Segmenting CT prostate images using population and patient-specific statistics for radiotherapy

Qianjin Feng[a)]
Biomedical Engineering College, South Medical University, Guangzhou, China and Department of Radiology, University of North Carolina, Chapel Hill, North Carolina 27510

Mark Foskey[b)]
Department of Radiation Oncology, University of North Carolina, Chapel Hill, North Carolina 27599

Wufan Chen[c)]
Biomedical Engineering College, South Medical University, Guangzhou 510510, China

Dinggang Shen[d)]
Department of Radiology, University of North Carolina, Chapel Hill, North Carolina 27510

(Received 27 April 2010; revised 1 June 2010; accepted for publication 25 June 2010; published 19 July 2010)

Purpose: In the segmentation of sequential treatment-time CT prostate images acquired in image-guided radiotherapy, accurately capturing the intrapatient variation of the patient under therapy is more important than capturing interpatient variation. However, using the traditional deformable-model-based segmentation methods, it is difficult to capture intrapatient variation when the number of samples from the same patient is limited. This article presents a new deformable model, designed specifically for segmenting sequential CT images of the prostate, which leverages both population and patient-specific statistics to accurately capture the intrapatient variation of the patient under therapy.

Methods: The novelty of the proposed method is twofold: *First*, a weighted combination of gradient and probability distribution function (PDF) features is used to build the appearance model to guide model deformation. The strengths of each feature type are emphasized by dynamically adjusting the weight between the profile-based gradient features and the local-region-based PDF features during the optimization process. An additional novel aspect of the gradient-based features is that, to alleviate the effect of feature inconsistency in the regions of gas and bone adjacent to the prostate, the optimal profile length at each landmark is calculated by statistically investigating the intensity profile in the training set. The resulting gradient-PDF combined feature produces more accurate and robust segmentations than general gradient features. *Second*, an online learning mechanism is used to build shape and appearance statistics for accurately capturing intrapatient variation.

Results: The performance of the proposed method was evaluated on 306 images of the 24 patients. Compared to traditional gradient features, the proposed gradient-PDF combination features brought 5.2% increment in the success ratio of segmentation (from 94.1% to 99.3%). To evaluate the effectiveness of online learning mechanism, the authors carried out a comparison between partial online update strategy and full online update strategy. Using the full online update strategy, the mean DSC was improved from 86.6% to 89.3% with 2.8% gain. On the basis of full online update strategy, the manual modification before online update strategy was introduced and tested, the best performance was obtained; here, the mean DSC and the mean ASD achieved 92.4% and 1.47 mm, respectively.

Conclusions: The proposed prostate segmentation method provided accurate and robust segmentation results for CT images even under the situation where the samples of patient under radiotherapy were limited. A conclusion that the proposed method is suitable for clinical application can be drawn. [DOI: 10.1118/1.3464799]

Key words: deformable model, shape statistics, segmentation, prostate CT images

I. INTRODUCTION

Segmentation of the prostate from CT images is an important and challenging task for prostate cancer radiotherapy. The treatment is usually planned on a planning CT on which the prostate and nearby critical structures are manually contoured. The treatment is delivered in daily fractions over a period of several weeks. In image-guided radiotherapy (IGRT), a new CT image is acquired before each individual treatment to enable adjustment of the patient setup. Segmentations of these images are useful for a number of reasons, such as for calculating daily dose to the prostate to judge the progress of the treatment. However, repeated manual

segmentation of images from the same patient is unacceptably burdensome. The purpose of this article is to eliminate the need for multiple manual segmentations by using both shape and appearance information learned from previous days to guide automatic segmentation of new treatment images.

Deformable-model-based segmentation methods, which combine statistical information about organ shape and image appearance, have been proven effective for segmenting the prostate from CT,[1–5] MRI,[6–8] or ultrasound images.[9–13] An extensive review of statistical-shape-model-based medical object segmentation can be found in Ref. 14. Specifically, research efforts on deformable-model-based methods have focused on the following.

I.A. Improving the statistical modeling of shape

A shape model is highly desirable to be as informative as possible by capturing as tightly as possible the amount of variation that might actually occur. If the intrapatient variation for the image being segmented is known, then it is likely to make a better prior because it will generally be a tighter distribution than interpatient variation. To illustrate this, we have investigated 24 anonymous patients in our training set and found that the shape variation within a same patient is generally much less than that between patients. This is evident after glancing at the prostate shapes of different patients shown in Fig. 1(a) and comparing the inter- and intrapatient average shape distances [see Fig. 1(b)]. Both active shape models (ASMs) (Ref. 15) and medial models (termed m-reps)[16] can capture intra- as well as interpatient variation, but it is not a trivial task to do this from population training samples. Obviously, the larger interpatient variation will dominate the outputs of principal component analysis (PCA) in ASMs or principal geodesic analysis (PGA) (Ref. 17) in m-reps and thus cause loss of information of intrapatient shape variation. To eliminate the effect of interpatient variation, the most common approach is to build a solely intrapatient model[1,3] using patient-specific samples; that is, training samples and images to be segmented would come from the same patient. This approach is useful to demonstrate properties of an appearance or a shape model, but it is impractical for clinical radiotherapy because of the difficulty of collecting enough training samples for a patient under therapy. In particular, for a patient's first daily treatment, only one planning image is available for training. Thus, it is uncertain whether or not these approaches will work well.

I.B. Identification of optimal image features for appearance modeling

Numerous features have been proposed to build an appearance model. One class of features that have been tested for prostate segmentation are the *profile-based features*, which consist of sequences of intensities or derivatives acquired along the normal to the object boundary.[9,10] These schemes are effective in situations where objects have distinctive patterns along the boundary normal, such as a high-contrast edge, and when the statistics of the profiles accords

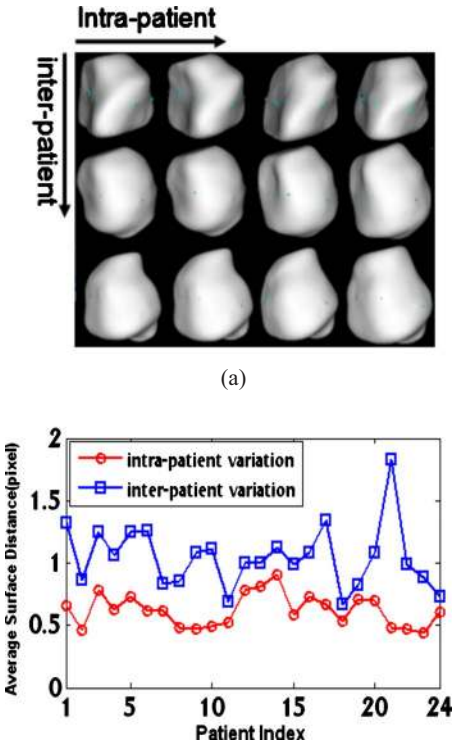

FIG. 1. Comparison between intra- and interpatient variations. (a) Prostate shapes of three patients at four time points. Shapes in each row are from the same patient; shapes in different columns are from different time points. (b) The ASDs among 24 patients. For one patient, the intrapatient ASD is the mean of distances between each shape and the mean shape of that patient, and the interpatient ASD is the distance between the mean shape of that patient and the total mean shape of all patients.

with a specific parametric distribution, e.g., a Gaussian distribution. Heavy variations in soft tissue surrounding the prostate are difficult to capture by these approaches. To avoid this problem, *probability distribution functions* (PDFs) of some chosen photometric variables taken over some regions of interest have been exploited for prostate segmentation.[1–6] These approaches are capable of detecting objects whose boundaries are not well-represented in terms of gradients, perhaps lacking sharp edges. Nonparametric probability distribution functions can be used to capture the prostate appearance variations. Being based on regional information, these features exhibit robustness to noise and to poor initialization. However, region-based features sacrifice some segmentation precision because they depend on the statistics of the whole region, without regard to local image patterns.

