000
Number of scans	14	17	17
median v_d , prostate	0.829	0.864	0.803
median v_{fa} , prostate	0.135	0.123	0.186
Median of centroid distance, prostate (pixels)	4.86	3.67	5.87
Median surface distance, prostate (pixels)	0.93	0.79	1.19
median v_d , rectum	0.745	0.765	0.711
median v_{fa} , rectum	0.184	0.105	0.176
Median of centroid distance, rectum (pixels)	6.92	4.92	6.81
Median surface distance, rectum (pixels)	1.43	2.07	2.28
Mean fitting time (sec)	47.7	32.5	34.6
Mean number of iterations	12.6	7.1	6.4

V. CONCLUSION AND FUTURE WORK

We have demonstrated a segmentation algorithm that matches a learned model of shape and appearance to an image by comparing empirical and model probability distributions. The algorithm produces good results on real, challenging, medical images. In our experiments, the algorithm returns a segmentation result in a matter of seconds for the prostate-only model, and well under a minute for the joint model, compared to the 30 minutes it might take a resident to contour the same images by hand. We would expect an order of magnitude speed-up if the algorithm were implemented in C++ and optimized for efficiency.

Therefore, we conclude that our segmentation algorithm has substantial promise for the problem of rapid, automatic contouring.

Since the algorithm relies on local optimization, future work will focus on coarse-to-fine methods of sampling the model space as well as smarter initialization, to ensure that the algorithm is not trapped in an incorrect optimum. Our experiments

show that in some cases, there exist values for the model parameters that can produce better alignment with the ground truth than those to which the fitting algorithm converged. It would also be straightforward to incorporate constraints on the magnitude of the parameters to prevent the model from assuming implausible shapes (e.g., self-intersections), though this has not yet proved to be a problem in our experiments.

Another direction for future work relates to the model we chose for our application. Strict linear models can require many modes to capture complex variations. We are investigating multilinear [12] or nonlinear [26], [32] models to that will allow us to use fewer parameters, and will ensure that any model parameter values produce a "valid" instance of shape and appearance. We would also like to perform a careful experimental comparison of this method with techniques based on more complex models that use point correspondences.

We have demonstrated results on intra-patient datasets in this work; we have also investigated inter-patient prostate segmentation (see [10]). In our experiments, we found that the inter-patient shape variability is high enough that a substantial number of training sets would be required to build a model that would apply to the general population. However, we are currently analyzing a corpus of training data that includes CT volumes and contours that vary along both inter-patient and intra-patient axes. That is, for each of many patients we have many examples of that patient's bodily state. Using this data, we plan to learn a two-axis model of inter-patient and intra-patient variation that is more suitable for clinical treatment planning than the solely intra-patient models described here.

