Online updating of context-aware landmark detectors for prostate localization in daily treatment CT images

Xiubin Dai
College of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210015, China and IDEA Lab, Department of Radiology and BRIC, University of North Carolina at Chapel Hill, 130 Mason Farm Road, Chapel Hill, North Carolina 27510

Yaozong Gao
IDEA Lab, Department of Radiology and BRIC, University of North Carolina at Chapel Hill, 130 Mason Farm Road, Chapel Hill, North Carolina 27510

Dinggang Shen[a)]
IDEA Lab, Department of Radiology and BRIC, University of North Carolina at Chapel Hill, 130 Mason Farm Road, Chapel Hill, North Carolina 27510 and Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea

(Received 18 July 2014; revised 22 February 2015; accepted for publication 20 March 2015; published 24 April 2015)

Purpose: In image guided radiation therapy, it is crucial to fast and accurately localize the prostate in the daily treatment images. To this end, the authors propose an online update scheme for landmark-guided prostate segmentation, which can fully exploit valuable patient-specific information contained in the previous treatment images and can achieve improved performance in landmark detection and prostate segmentation.
Methods: To localize the prostate in the daily treatment images, the authors first automatically detect six anatomical landmarks on the prostate boundary by adopting a context-aware landmark detection method. Specifically, in this method, a two-layer regression forest is trained as a detector for each target landmark. Once all the newly detected landmarks from new treatment images are reviewed or adjusted (if necessary) by clinicians, they are further included into the training pool as *new patient-specific information* to update all the two-layer regression forests for the next treatment day. As more and more treatment images of the current patient are acquired, the two-layer regression forests can be continually updated by incorporating the patient-specific information into the training procedure. After all target landmarks are detected, a multiatlas random sample consensus (multiatlas RANSAC) method is used to segment the *entire* prostate by fusing multiple previously segmented prostates of the current patient after they are aligned to the current treatment image. Subsequently, the segmented prostate of the current treatment image is again reviewed (or even adjusted if needed) by clinicians before including it as a new shape example into the prostate shape dataset for helping localize the entire prostate in the next treatment image.
Results: The experimental results on 330 images of 24 patients show the effectiveness of the authors' proposed online update scheme in improving the accuracies of both landmark detection and prostate segmentation. Besides, compared to the other state-of-the-art prostate segmentation methods, the authors' method achieves the best performance.
Conclusions: By appropriate use of valuable patient-specific information contained in the previous treatment images, the authors' proposed online update scheme can obtain satisfactory results for both landmark detection and prostate segmentation. [http://dx.doi.org/10.1118/1.4918755]

Key words: online update, prostate segmentation, context-aware landmark detection, regression forest

1. INTRODUCTION

Prostate cancer is one of the most common cancers for American males.[1] As a recently developed technology for prostate cancer radiation treatment, image guided radiation therapy (IGRT) provides an effective tool for cancer therapy, which often consists of a planning stage and a treatment stage.[2–4] *In the planning stage*, a planning CT scan is acquired from the patient, in which the prostate is manually delineated by a clinician for designing a patient-specific treatment plan. *In the treatment stage*, at each treatment day right before the radiation therapy, a treatment image will be obtained for localizing the prostate. Then, the treatment plan designed for the planning image can be transformed and adapted to the treatment image. Since the radiation dose should be precisely delivered to the current position of the tumor(s) and also minimized to the neighboring healthy tissues, it is important to accurately and quickly localize the prostate in the daily treatment images when the patient is still lying on the table.

However, as investigated in the previous works,[5–7] there are mainly three challenges making prostate localization in the treatment CT images quite challenging: (1) low contrast of treatment CT images; (2) dramatic appearance change of treatment images across different treatment days; and (3) unpredicted daily prostate motions.