Combining multiple candidate features[11,18] is another popular approach for appearance modeling. For these methods, additional complicated feature selection strategies, such as support vector machines (SVMs), usually are adopted to get optimal combination features for a specific segmentation task. Finally, some other image descriptors have been tested for medical object segmentation. The scale invariant feature transform (SIFT), Gabor features, and wavelet features have been employed for prostate segmentation from CT,[19] ultrasound,[12] and MRI images,[7] respectively.

In this article, we propose a novel deformable-model-based segmentation approach for CT prostate segmentation. There are two main contributions of our method. First, it combines cross-patient information with patient-specific information in a weighted fashion that depends on how much patient-specific information is available. To create the initial model used for the first daily treatment, the shape and appearance from the planning image are combined with the statistics of intrapatient variation learned from the population. As more subsequent images are segmented, the shape and appearance statistics are updated in an online learning process, and the patient-specific information takes an increasing role in capturing the variation in the specific patient more accurately. In our second main contribution, to address limitations of region- and profile-based features, we introduce a combination feature, called the gradient-PDF combination feature, which produces more accurate and robust segmentations than general profile-based features. This also extends our prior work.[19] Here we replace the SIFT features adopted in Ref. 19 with the proposed gradient-PDF combination features to save computational cost.

II. METHOD

Our method has three main components. First, there is a statistical model of prostate shape, represented implicitly as a collection of the corresponding surfaces for a range of prostate shapes. Second, there is a model of image appearance, embodied in the objective function that is optimized during segmentation. Finally, there is an online learning mechanism that incorporates patient-specific information as additional images are segmented.

The shape model is developed by a procedure we refer to as "surface construction," in which a surface deformation algorithm is used to generate the individual surface of each prostate in the training dataset. A template-based framework is used to ensure that all created surfaces have good point-to-point correspondence. In the appearance model, we aim to improve the accuracy and robustness of our method by introducing local intensity PDFs, combined with gradient features, to guide the segmentation. In the online learning mechanism, information from each just-completed segmentation is incorporated into the patient's appearance and shape statistics so that the weight of patient-specific information can be enlarged gradually from day to day, better corresponding to that patient's images. The following sections describe each part in detail. In Sec. II A, we describe the surface construction method for generating the shape model. In Sec. II B, we describe the method for segmenting an individual image, focusing on the appearance model. Finally, in Sec. II C, we discuss how the statistical-shape model can be modified to reflect the within-patient statistics gathered over the course of treatment.

II.A. Surface construction

Clearly, the accuracy of the shape model is largely dependent on the performance of shape correspondence, and thus developing more accurate and efficient shape correspondence algorithms for use in statistical-shape analysis has been widely investigated over the past several years.[20–23] In this study, we developed a deformable-model-based algorithm to fit a surface to each segmented prostate image in the training set. Also, a template-based framework is employed to ensure all the generated surfaces have good point-to-point correspondence. At the time of segmentation, each surface will be registered to the existing planning surface for the patient and the modes of variation will be calculated as described in Sec. II C 1.

II.A.1. Deformable-model-based algorithm for surface construction

A shape is represented by a triangle mesh $s=(V,T)$, where $V=\{v_i \in \mathbb{R}^3\}_{i=1}^N$ is the set of N mesh vertices and $T=\{t_i \in \mathbb{Z}_+^3\}_{i=1}^M$ is the set of M triangles, with each triangle represented as a triple of vertex indices. The cost function of the proposed surface deformation algorithm is as follows:

$$s^* = \arg\min_s \{\boldsymbol{E}^{\text{Distance}}(s,\boldsymbol{I}_{\text{seg}}) + \alpha \boldsymbol{E}^{\text{Smooth}}(s)\}, \quad (1)$$

where $\boldsymbol{E}^{\text{Distance}}=\sum_{i=1}^N |v_i - c_i|$ and c_i is the point of intersection between the surface normal at v_i and the prostate boundary in the segmented image $\boldsymbol{I}_{\text{seg}}$. Obviously, by minimizing $\boldsymbol{E}^{\text{Distance}}$, all surface vertices should move to the prostate boundary. $\boldsymbol{E}^{\text{Smooth}}$ is a smooth regularization term. Equation (1) is iteratively optimized by alternately minimizing $\boldsymbol{E}^{\text{Distance}}$ and smoothing the new generated surface to minimize $\boldsymbol{E}^{\text{Smooth}}$. The smooth regularization term is crucial to the robustness of the algorithm, especially for the initial iteration steps, where the surface deformation is generally large. However, a simple surface smoothing algorithm will lead to an additional erroneous shrinking of the surface. To address this issue, the smoothing method proposed by Taubin[24] is employed here. This method is capable of smoothing the surface while preventing additional surface shrinkage. The parameter α is used to weigh the smoothness constraint dynamically, e.g., α is designed to be large initially and decreases later during the progression of the algorithm. In our experiment, we find that dynamically adjusting the parameter α from 1 to 0.2 according to an exponential function during iterative optimization leads to robust and accurate results. More details on this weighing can be found in Sec. II.D.

II.A.2. Template-based framework for surface construction

To obtain point correspondences of the surfaces, a template-based framework similar to the method proposed in Ref. 22 is adopted (see Fig. 2). In this framework, all segmented prostate images in the training dataset are first affinely aligned to a uniform space, after which the aligned images are averaged, and the resulting mean image is thresholded to get a template image. In this uniform space, a predefined cylinder surface is deformed to match this template image using the method described in Sec. II A 1, yielding a model surface. Next, this model surface is transformed back to each segmented prostate image's space using the inverse

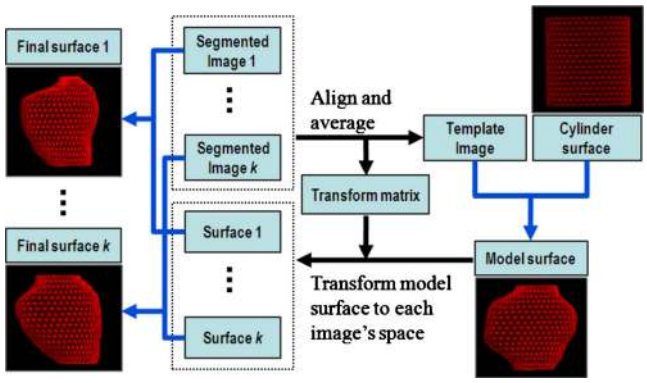

FIG. 2. Flow diagram of template-based surface construction framework. The arrows with light color (blue in online version) represent the surface deformation procedures using the method proposed in Sec. II A 1.

of the transform matrix created in the previous alignment step. Finally, these warped model surfaces are respectively refined according to their corresponding segmented images using the method described in Sec. II A 1, and thus the boundary surfaces of individual prostates are obtained. Because each prostate boundary model is derived from a common model that is based on an average of all aligned prostates, the corresponding vertices can be expected to approximately reflect geometrically corresponding locations in the different prostate images.

II.B. Deformable-model-based segmentation

In this section, we describe how a new image is segmented once a statistical-shape model has been trained, focusing on the appearance features that are incorporated into the objective function. Before being segmented, the new image $\boldsymbol{I}$ is normalized to a common reference frame, as described below in Sec. II B 4, and it will be referred to as $\boldsymbol{I}^{\text{norm}}$.