REFERENCES

- [1] V. Caselles, R. Kimmel, and G. Sapiro, "On geodesic active contours," *Int. J. Comput. Vis.*, vol. 22, no. 1, pp. 61–79, 1997.
- [2] A. Chakraborty, L. Staib, and J. Duncan, "Deformable boundary finding in medical images by integrating gradient and region information," *IEEE Trans. Med. Imag.*, vol. 15, no. 6, pp. 859–870, Dec. 1996.
- [3] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
- [4] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 5, pp. 564–577, May 2003.
- [5] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.
- [6] T. Cootes and C. Taylor, "Statistical models of appearance for medical image analysis and computer vision," *Proc. SPIE (Medical Imaging 2001: Image Processing, M. Sonka, K. M. Hanson, Eds.)*, vol. 4322, pp. 236–248, Jul. 2001.
- [7] R. Davies, C. Twining, T. Cootes, and C. Taylor, "A minimum description length approach to statistical shape modeling," *IEEE Transactions on Medical Imaging*, vol. 21, pp. 525–537, 2002.
- [8] R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor, "3D statistical shape models using direct optimization of description length," in *Proc. ECCV 2002*, vol. 3, 2002, pp. 3–20.
- [9] C. Fiorino, M. Reni, A. Bolognesi, G. Cattaneo, and R. Calandrino, "Intra- and inter-observer variability in contouring prostate and seminal vesicles: Implications for conformal treatment planning," *Radiotherapy Oncol.*, vol. 47, no. 3, pp. 285–292, Jun. 1998.
- [10] D. Freedman, R. Radke, T. Zhang, Y. Jeong, and G. Chen, "Model-based multi-object segmentation via distribution matching," presented at the IEEE Workshop on Articulated and Nonrigid Motion 2004, in Conjunction with CVPR 2004, Washington, DC, Jun. 2004.
- [11] D. Freedman and T. Zhang, "Active contours for tracking distributions," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 518–526, Apr. 2004.
- [12] W. T. Freeman and J. B. Tenenbaum, "Learning bilinear models for two-factor problems in vision," in *Proc. IEEE Computer Vision and Pattern Recognition 1997*, Jun. 1997, pp. 554–560.
- [13] A. Ghanei, H. Soltanian-Zadeh, A. Ratkiewicz, and F. Yin, "A three-dimensional deformable model for segmentation of human prostate from ultrasound images," *Med. Phys.*, vol. 28, no. 10, pp. 2147–2153, Oct. 2001.
- [14] L. Gong, S. Pathak, D. Haynor, P. Cho, and Y. Kim, "Parametric shape modeling using deformable superellipses for prostate segmentation," *IEEE Trans. Med. Imag.*, vol. 23, no. 3, pp. 340–349, Mar. 2004.
- [15] A. Hill, A. Thornham, and C. Taylor, "Model-based interpretation of 3-D medical images," in *Proc. 4th Br. Machine Vision Conf.*, Sep. 1993, pp. 339–348.
- [16] N. Hu, D. B. Downey, A. Fenster, and H. M. Ladak, "Prostate boundary segmentation from 3D ultrasound images," *Med. Phys.*, vol. 30, no. 7, pp. 1648–1659, Jul. 2003.
- [17] C. Hua, D. Lovelock, G. Mageras, M. Katz, J. Mechalakos, E. Lief, T. Hollister, W. Lutz, M. Zelefsky, and C. Ling, "Development of a semi-automatic alignment tool for accelerated localization of the prostate," *Int. J. Radiation Oncol. Biol. Phys.*, vol. 55, no. 3, pp. 811–824, 2003.
- [18] S. Jehan-Besson, M. Barlaud, G. Aubert, and O. Faugeras, "Shape gradients for histogram segmentation using active contours," in *Proc. Int. Conf. Computer Vision*, Nice, France, 2003, pp. 408–415.
- [19] S. Joshi, "Large deformation diffeomorphisms and Gaussian random fields for statistical characterization of brain submanifolds," Ph.D. thesis, Washington Univ., St. Louis, MO, 1997.
- [20] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," presented at the Int. Conf. Computer Vision, London, U.K., Jun. 1987.
- [21] A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-D neuroradiological data sets," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 828–839, Oct. 1999.
- [22] M. Leventon, E. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proc. CVPR 2000*, 2000, pp. 316–323.
- [23] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulation," *J. Comput. Phys.*, vol. 79, pp. 12–49, 1988.
- [24] N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 3, pp. 266–280, Mar. 2000.
- [25] S. Pizer, G. Gerig, S. Joshi, and S. Aylward, "Multiscale medial shape-based analysis of image objects," *Proc. IEEE*, vol. 91, no. 10, pp. 1670–1679, Oct. 2003.
- [26] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, 2000.
- [27] Y. Rubner, C. Tomasi, and L. Guibas, "The earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vis.*, vol. 40, no. 2, pp. 99–121, 2000.
- [28] D. Shen and C. Davatzikos, "An adaptive-focus deformable model using statistical and geometric information," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 8, pp. 906–913, Aug. 2000.
- [29] D. Shen, Y. Zhan, and C. Davatzikos, "Segmentation of prostate boundaries from ultrasound images using statistical shape model," *IEEE Trans. Med. Imag.*, vol. 22, no. 4, Apr. 2003.
- [30] M. Stricker and M. Orengo, "Similarity of color images," *Proc. SPIE*, vol. 2420, pp. 381–392, 1995.
- [31] M. Styner, G. Gerig, S. Pizer, and S. Joshi, "Automatic and robust computation of 3D medial models incorporating object variability," *Int. J. Comput. Vis.*, vol. 55, no. 2–3, pp. 107–122, 2002.
- [32] J. Tenenbaum, V. de Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319–2323, 2000.
- [33] A. Tsai, W. Wells, C. Tempny, E. Grimson, and A. Willsky, "Coupled multi-shape model and mutual information for medical image segmentation," in *Proc. IPMI 2003*, 2003, pp. 185–197.
- [34] A. Tsai, A. Yezzi, W. Wells, C. Tempny, D. Tucker, A. Fan, E. Grimson, and A. Willsky, "A shape based approach to curve evolution for segmentation of medical imagery," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, Feb. 2003.
- [35] G. Turk and J. F. O'Brien, "Shape transformation using variational implicit functions," in *Proc. SIGGRAPH'99 (Computer Graphics)*, 1999, pp. 335–342.
- [36] M. Van Herk, A. Bruce, A. Kroes, T. Shouman, A. Touw, and J. Lebesque, "Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration," *Int. J. Radiation Oncology Biol. Phys.*, vol. 33, no. 5, pp. 1311–1320, 1995.
- [37] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "A unifying framework for partial volume segmentation of brain MR images," *IEEE Trans. Med. Imag.*, vol. 22, no. 1, pp. 105–119, Jan. 2003.
- [38] Y. Wang and L. H. Staib, "Integrated approaches to nonrigid registration in medical images," in *Proc. IEEE WACV 1998*, Oct. 1998, pp. 102–108.
- [39] T. Zhang and D. Freedman, "Tracking objects using density matching and shape priors," in *Proc. Int. Conf. Computer Vision*, France, 2003, pp. 1056–1062.