To cope with the aforementioned challenges, numerous works have been carried out. One of the most popular prostate segmentation methods is the *deformable model based method*.[8–12] For example, Chen *et al.*[9] constructed a deformable model embedded with anatomical constraints from bones to jointly segment the prostate and rectum. Feng *et al.*[10] combined the gradient profile features and region-based features to guide deformable segmentation. Nevertheless, high dependence on good initialization of the deformable model, which is difficult to achieve in the CT prostate image, makes deformable models that have limited performance.

Another kind of method is the *deformable registration based method*.[13–21] For example, Liao and Shen[14] proposed a feature-guided deformable registration method using patient-specific information. To eliminate bowel gas before registration, Foskey *et al.*[13] presented a deflation method. Although deformable registration based methods are more robust than the deformable model based methods, it often takes more time for nonrigid registration and prostate localization, which is generally beyond the limit of computational time allowed in the clinical practice.

Classification based methods are also widely used in prostate segmentation.[7,22–24] For instance, Gao *et al.*[22] proposed a method to train the sparse representation classifier and then use multiatlas labeling to segment the prostate. In Li's work,[23] the auto-context model[25] was employed to iteratively refine the classification response map, and then the level-set method was used to segment the prostate based on the final classification response map. Shi *et al.*[24] proposed a spatially constrained transductive LASSO to produce a 3D prostate-likelihood map for guiding the segmentation. In spite of high accuracy in CT prostate segmentation, the classification based methods require a sufficient number of manually segmented patient-specific images (i.e., at least three images) for training. Thus, these methods cannot be used to segment the initial treatment images. Additionally, longer computational time is often needed by these classification based methods because of the complexity of the techniques used (e.g., sparse representation).

Recently, *learning-based landmark detection methods* have been widely used for many medical image analysis tasks.[6,26–38] For example, Criminisi *et al.*[27] built landmark detectors upon randomized decision forests for detection and localization of anatomical structures in the CT volumes. Zhan *et al.*[28] used a set of extended Haar-wavelet features and also the Adaboost learning method to train detectors for MR knee landmark detection. In the Criminisi's work,[38] they learned a regression model to detect anatomical structures in the CT images. Unlike most of the previous prostate segmentation methods, which need to manually annotate the entire prostate on all training images, the learning-based landmark detection methods require only the manual annotations of several anatomical landmarks. Obviously, this will significantly reduce the labors required for manual annotations. Thus, it is appealing to apply learning-based landmark detection methods for CT prostate localization. However, (1) the anatomical structures across different subjects, or across different treatment days of the same subject, might be quite different, as can be observed from Fig. 1; (2) image appearances of landmarks are often not same because of low image contrast (see Fig. 1). These two difficulties make automatic landmark detection for CT prostate localization quite challenging.

To address these difficulties, Gao *et al.*[39] recently developed a context-aware landmark detection method, by employing a two-layer regression forest to train a landmark detector for each target anatomical landmark in the CT prostate image. By utilizing both the intralandmark and the interlandmark distance-based context features, this two-layer regression model can be trained to have spatial consistency both (1) within the displacement field of each target landmark and (2) across displacement fields of all different target landmarks. Note that, in this paper, each component of the *displacement field* is a displacement vector, which denotes the 3D displacement of each voxel in each image toward the corresponding target landmark on the same image. Since other landmark detection methods[26–38] cannot simultaneously fulfill these two types of spatial consistency for the estimated displacement fields, which are important for accurate landmark detection, their performance might be limited.

However, Gao's method[39] was specifically designed for prostate localization in the planning images, instead of the treatment images. For detecting landmarks in the treatment images, this method would neglect the abundant prostate appear-