In order to get robust and accurate segmentation, we build the appearance model using image features based on intensity PDFs over local regions, as well as profile-based gradient features. Each feature, whether it is based on profiles or PDFs, is tied to a particular vertex of the surface model, which we will often refer to as a landmark. A profile feature consists of an ordered sequence of samples along a normal to the surface at a landmark, and a PDF-based feature consists of a probability distribution of values taken over a region surrounding the landmark. The challenge is how best to combine these features. As discussed in Sec. I, the region-based feature is more robust to noise and poor initialization, but somewhat less precise, while the profile-based feature is effective for landmarks that have consistent, distinctive patterns along the surface normal. Accordingly, two motivations need to be considered. *First*, each landmark should be assigned an optimal feature which can be determined by the image pattern of its immediate surroundings. *Second*, region-based features should take a more important role at the initialization and beginning stages of the optimization process to obtain a good initialization, while the profile-based features should gradually begin to dominate later in the process to get a more accurate segmentation. In other words, we need a mechanism to adjust the weight between profile- and region-based features during the optimization.

Some classifiers such as SVMs are commonly used to select appropriate subfeatures to compose optimal features. However, for these methods, the feature selection processes are implicit and it is impossible to dynamically adjust the weights of subfeatures during segmentation. The most intuitive approach would be to integrate profile- and region-based features through a linear combination. However, since the value ranges of different subfeatures are so different, determining the coefficients of each subfeature is not a trivial task.

Based on the above analyses, we propose a novel approach to integrate the profile-based gradient feature and local-region-based PDF feature. For each specific vertex of the surface, only one subfeature is selected according to a predefined rule (described in Sec. II B 1). This essentially is the simplest case of linear combination of these two kinds of features, i.e., the values of the coefficients of the two subfeatures are limited to 0 or 1. From the view of the whole appearance model, these two kinds of features are integrated together since they separately are used at different landmarks simultaneously. Here the set of vertices $V=\{v_i \in \mathbb{R}^3\}_{i=1}^{N}$ is divided into two subsets denoted as V^{PDF} and V^{GRA}, which represent the vertex subsets that use PDF features and gradient features, respectively. In this scheme, the weight between subfeatures can be adjusted easily by reassigning vertices from one set to another.

II.B.1. Cost function

We use the following cost function based on gradient-PDF combination features:

$$\boldsymbol{s}^* = \arg\min_{\boldsymbol{s}}\{\boldsymbol{E}^{\text{Gradient}}(V^{\text{GRA}}) + \boldsymbol{E}^{\text{PDF}}(V^{\text{PDF}}) + \beta \boldsymbol{E}^{\text{Smooth}}(\boldsymbol{s})\}, \quad \boldsymbol{s} \in \boldsymbol{D}. \tag{2}$$

There are three terms in this function. The first two terms are image-based terms using gradient and PDF features, respectively. Bear in mind that $\boldsymbol{E}^{\text{Gradient}}$ and $\boldsymbol{E}^{\text{PDF}}$ are separately defined over V^{GRA} and V^{PDF}, so these two terms only act on the vertices in their corresponding subset. The third term is a smooth regularization term. $\boldsymbol{D}$ is the shape space learned from the training data. Deformed surfaces generated according to this function are constrained to this space.

The $\boldsymbol{E}^{\text{Gradient}}$ term: The energy term $\boldsymbol{E}^{\text{Gradient}}$ represents the differences in the corresponding gradient features between the model and the current image $\boldsymbol{I}^{\text{norm}}$. Minimizing it requires that the gradient features learned in image training match with the gradient features found in the image.

The profile-based gradient feature is similar to that described in Ref. 15. Each profile is a vector of gradient magnitudes of image intensities, sampled at locations separated by a fixed step size along the surface normal. The length of each profile varies for different landmark points according to

a model described in Sec. II B 3 below, with the goal of minimizing the inclusion of gas or bone intensities. Based on gradient features, $\boldsymbol{E}^{\text{Gradient}}$ is defined by

$$\boldsymbol{E}^{\text{Gradient}}(V^{\text{GRA}}) = \sum_{v_i \in V^{\text{GRA}}} \frac{1}{M_{v_i}} \sum_{i=1}^{M_{v_i}} \left(\frac{|\boldsymbol{g}_{v_i}(l) - \bar{\boldsymbol{g}}_{v_i}(l)|}{\sigma_{v_i}(l)} \right), \quad (3)$$

where M_{v_i} is the length (number of entries) of the profile at vertex v_i, $\bar{\boldsymbol{g}}_{v_i}(l)$ represents the lth entry of the average profile at vertex v_i from the model space, and $\boldsymbol{g}_{v_i}(l)$ is the lth entry of the corresponding profile from the image $\boldsymbol{I}^{\text{norm}}$. The quotient value $\sigma_{v_i}(l)$ is the standard deviation of the lth entry of the profiles at vertex v_i.

The $\boldsymbol{E}^{\text{PDF}}$ term: Similar to the $\boldsymbol{E}^{\text{Gradient}}$ term, $\boldsymbol{E}^{\text{PDF}}$ denotes the differences in the local intensity PDF features at the vertices in V^{PDF} between the model and the image. The local neighborhood of a given vertex v_i was split into inside (prostate) and outside (background) subregions using the surface mesh, then the local interior PDF $P^{\text{in}}_{v_i}$ and local exterior PDF $P^{\text{out}}_{v_i}$ can be estimated easily by using intensities from these two subregions, respectively.

In this article, the Bhattacharyya distance[25,26] is used to measure the difference between two distributions for its simplicity. The Bhattacharyya distance is a measure of similarity between two PDFs, P and P', defined as

$$B(P,P') = \int \sqrt{P(x)P'(x)}dx, \quad (4)$$

where $x \in X$ is an intensity variable living in some intensity space X. This measure varies between 0 and 1, where 0 indicates a complete mismatch and 1 indicates a complete agreement between two PDFs. Thus $(1-B)$ can be used to measure the dissimilarity (or difference) between two PDFs. Based on the Bhattacharyya distance, $\boldsymbol{E}^{\text{PDF}}(V^{\text{PDF}})$ is defined by

$$\begin{aligned}\boldsymbol{E}^{\text{PDF}}(V^{\text{PDF}}) = \Sigma_{v_i \in V^{\text{PDF}}} \{ &[1 - B(P^{\text{in}}_{v_i}, \bar{P}^{\text{in}}_{v_i})] + B(P^{\text{in}}_{v_i}, \bar{P}^{\text{out}}_{v_i}) \\ &+ [1 - B(P^{\text{out}}_{v_i}, \bar{P}^{\text{out}}_{v_i})] + B(P^{\text{out}}_{v_i}, \bar{P}^{\text{in}}_{v_i}) \}. \end{aligned} \quad (5)$$

Here, $\bar{P}^{\text{in}}_{v_i}$ is the interior PDF attached to v_i from the model estimated over the training dataset and $P^{\text{in}}_{v_i}$ is its corresponding interior PDF estimated from the current image $\boldsymbol{I}^{\text{norm}}$. $\bar{P}^{\text{out}}_{v_i}$ and $P^{\text{out}}_{v_i}$ are defined in the same way for the exterior. Minimizing $\boldsymbol{E}^{\text{PDF}}$ essentially aims to find new locations of v_i, where $P^{\text{in}}_{v_i}$ is more like $\bar{P}^{\text{in}}_{v_i}$ and less like $\bar{P}^{\text{out}}_{v_i}$, while $P^{\text{out}}_{v_i}$ is more like $\bar{P}^{\text{out}}_{v_i}$ and less like $\bar{P}^{\text{in}}_{v_i}$.