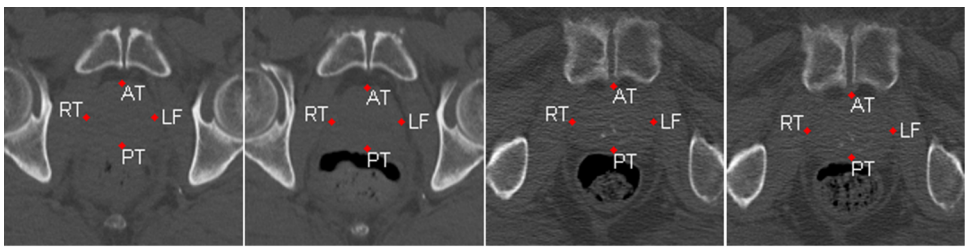

FIG. 1. Different anatomical appearances across different treatment days of two patients. (a) and (b) are the same slices obtained from different treatment images of patient A, and (c) and (d) are the same slices obtained from different treatment images of patient B. The red dots represent four selected anatomical landmarks, which are the right lateral point (RT), left lateral point (LF), posterior point (PT), and anterior point (AT), respectively.

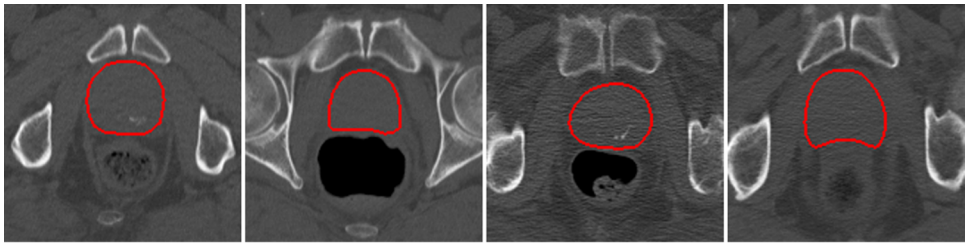

FIG. 2. The corresponding slices from four different patients [(a)–(d)] show quite different prostate appearances and shapes. Here, red contours denote the prostate shapes.

ance characteristics and shape information of each patient, which are embedded in the previous treatment images of the same patient. Besides, considering large differences of both prostate appearances and shapes among different patients (see Fig. 2), the performance of Gao's method, which uses only the planning images of different patients for training, could not be effective for prostate localization in the treatment images.

On the other hand, Gao *et al.*[6] also proposed a novel method, namely, the incremental learning with selective memory (ILSM), which blended the appearance statistics from both large population data and limited patient-specific data for detecting seven prostate anatomical landmarks and also localizing the prostate. Since ILSM fully exploits the appearance information embedded in the previous treatment images, it shows good performance in prostate localization of the treatment images. However, since ILSM detects landmarks independently without considering interlandmark spatial relationships during the landmark detection procedure, its accuracy could be further improved.

Motivated by the above-mentioned methods,[6,39] we propose an online-updated landmark-guided prostate localization method, by extending Gao's context-aware landmark detection method with an online update scheme. Specifically, it first uses an online-updated context-aware landmark detection method to train two-layer regression forests as detectors for automatically detecting six anatomical landmarks on the prostate boundary. Second, as soon as all six target landmarks are detected, a multiatlas RANSAC method is presented to localize the prostate by fusing multiple patient-specific prostate shapes that have been aligned onto the new treatment image (of the same patient).

With the help of the proposed online update scheme, the two-layer regression forests can be continually updated by incorporating the new appearance characteristics embedded in the newly acquired treatment image during the therapy. In this manner, the abundant patient-specific information (such as appearance and anatomical features of the prostates) can be better utilized to reduce the effect of large interpatient variation in training the initial landmark detectors. As a result, the online-updated two-layer regression forests can better adapt to the current treatment image than the conventional context-aware landmark detection method. In addition, the patient-specific prostate shape information used in our multiatlas RANSAC method can also be updated by the proposed online update scheme. Since our multiatlas RANSAC method takes advantage of the patient-specific prostate shape information contained in both the planning and previous treatment images, it can effectively lessen the performance degradation caused by the daily shape variations in the treatment images of the same patient and, thus, achieves more accurate results for prostate localization.

Compared to the ILSM method, our method will detect all the landmarks jointly by combining the context features for each target landmark with the context features for all other landmarks. That is, our proposed method can fully consider both the intralandmark and the interlandmark spatial relationships when detecting all the target landmarks. Thus, it can capture more intralandmark and interlandmark spatial relationships than the ILSM method, hence producing more accurate landmark detection than independent landmark detection.