II.B.2. Feature selection

As described earlier in Sec. II B, how to choose the sets V^{PDF} and V^{GRA} is a key point for building the gradient-PDF combination features. We use the following principle: For landmarks whose gradient features are consistent over the training dataset, gradient features are selected. Otherwise, PDF features are used. Here we define a measure C_{v_i} of the consistency of gradient features at a given vertex v_i by

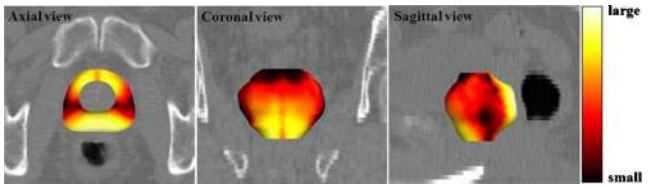

FIG. 3. Demonstration of the map of the prostate. In this figure, a prostate surface was color coded according to the value at each vertex.

$$C_{v_i} = M_{v_i} / \Sigma_l^{M_{v_i}} \sigma_{v_i}(l), \quad (6)$$

where $\sigma_{v_i}(l)$ is the standard deviation of the lth entry of the gradient feature defined in Sec. II B 1. In other words, C_{v_i} is the reciprocal average of the standard deviations so that a profile with a large C_{v_i} will have smaller standard deviations for its entries, and thus be more consistent across images and thus more prone to be selected for the appearance modeling. Figure 3 illustrates how C_{v_i} varies across a prostate boundary. In order to adjust the relative weights of the gradient features and PDF features dynamically, all the vertices are ranked according to the C_{v_i} of their gradient features, after which a certain proportion of vertices with smaller C_{v_i} is selected to compose the subset V^{PDF}, and the remaining ones make up V^{GRA}. A parameter ρ is defined to control the sizes of V^{GRA} and V^{PDF} as follows:

$$\rho = N^{\text{GRA}}/N, \quad (7)$$

where N^{GRA} is the number of vertices in V^{GRA} and N is the number of vertices in V. Thus, the number of vertices in V^{PDF} is equal to $(N - N^{\text{GRA}})$. In the iterative optimization stage, the parameter ρ is increased from 0 to 1 according to an exponential function, aiming to dynamically weigh the gradient features and PDF features.

II.B.3. Rectum gas and bone

The presence of rectum gas and bony structures near the prostate boundary is troublesome for segmentation. Regions of gas appear as black blobs surrounded by gray tissues, and there is no consistency from image to image in which, or even whether, gas can be expected to appear, thus making profile features inconsistent. In the same way, bone has much higher voxel intensities which create strong edges near the true boundary of the prostate, and there is enough variability in the relative positions of the prostate and bony structures where the strong edges of bone often lead to large errors of segmentation. Because of these problems, samples from regions of gas and bone should be regarded as outliers in feature extraction for prostate segmentation. Other work has been done to address these issues. Broadhurst *et al.*[27] separately modeled the probability of gas and bone tissue intensities, and Davis *et al.*,[28] in the context of deformable registration, introduced a "deflation approach" to deal specifically with the problem of gas. In our method, since two classes of features are employed to build the appearance model, two

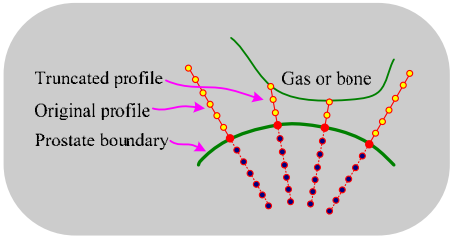

FIG. 4. Demonstration of profile truncation. The profiles which lie across the regions associated with gas or bone are truncated. In this figure, there are two profiles truncated to an appropriate length.

different strategies are proposed to alleviate the effect of gas and bone: One for the profile-based gradient features and one for the local PDF features.

For the profile-based features, the profile length at each landmark is determined by statistically investigating the intensity profile of all training images. The goal is to choose a profile length for each vertex v_i so that, across all training images, no entry in v_i's profile has a greater than 50% probability of being in bone or gas. Making this choice is a three-step process. First, the intensity profiles of all training samples are calculated, and then the samples located in gas and bone regions are detected according to their value in Hounsfield units (HU) by using two predefined thresholds (-150 HU for gas and $+100$ HU for bone). Second, in the model space, for each entry of each profile, the probability of being an outlier is estimated by using the proportion of outliers at this location in the training dataset. Then a threshold for probability is selected to determine whether this sample is an outlier. In this article, the threshold is 0.5. In the *last* step, each profile is truncated from the outlier location closest to the surface to the outside end point. The resulting length is taken as the optimal length for each profile (Fig. 4). The optimal length profiles avoid capturing the voxel intensities of gas and bone regions, and thus the segmentation results should suffer less from the inconsistency in image appearance across the training samples in those parts of the image. Figure 5 shows some examples to illustrate the effectiveness of the optimal length profile.

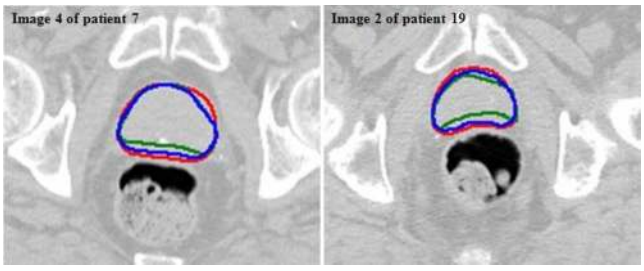

FIG. 5. Typical results illustrate the effectiveness of the optimal length profile. The dark (red in online version) contours show the hand-drawn ground-truth contours, and the light (green in online version) and gray (blue in online version) contours represent the results using the fixed length profiles and the proposed optimal length profiles, respectively. It is clear that the segmentation accuracy is significantly improved for the boundary segment adjacent to the rectum in (a) and (b), as well as for the boundary segment adjacent to the bone in (b) by using the proposed optimal length profiles.

For the local PDF features, it is easy to alleviate the effect of gas or bone. The appearance model and fitting algorithm are applied to the images using values $[-150$ HU, $+100$ HU$]$ to build 125-bin histograms, and the extremely low and high CT values corresponding to gas and bone are naturally thresholded out to remove the effects of these outliers.

II.B.4. Initialization and optimization strategy

In our method, for both training and deformable segmentation steps, all calculations are done in a specified benchmark frame of reference, which we define to be the space of the planning image for the patient being treated. All images and surfaces are transformed to this benchmark space before any further operations. When a new treatment image is acquired, it is rigidly transformed into the benchmark space according to the pelvic bone. Thus, we obtain a pose-normalized image, denoted as $\boldsymbol{I}^{\text{norm}}$. It is worth noting that the model shape $\boldsymbol{s}^{\text{mdl}}$, i.e., the mean shape $\boldsymbol{s}^{\text{mean}}$ of the current patient (defined in Sec. II C 1), is in the same space as $\boldsymbol{I}^{\text{norm}}$. Thus, $\boldsymbol{s}^{\text{mdl}}$ will be close to the prostate in $\boldsymbol{I}^{\text{norm}}$. For a more accurate initialization, $\boldsymbol{s}^{\text{mdl}}$ is shifted and rotated over a limited range while the cost function defined by Eq. (2) is evaluated. When this function reaches its minimum, a good initialization is obtained. Here, we are not guaranteed to find the global minimum of the cost function, but we can get a local optimum that is a good solution for the initial position of the model, and the experimental results have shown it was good enough to ensure the robustness and accuracy of the final segmentation. At this initialization stage, the parameter ρ defined by Eq. (7) is set to 0, i.e., only the local PDF features are used to guide the initialization, for greater robustness.

To solve Eq. (2), an iterative optimization strategy is used to compute the deformed surface by alternately minimizing the energy term (without regard to the shape space) and then correcting the new surface into the learned shape space $\boldsymbol{D}$. At the tth iteration step, we perform the following: First, the deformable surface $\boldsymbol{s}^t$ is updated via a local search around its current location so that the newly updated surface $\hat{\boldsymbol{s}}^t$ has a better match to the gradient and PDF features of the testing image, which yields a smaller $\boldsymbol{E}^{\text{Gradient}}(\hat{\boldsymbol{s}}^t)+\boldsymbol{E}^{\text{PDF}}(\hat{\boldsymbol{s}}^t)$.

Second, the newly updated surface $\hat{\boldsymbol{s}}^t$ is smoothed using the Taubin method[24] and a smooth surface is obtained, denoted as $\hat{\boldsymbol{s}}^t_{\text{smo}}$. To dynamically adjust the strength of the smoothness constraint, the surface $\hat{\boldsymbol{s}}^t$ and $\hat{\boldsymbol{s}}^t_{\text{smo}}$ are combined using $\hat{\hat{\boldsymbol{s}}}^t_{\text{smo}}=\hat{\boldsymbol{s}}^t+\beta(\hat{\boldsymbol{s}}^t_{\text{smo}}-\hat{\boldsymbol{s}}^t)$ to get a partially smoothed version of $\hat{\boldsymbol{s}}^t$, denoted as $\hat{\hat{\boldsymbol{s}}}^t_{\text{smo}}$. This is the same way as the smoothness constraint is applied during surface construction (Sec. II A 1), with the parameter β here adjusted the same as α in Eq. (1).

Third, the population-based and patient-specific shape statistics are used to constrain the newly created smooth surface $\hat{\hat{\boldsymbol{s}}}^t_{\text{smo}}$. A reasonable surface $\boldsymbol{s}^t_{\boldsymbol{D}}$ is obtained by finding the nearest surface to $\hat{\hat{\boldsymbol{s}}}^t_{\text{smo}}$ in the learned shape space $\boldsymbol{D}$.

Last, the surface $\hat{\hat{s}}^t_{\text{smo}}$ and the surface s^t_D from the shape space are combined by a weighted average using $s^{t+1}=\hat{\hat{s}}^t_{\text{smo}}+\gamma(s^t_D-\hat{\hat{s}}^t_{\text{smo}})$ to get the final surface s^{t+1} of this iteration step. The parameter γ is used to adjust the strength of the shape constraint. In our implementation, γ was decreased from 1 to 0.2 according to an exponential function to relax the shape constraint gradually during iterative optimization. The rationale is that the latter iteration steps are trying to improve a segmentation that is already somewhat accurate, so image information is more trustworthy and should be given more weight relative to the shape prior. It is reasonable and experimental results have justified its robustness.

II.C. Online learning mechanism

We propose an online learning mechanism with the aim of accurately capturing shape and appearance statistics of the patient under treatment. At the beginning, we use the shape and appearance of the planning image as the mean shape and the mean appearance, with the residue derived from the training set as our residue model. As more images are segmented, the shape and appearance statistics are updated online, and the patient-specific information takes a larger and larger role. There are two major advantages of this approach. First, it becomes tractable to establish a shape and appearance model, even if the samples of the current patient are very limited, by using the residue information from the population. Second, by increasing the weight of patient-specific information as more samples of the current patient are processed, the models will more accurately capture the shape and appearance variation of the current patient's prostate, leading to improved segmentation.

II.C.1. Shape model online learning

As mentioned in Sec. II B 4, for the training step, all the calculations are performed in the space of the planning image for the patient being treated. All surfaces are affinely aligned to the known surface already provided in the planning image by a least-squares fit before further operations. Figure 6 demonstrates the main procedures of shape model online training.

First, for each treatment image of a given patient sample, its surface is mapped to the benchmark space of the current patient, yielding a surface that we denote by s^{align}. The mean shape s^{mean} of each patient is calculated by averaging the corresponding vertex coordinates of the s^{align} surfaces, and residual shapes s^{res} are obtained by subtracting the mean shape s^{mean} from the aligned shape s^{align}. This process is performed for all patient samples and also for the prior images of the patient under treatment.

Second, a weighted PCA is done to all s^{res}, with weight factors ω^s for the current patient and $1-\omega^s$ for the training set. By adjusting the parameter ω^s, we can flexibly control the relative weights between the patient-specific and population information.

Last, after performing weighted PCA, a reasonable residual shape space is obtained. Then, by just shifting this space using s^{mean} of the current testing patient, we can get a reasonable shape space $\boldsymbol{D}$ for the current patient. Thus, $\boldsymbol{D}$ is available to guide the deformable model to segment the later time images of the current patient. In the segmentation step, we use the mean shape of the current patient as the model shape, denoted s^{mdl}.

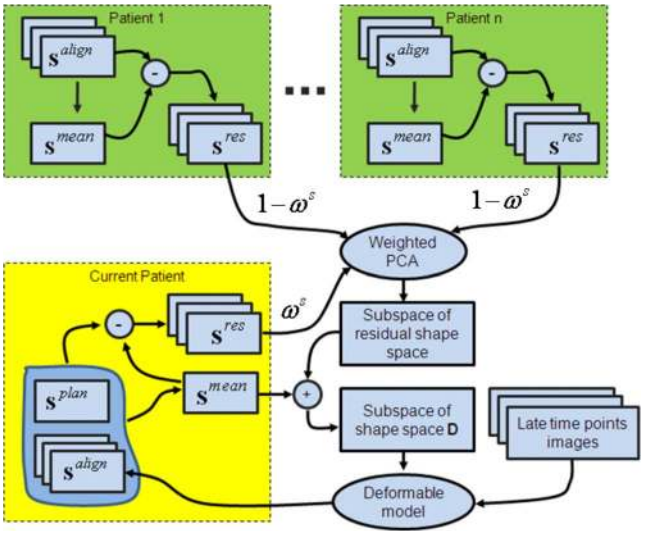

FIG. 6. Flow diagram of shape model online training. In this figure, the blocks with dark color (green in online version) represent the population information, and the light (yellow in online version) one represents the patient-specific information.

It should be noted that the training is dynamic: Along with each new treatment image that is segmented, s^{mean} and s^{res} of the current patient are changed. At the same time, we increase ω^s according a sigmoid function (see Fig. 7). Thus, as more subsequent images of the current patient are acquired, the patient-specific shape statistics derived from those images gain more influence. In this framework, the mean shapes of other patients are replaced by the mean shape of the current patient by doing the PCA on the residual shape space and adding the mean shape of the current patient to the resulting residual shape subspace. This process reduces the effect of interpatient shape variation.

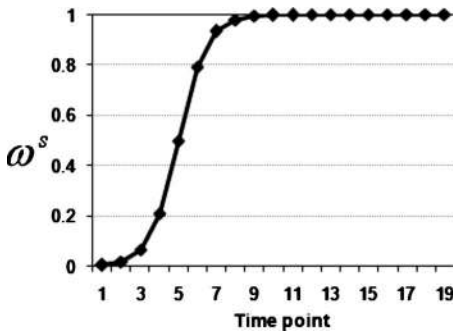

FIG. 7. Illustration of the values of parameter ω^s at different time points. In this figure, after the number of captured images of the current patient reaches 9, it approximately equals 1. That implies that patient-specific information totally controls the shape model after that time.

II.C.2. Appearance model online learning

For the gradient feature, similar to the use of the shape statistics, the average feature vectors $\bar{g}$ come from the current patient, while the standard deviation σ is a weighted combination of the standard deviation from the current patient and that from the population. The weight factor ω^s is dependent on the number of images we have captured and processed from the current patient. This is consistent with the use of the patient-specific shape statistical model.

For the PDF feature, the interior PDFs and exterior PDFs are estimated only using the intensities sampled from images of the current patient. Here the population information is not employed for the estimation of the PDF features. The reason we abandon the population information is that we do not require statistical information about the day-to-day variability of these PDF features from other patient samples.