The organization of this paper is as follows. Section 2 presents the proposed method based on the two-layer regression

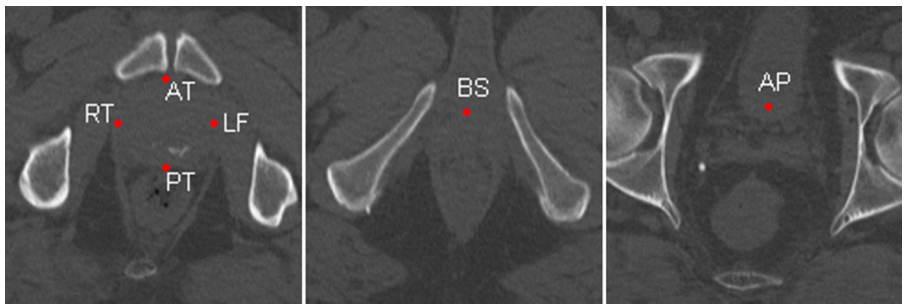

FIG. 3. Six prostate anatomical landmarks used in our study. In this paper, we choose four extreme points in the middle slice of the prostate [i.e., the leftmost (LF), the rightmost (RT), the most anterior (AT), and the most posterior points (PT)], one center point in the apex slice (AP) and one center point in the base slice (BS) as six anatomical landmarks of the prostate.

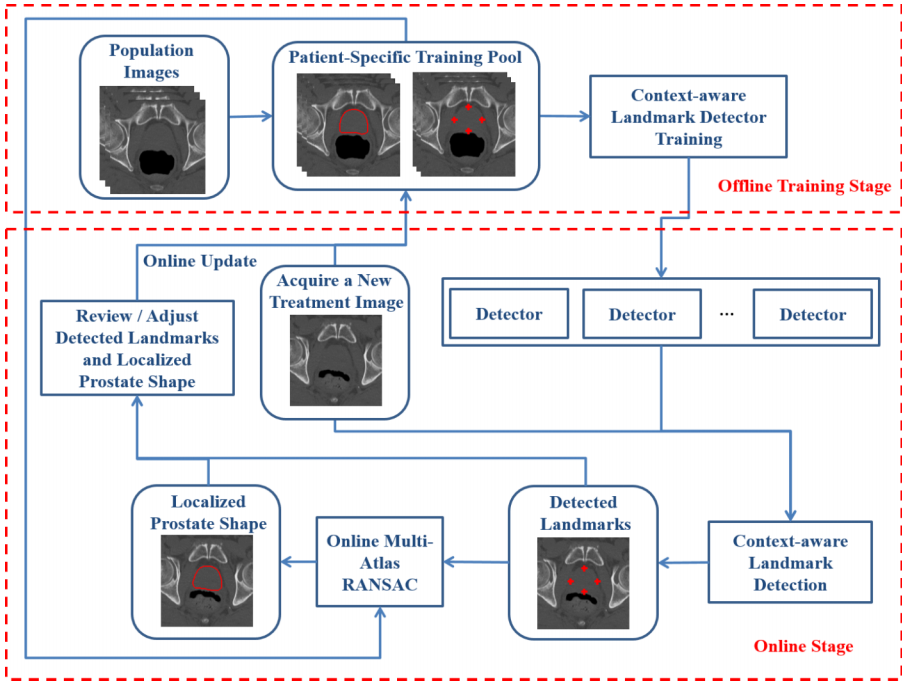

FIG. 4. The flowchart of the proposed method.

forest. The performance of our method is evaluated in Sec. 3. Finally, Sec. 4 concludes the paper.