III. EXPERIMENTAL RESULTS

III.A. Testing data and quantitative measures

Our data consist of 24 patient image sets, each with more than nine daily CT scans of the male pelvic area taken during a course of radiotherapy, with a total image count of 330. The images have an in-plane resolution of 512×512 with voxel dimensions of 0.98×0.98 mm^2 and an interslice distance of 3 mm. We are also provided with expert manual segmentations of the prostate in every image. Considerable efforts were made to minimize the intraoperator variability. A careful protocol was established by a radiologist and a radiation oncologist working together. Earlier segmentations were reviewed in comparison with later ones to make sure that consistency was maintained. For each patient, manual contouring was done after rigidly registering all of that patient's images together. Successive images could be viewed overlaid on one another so that consistency could be maintained in how each day was segmented.

We consider the patients separately, i.e., segmenting the images from one patient in a leave-one-patient-out strategy. For one patient whom we want to test, the first treatment image is regarded as the planning image and the remaining images are segmented for evaluation. At same time, the patient-specific and population training datasets are built by using processed images of this patient and all images of other patients, respectively. So, leaving out the planning image of each patient, there are a total of $330-24=306$ test images for evaluating the performance of segmentation. Two quantitative measures are used to evaluate the performance of the algorithms, i.e., the Dice similarity coefficient (DSC) (Ref. 29) and the average surface distance (ASD) between the automated and the manual segmentation results.

Our method was performed on an Intel Core 2 2.33 GHz processor. The initialization required 15.0 s. For the gradient feature, the surface deformation step required 42.8 s; if the gradient-PDF combination feature was used, the time increased to 66.2 s. In the online learning step, the shape model and the appearance model have to be updated, requiring 14.6 s. The total computing time is 95.8 s.

III.B. Experiments

A series of experiments was performed to evaluate the effectiveness of the proposed segmentation algorithm as outlined next.

Evaluation of gradient-PDF combination features: In order to evaluate the effectiveness of the gradient-PDF combination features, we carried out a comparison between gradient features taken alone and gradient-PDF combination features in our algorithm.

Evaluation of the online learning mechanism: In this experiment, we tested the performance under three different update strategies, labeled as follows: (1) Partial online update, (2) full online update, and (3) manual modification before online update.

Under the *partial online update* strategy, for the shape model, the mean shape of the current patient is recalculated after each new image is segmented, while the residual shape subspace calculated from the population training dataset remained unchanged. For the appearance model, the average model gradient profiles $\bar{g}$ and model PDFs $\bar{P}^{\text{in}}$ and $\bar{P}^{\text{out}}$ are recalculated and updated online, but the standard deviations σ of the gradient profiles are those of the population and not updated. Under the *full online update* strategy, all parameters were updated online according to the procedure described in Sec. II C.

Under the *manual modification before online update* strategy, the previous automatic segmentation results are refined manually before the next online update. The reasons that we employ this strategy are as follows: (1) We aim to get more accurate appearance and shape statistics to improve the segmentation performance and (2) this strategy is clinically feasible. That is, checking and repairing segmentations for the first few days may be faster than manually segmenting them from scratch, and once good patient-specific training data is achieved, additional manual corrections may not be necessary. Also, it may be possible to run this algorithm without corrections with the patient on the table, so that the results can be used to adjust the treatment after they are deemed acceptable by the therapist. Then the physician can refine the model to improve the training at their convenience.

Evaluation of the comparison with the existing methods: In order to evaluate the effectiveness of our method, we carried out comparison with two deformable-model-based methods proposed in Refs. 1 and 3.

III.B.1. Effectiveness of gradient-PDF combination features

The performance of the gradient-PDF combination features is evaluated by way of a comparison between the segmentation results for the gradient-PDF combination features and gradient features. We segmented 306 images of 24 patients (except the planning image of each patient) using these two kinds of features, respectively.

Table I shows the average DSC and the average ASD of all 306 segmentation results. It can be seen that when the gradient-PDF combination features are used, the average DSC (89.3%) is much higher than that of gradient features

TABLE I. Average DSC and average ASD between the manual and the automated segmentations of all 306 images using gradient features and gradient-PDF combination features, respectively.

		Mean ± std	Min	Max	Median
DSC (%)	Gradient	86.8 ± 11.3	19.1	96.9	90.2
	Gradient+PDF	**89.3 ± 5.0**	**42.4**	**96.7**	**90.6**
ASD (mm)	Gradient	2.47 ± 1.57	0.72	10.48	1.98
	Gradient+PDF	**2.08 ± 0.79**	**0.69**	**7.89**	**1.87**

(86.8%). Also, the average ASD is reduced from 2.47 to 2.08 mm. We performed a paired t-test on the DSC values, a one-tailed test shows that the DSCs for the gradient-PDF combination feature comparisons are significantly ($p < 0.001$) greater than the DSCs for the gradient-only feature comparisons. These results indicate that the segmentation accuracy is improved appreciably by using gradient-PDF combination features against using gradient features.

In addition, when we investigate all results of these 306 images (see Fig. 8), it is clear that the resulting curves using only gradient features have more spikes than those using gradient-PDF features. These spikes correspond to the bad segmentation results. As a DSC value of 0.7 or greater is generally considered to be a high level of coincidence between segmentations,[30] we call the segmentation a failure if its DSC value is less than 0.7. Under this measure, we find that the 304 results are successful and the success ratio is 99.3% for using gradient-PDF combination features, while for the gradient features case, the number of successful segmentation is 288 with a success ratio of 94.1% only. In short, by using gradient-PDF combination features, the success ratio increases by 5.2%. These results indicate that the proposed method using combination features is more robust.

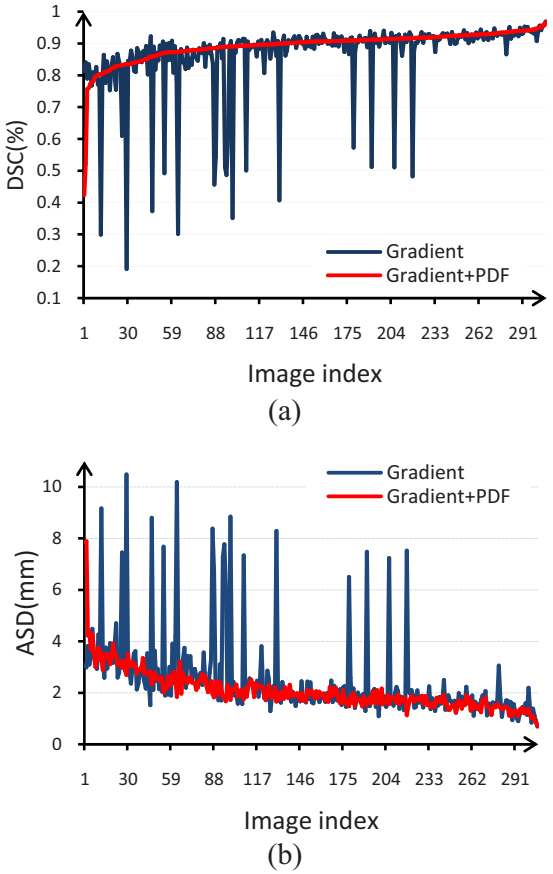

FIG. 8. DSCs (a) and ASDs (b) of all 306 images by using gradient and gradient-PDF combination features, respectively. The image order is ranked according to the DSC values of segmentation results using gradient-PDF combination features for better illustration.