2. METHODS

To localize prostate in the daily treatment images, we present an online-updated landmark-guided prostate localization method. First, the context-aware landmark detection method[39] is used to train two-layer regression forests to detect six anatomical landmarks, which are displayed in Fig. 3. With these newly detected landmarks, a new multiatlas RANSAC method is developed to localize the entire prostate in the new treatment image. Finally, the newly detected landmarks and localized prostate shapes can be reviewed or even adjusted by a clinician and then added as new patient-specific data to the pool of the planning image and previous treatment images of the same patient for updating both the two-layer regression forests and prostate shape information. The updated two-layer regression forests and prostate shape information can be used for landmark detection and prostate localization in the next treatment day. Figure 4 shows the entire procedure of our method. In Subsection 2.A, our method will be detailed.

2.A. The context-aware landmark detection

Although the context-aware landmark detection method also uses regression forests[40] as landmark detectors, it trains the two-layer regression forests, instead of only one-layer regression forests, as the traditional landmark detection methods,[37,38] by using both image appearance features and the two types of *high-level* distance-based context features (i.e., intralandmark and interlandmark context features). In this way, both the spatial consistency within the displacement field of each target landmark and the spatial consistency across the displacement fields of all different target landmarks can be fulfilled. In the following paragraphs, we will elaborate the context-aware landmark detection method,[39] which is composed of two main steps: (1) training of the context-aware landmark detectors and (2) landmark detection.

2.A.1. The training of the context-aware landmark detectors

In the context-aware landmark detection method, we train a two-layer regression forest as the detector for each target landmark. To do that, we first compute image appearance features of each sample point by using patch-based Haar-like features, which are defined as

$$H(P_I) = \sum_{h=1}^{Z} p_h \sum_{\|\mathbf{x}-a_h\|_\infty \le s_h} P_I(\mathbf{x}), \quad (1)$$

where P_I is a local patch in the image I; Z is the number of 3D cubic functions used in the Haar-like feature; $p_h \in \{-1,1\}$, $a_h \in R^3$, and s_h are the polarity, position, and scale of the hth 3D cubic function, respectively. By randomly sampling the number of 3D cubic functions and parameters (e.g., p_h, a_h, and s_h) of each cubic function in Eq. (1), we can generate a set of random patch-based Haar-like features to represent each sample point. Random Haar-like features capture both average intensity and intensity difference information within the patch. It should be noted that only informative Haar-like features will be finally selected during the regression forest training.

With the random patch-based Haar-like features and also the displacements from each sample point to the target landmarks, we can train the first-layer regression forest for each target landmark to learn the nonlinear mapping from the local appearance of an image voxel to its 3D displacement vector toward the corresponding target landmark. After that, these

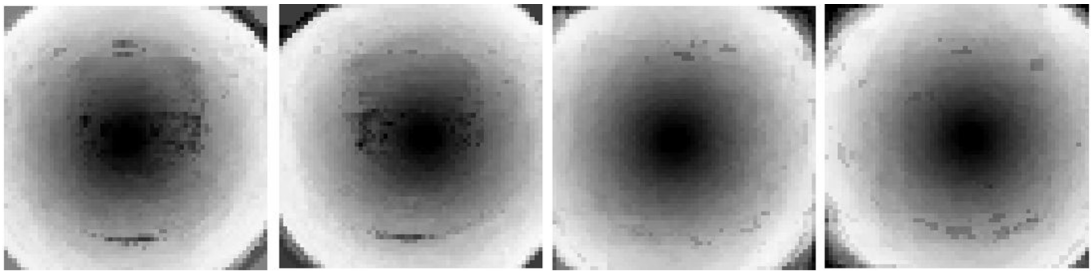

FIG. 5. Examples of the first-layer and second-layer distance maps for the two different landmarks. (a) and (b) are the first-layer distance maps for landmarks RT and LF, respectively; (c) and (d) are the second-layer distance maps for landmarks RT and LF, respectively.

trained first-layer regression forests are exerted on each training image to estimate a *first-layer displacement field* for every target landmark. By taking $L2$ norm on each displacement vector in the first-layer displacement fields, we can convert the first-layer displacement fields into the *first-layer distance maps* [Figs. 5(a) and 5(b)]. The $L2$ norm of the displacement vector $\mathbf{w} = [w_1, w_2, w_3]$ is denoted as $\sqrt{w_1^2 + w_2^2 + w_3^2}$.

Given the first-layer distance maps of all the target landmarks, the context features of the sample points can be extracted from the first-layer distance map of each target landmark (namely, *intralandmark context features*) by computing the random patch-based Haar-like features as in Eq. (1). Besides, the context features of the sample points are also extracted from the first-layer distance maps of all other landmarks (namely, *interlandmark context features*). Afterward, using these two different kinds of context features, along with the image appearance features, as new input features, we train the second-layer regression forests accordingly. Then, we can obtain the *second-layer displacement field* for each target landmark in every training image by applying the respective trained second-layer regression forest. Again, by taking $L2$ norm on each displacement vector in all the second-layer displacement fields, we can obtain their respective *second-layer distance maps* as displayed in Figs. 5(c) and 5(d). It is clear that, in contrast to the initial distance maps (i.e., the first two maps in Fig. 5), the qualities of second-layer distance maps (i.e., the last two maps in Fig. 5) are obviously improved by using our two-layer regression models.

2.A.2. Landmark detection

After finishing the training of the two-layer regression forests, the target landmarks of a new treatment image can be detected in a similar manner as in the training procedure. Specifically, in the first step of landmark detection, the random patch-based Haar-like features in Eq. (1) are computed as the image appearance features for each voxel in the new treatment image. Taking these image appearance features as the input, we acquire the first-layer 3D displacement fields for all the target landmarks by exerting the trained first-layer regression forests on the new treatment image. Then, by computing the $L2$ norm, the first-layer 3D *displacement* fields can be converted into the first-layer *distance* maps of all different target landmarks, from which *intralandmark* and *interlandmark* context features can be extracted. Afterward, we combine the image appearance features with these two types of context features as the input features to the trained second-layer regression forests, which will estimate the second-layer 3D displacement fields for all the target landmarks in the new treatment image. Once the second-layer displacement field for each target landmark is obtained, a method called regression voting[39] is adopted to detect the location of the target landmark. Specifically, in regression voting, each voxel $\mathbf{y}$ in the new treatment image can cast one vote to the location $\mathbf{y} + \mathbf{d}_\mathbf{y}$, where $\mathbf{d}_\mathbf{y}$ is the estimated displacement vector of voxel $\mathbf{y}$ toward the target landmark. After computing the voting for every location in the new treatment image, the target landmark can be detected as the location which receives the most votes.

2.A.3. Multiresolution training and landmark detection

In order to save the computational time and also improve the accuracy of landmark detection, the training process and landmark detection described in Subsections 2.A.1 and 2.A.2 can be further implemented in a multiresolution way.

For the multiresolution landmark detection, the landmark position detected in the coarser resolution can be used as a good initialization for landmark detection in the next finer resolution. Specifically, if the position of a target landmark has been detected in the coarser resolution, the regression voting in the next finer resolution will be performed only within a local neighborhood centered at the position detected by the coarser resolution. In this way, in the next finer resolution, we only need to search for the candidate target landmark within a small neighborhood, instead of the whole image. Thus, the computational time of landmark detection can be reduced greatly. Moreover, this also means that the voxels, which are far away from the target landmark, will not participate in the regression voting. Since those far-away voxels are often not informative to detect the target landmark, excluding these far-away voxels in the regression voting could also improve the accuracy of landmark detection.

Under such a multiresolution detection scheme, one regression forest is trained for each resolution in the training stage. Specifically, in the coarsest resolution, all training samples are drawn from the entire image domain to ensure the robustness of the initial landmark detection. In the finer resolution, the training samples are only drawn near the previously estimated location of target landmark, thus, gradually improving the accuracy of landmark detection.

The framework of this multiresolution landmark detection is given below for one anatomical landmark.