The above results can be justified as follows. First, the profile-based gradient features use only a few samples along the boundary normal and are thus sensitive to noise. Further, the Gaussian model assumption of their distribution limits their capability to capture heavy variations in soft tissue surrounding the prostate. Second, on the other hand, PDFs based on statistical information of the regions exhibit robustness to noise and poor initialization. Additionally, without imposing a specific parametrization, e.g., a Gaussian distribution used for the profile-based features, region-based features are beneficial to capture large variations in the prostate appearance. Therefore, the gradient-PDF combination features can use the advantage of both types of features and thus provide better segmentation results, as shown above.

III.B.2. Effectiveness of the online learning mechanism

Table II shows the average DSC and average ASD of all 306 segmentation results. Compared to the partial online update strategy, when we explore the full online update strategy, the average DSC is improved from 86.6% to 89.3% with 2.8% increase, and the average ASD decreased from 2.34 to 2.08 mm. It is clear that the full online update strategy produces more accurate segmentation results. In addition, when the third strategy—manual modification before online update—is adopted, we get the best segmentation performance, with the average DSC and average ASD reaching 92.4% and 1.47 mm, respectively. Compared to the partial online update, the average DSC is improved for 4.8% from 86.6% to 92.4%, and the average ASD is decreased for 0.87 mm from 2.34 to 1.47 mm. Figure 9 shows a typical result with the DSC 92.2%. From this figure, we can see the prostate boundary created by the proposed method is very close to that of expert manual segmentations when the DSC is about 92%.

To justify the validity of the proposed method, we further investigate whether the use of the online learning mechanism can improve the segmentation for the initial treatment images, or whether the segmentation accuracy can be improved

TABLE II. Average DSC and average ASD between the manual and the automated segmentations of all 306 images under three different update strategies, respectively.

		Mean ± Std	Min	Max	Median
DSC (%)	Partial online update	86.6 ± 5.8	32.1	96.7	87.7
	Full online update	89.3 ± 5.0	42.4	96.7	90.6
	Manual modification	**92.4 ± 3.4**	**64.2**	**96.7**	**93.3**
ASD (mm)	Partial online update	2.34 ± 1.0	1.88	14.99	2.15
	Full online update	2.08 ± 0.79	0.72	7.89	1.87
	Manual modification	**1.47 ± 0.57**	**0.71**	**6.16**	**1.33**

by collecting more treatment images. Figure 10 shows our obtained average DSCs of 24 patients at each treatment day. It can be observed that the use of the online learning mechanism improves segmentation even for the initial treatment images, and segmentation keeps improving until a sufficient number of treatment images have been collected for training patient-specific models. We calculated the average DSC and average ASD of 24 patients over the last 9 (9–17) treatment days, finding that the average DSC and average ASD reach 93.0% and 1.38 mm, respectively. Compared to the partial online update, the average DSC is improved from 86.2% to 93.0% with a 5.8% increase, and the average ASD is decreased from 2.40 to 1.38 mm with an error reduction of 1.02 mm.

III.B.3. Comparison with the other methods

To compare to the method proposed in Ref. 3, which has the best segmentation accuracy in literature to the best of our knowledge, our method was further evaluated on the same dataset as used in Ref. 3. This dataset consists of 80 images from five patients, each having 13–18 images from multiple treatment days. For our method, due to the fact that requirement of a planning image is required for training, there are 80−5=75 test images for evaluation in our experiment, instead of 80 test images in Ref. 3. Ignoring this negligible difference, our results were compared to those reported in Ref. 3. We find that our results are comparable, if with a slight decrease in accuracy. The average DSC of five patients is decreased from 93.0% to 92.1% with a 0.9% decrease, and the average ASD is increased from 0.99 to 1.43 mm with an error increase of 0.44 mm. It is worth noting the difference of the training strategies used in the two methods, with our training strategy more suitable for clinical application. In Ref. 3, for each patient, the authors successively left each day image out, trained on all remaining days, and segmented the left-out day image using the trained shape and appearance statistics. Namely, even for the patient with the fewest daily images (13), there are 12 patient-specific images available for training for the method described in Ref. 3. In our method, to segment a given day's image, only the preceding daily images of the same patient are used for gathering the patient-specific information, i.e., if you want to segment the second day image, only the first day image is available and used for training. Obviously, our segmentation framework more coincides with the actual clinical condition of radiotherapy.

In addition, each registration in Ref. 3 was begun with an initialization by means of manually placed landmarks. Our method is fully automatic for successive patient images, requiring no human intervention.

To compare our method to that in Ref. 1, we quantified our results using two measures v_d and V_{fa}, which are used in Ref. 1. The measure v_d is calculated as the fraction of the ground-truth organ that was contained by the estimated organ, while V_{fa} is calculated as the fraction of the estimated organ that lies outside the ground-truth organ. For a good segmentation, v_d and V_{fa} should be close to 1 and 0, respectively. The average median v_d and average median V_{fa} of 24 patients of our method are 91.1% and 7.0%, as compared to

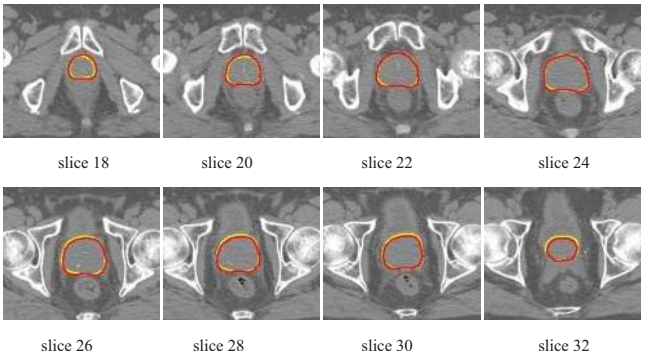

FIG. 9. Segmentation results for evenly spaced slices of image 8 of patient 24 (DSC of 92.2%). The light (yellow in online version) contours show the result of the proposed method. The dark (red in online version) contours show the hand-drawn (ground-truth) contours supplied by a radiation oncologist.

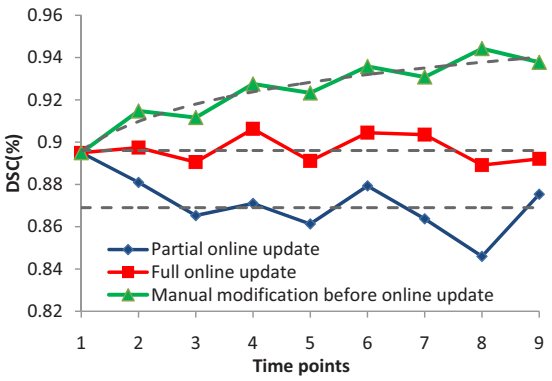

FIG. 10. Average DSCs of 24 patients at each treatment day (1–9). Dotted lines show the trends of segmentation performance.

85.45% and 13.4% of the three patients reported in Ref. 1. It is clear that our results indicate a significant improvement of the accuracy of segmentation, although the respective results were obtained with different datasets. Also, like the method in Ref. 3, the method in Ref. 1 is also based on intrapatient models. In their experiment, at least 13 patient-specific samples were used to train the intrapatient models for a test patient. This method will face the same problem as that in Ref. 3 when applied to clinical radiotherapy.

Based on the above analysis, we can draw a conclusion that the proposed method produces a competitive accurate result for prostate segmentation. More importantly, unlike methods requiring clinically infeasible numbers of patient-specific samples for intrapatient shape and appearance modeling, our method is more suitable for clinical application. Even in the situation where only one plan image is available for training, our method is capable of producing a reasonable segmentation. Furthermore, as more sequent treatment images are acquired, the models are updated dynamically and more accurate results can be obtained, as reported in Sec. III B 2. This point is very important for designing and evaluating a segmentation method for clinical application.

IV. CONCLUSION

We have presented a new deformable model for segmenting the prostate in serial CT images by using both population- and patient-specific statistics. The patient-specific statistics are learned online and incrementally from the segmentation results of previous treatment images of the same patient. In particular, for initial treatment images, the population-based shape statistics plays the primary role for statistically constraining the deformable surface. As more and more segmentation results are obtained, the patient-specific statistics start to constrain the segmentation and gradually take the major role for statistical constraining. In order to improve the robustness and accuracy of the segmentation, the gradient-PDF combination feature is also used to dynamically characterize the image information around each surface point. Experimental results show that the proposed method is robust and accurate and is suitable for clinical application.

In the future, we will speed up our segmentation algorithm using GPU. In our method, the most time-consuming procedure is the surface deformation step, which includes an iterative process. In each iteration, for each landmark on the surface, a local search is needed to find an optimal location to minimize the cost function. Fortunately, because the local search step (applied to each landmark) is almost independent, it is straightforward to align one computation thread of the GPU to one landmark for its corresponding calculation so that computation of hundreds of landmarks can be run in parallel. In this way, the overall computation time of our method could be reduced significantly.

ACKNOWLEDGMENTS

This research was supported by the grants from the National Basic Research Program of China (973 Program) (Grant No. 2010CB732505), National Natural Science Funds of China (Grant No. 30800254, No. 30730036, and No. 30900380), Natural Science Funds of Guangdong Province (No. 9151051501000026) and National Institutes of Health under Grant NIH grant R01CA140413.

[a]Electronic mail: qianjinfeng08@gmail.com
[b]Electronic mail: mark_foskey@unc.edu
[c]Author to whom correspondence should be addressed. Electronic mail: wufanchen@gmail.com
[d]Electronic mail: dgshen@med.unc.edu

[1]D. Freedman, R. J. Radke, T. Zhang, Y. Jeong, D. M. Lovelock, and G. T. Chen, "Model-based segmentation of medical imagery by matching distributions," IEEE Trans. Med. Imaging **24**, 281–292 (2005).
[2]Y. Jeong and R. J. Radke, "Modeling inter- and intra-patient anatomical variation using a bilinear model," presented at IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA, 2006), 2006 (unpublished).
[3]J. V. Stough, R. E. Broadhurst, S. M. Pizer, and E. L. Chaney, "Regional appearance in deformable model segmentation," Inf. Process. Med. Imaging **20**, 532–543 (2007).
[4]S. M. Pizer, R. E. Broadhurst, Y. Jeong, Q. Han, R. Saboo, J. V. Stough, G. Tracton, and E. Chaney, "Intra-patient anatomic statistical models for adaptive radiotherapy," presented at MICCAI Workshop from Statistical Atlases to Personalized Models: Understanding Complex Diseases in Populations and Individuals, 2006 (unpublished).
[5]M. J. Costa, H. Delingette, S. Novellas, and N. Ayache, "Automatic segmentation of bladder and prostate using coupled 3D deformable models," presented at International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI '07), 2007 (unpublished).
[6]N. Makni, P. Puech, R. Lopes, A. S. Dewalle, O. Colot, and N. Betrouni, "Combining a deformable model and a probabilistic framework for an automatic 3D segmentation of prostate on MRI," Int J Comput Assist Radiol Surg **4**, 1861 (2009).
[7]D. Flores-Tapia, G. Thomas, N. Venugopal, B. McCurdy, and S. Pistorius, "Semi automatic MRI prostate segmentation based on wavelet multiscale products," Conference Proceedings IEEE Engineering in Medicine and Biology Society, 2008, Vol. 2008, pp. 3020–3 (unpublished).
[8]D. Pasquier, T. Lacornerie, M. Vermandel, J. Rousseau, E. Lartigau, and N. Betrouni, "Automatic segmentation of pelvic structures from magnetic resonance images for prostate cancer radiotherapy," Int. J. Radiat. Oncol., Biol., Phys. **68**, 592–600 (2007).
[9]F. A. Cosio, "Automatic initialization of an active shape model of the prostate," Med. Image Anal. **12**, 469–483 (2008).
[10]A. C. Hodge, A. Fenster, D. B. Downey, and H. M. Ladak, "Prostate boundary segmentation from ultrasound images using 2D active shape models: Optimisation and extension to 3D," Comput. Methods Programs Biomed. **84**, 99–113 (2006).
[11]Y. Zhan and D. Shen, "Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method," IEEE Trans. Med. Imaging **25**, 256–272 (2006).
[12]D. Shen, Y. Zhan, and C. Davatzikos, "Segmentation of prostate boundaries from ultrasound images using statistical shape model," IEEE Trans. Med. Imaging **22**, 539–551 (2003).
[13]A. Ghanei, H. Soltanian-Zadeh, A. Ratkewicz, and F. F. Yin, "A three-dimensional deformable model for segmentation of human prostate from ultrasound images," Med. Phys. **28**, 2147–2153 (2001).
[14]T. Heimann and H. P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," Med. Image Anal. **13**, 543–563 (2009).
[15]T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," IEEE Trans. Pattern Anal. Mach. Intell. **23**, 681–685 (2001).
[16]S. M. Pizer, P. T. Fletcher, S. Joshi, A. G. Gash, J. Stough, A. Thall, G. Tracton, and E. L. Chaney, "A method and software for segmentation of anatomic object ensembles by deformable m-reps," Med. Phys. **32**, 1335–1345 (2005).
[17]P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi, "Principal geodesic analysis for the study of nonlinear statistics of shape," IEEE Trans. Med. Imaging **23**, 995–1005 (2004).
[18]B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "Active shape model segmentation with optimal features," IEEE Trans. Med. Imaging **21**, 924–933 (2002).

[19]Q. Feng, M. Foskey, S. Tang, W. Chen, and D. Shen, "Segmenting CT prostate images using population and patient-specific statistics for radiotherapy," presented at IEEE International Symposium on Biomedical Imaging (ISBI'90), 2009 (unpublished).
[20]B. C. Munsell, P. Dalal, and S. Wang, "Evaluating shape correspondence for statistical shape analysis: A benchmark study," IEEE Trans. Pattern Anal. Mach. Intell. **30**, 2023–2039 (2008).
[21]R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," IEEE Trans. Med. Imaging **21**, 525–537 (2002).
[22]A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," IEEE Trans. Med. Imaging **21**, 1151–1166 (2002).
[23]T. Heimann, I. Wolf, and H. P. Meinzer, "Automatic generation of 3D statistical shape models with optimal landmark distributions," Methods Inf. Med. **46**, 275–281 (2007).
[24]G. Taubin, "Curve and surface smoothing without shrinkage," presented at the Fifth International Conference on Computer Vision (ICCV '95), 1995 (unpublished).
[25]Y. Rathi, O. Michailovich, J. Malcolm, and A. Tannenbaum, "Seeing the unseen: Segmenting with distributions," presented at International Conference on Signal and Image Processing, 2006 (unpublished).
[26]T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," IEEE Trans. Commun. Technol. **15**, 52–60 (1967).
[27]R. E. Broadhurst, J. V. Stough, S. M. Pizer, and E. L. Chaney, "A statistical appearance model based on intensity quantile histograms," presented at IEEE International Symposium on Biomedical Imaging (ISBI '06), 2006 (unpublished).
[28]B. C. Davis, M. Foskey, J. Rosenman, L. Goyal, S. Chang, and S. Joshi, "Automatic segmentation of intra-treatment CT images for adaptive radiation therapy of the prostate," presented at International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI '05), 2005 (unpublished).
[29]L. R. Dice, "Measures of the amount of ecologic association between species," Ecology **26**, 297–302 (1945).
[30]K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells III, F. A. Jolesz, and R. Kikinis, "Statistical validation of image segmentation quality based on a spatial overlap index," Acad. Radiol. **11**, 178–189 (2004).