Framework 1: The multiresolution landmark detection for one landmark.

Input:

- The trained first-layer regression forest F^1 and second-layer regression forest F^2 for the target landmark.
- The new treatmentimage.

Output: The detected target landmark L'.
For each image resolution, repeat the following steps:

(1) Down-sample the image $\tilde{I}$ and calculate the image appearance features by using random patch-based Haar-like features in Eq. (1);
(2) Use the trained regression forests F^1 and F^2 to obtain the second-layer 3D displacement field D^2;
(3) Based on D^2, perform regression voting to detect the location of the target landmark L':
   a. For the coarsest resolution, every voxel in $\tilde{I}$ will vote as described in Subsection 2.A.2;
   For other finer resolutions, only voxels in $\tilde{I}$, which are within a local neighborhood centered at the position detected in the previous coarser resolution, are chosen to vote as described in Subsection 2.A.2;
   b. The location of the target landmark L' is detected as the voxel which receives the most votes;
(4) If not reaching the finest resolution, go to Step (1) and L' will be used as the initialization for the next finer image resolution. Otherwise, stop.

2.B. Online update scheme

To improve the specificity of the landmark detectors and also the accuracy of prostate localization in the treatment images, we develop an online update scheme to update the two-layer regression forests by the previously acquired treatment images of the same patient. Besides, we also propose a multiatlas RANSAC method, which can online update the prostate shape information of the same patient. By updating both landmark detectors and prostate shape information, we can achieve accurate prostate localization in the treatment images.

2.B.1. Online-updated two-layer regression forest

Figure 4 shows the main procedure of training the two-layer regression forests with the online update scheme. In the beginning of radiation treatment, since only one planning image of the current patient is available, we use the planning images of other patients, namely, population images, as training data to train the two-layer regression forests, which will be utilized to detect the target landmarks in the first-day treatment image of the current patient. After finishing the radiotherapy at each treatment day, a clinician can review or even adjust those automatically detected landmarks. Then, the reviewed or adjusted landmarks can be added into the (patient-specific) training pool to retrain and update the two-layer regression forests. When the radiotherapy starts in the next treatment day, the newly updated two-layer regression forests will be used for landmark detection in the new treatment image. It is worthwhile to note that the above landmark review/adjustment and also the updating/retraining of the two-layer regression forests will be repeated in the subsequent treatment days until the completion of radiotherapy for the current patient.

As introduced above, our two-layer regression forests, which will be used to detect landmarks, are actually trained by both population training images and patient-specific training images. Specifically, for the treatment day M, we use 1 planning image and $M-2$ previous treatment images as patient-specific training images. As more subsequent treatment images and their corresponding landmarks are obtained and further added into the training pool during the treatment course, the patient-specific appearance statistics can be updated constantly and will take a larger and larger role in training the two-layer regression forests. Thus, the online-updated two-layer regression forests will more accurately capture intrapatient variability and then better adapt to the current patient. In this way, the performance of landmark detection for each of the treatment images would be improved.

2.B.2. Multiatlas RANSAC

Based on the detected landmarks, one can use the RANSAC algorithm[41] to remove the wrongly detected landmarks and then estimate the optimal rigid transformation to fit the prostate shape (segmented in the planning image) onto the detected landmarks in the new treatment image for prostate localization. This is referred as the single-atlas RANSAC method.

To achieve the optimal rigid transformation, RANSAC has to do the estimation iteratively. In each iteration, RANSAC first chooses a random subset of the landmarks. Then, based on the selected landmarks, we can get two landmark subsets, namely, subsets $\mathbf{\Omega}$ and $\mathbf{\Psi}$, which contain the positions of selected landmarks encoded in the shape atlas and detected in the new treatment image, respectively. Afterward, a rigid transformation T can be estimated that minimizes the sum of distances between the transformed landmarks of subset $\mathbf{\Omega}$ by T and their corresponding landmarks of subset $\mathbf{\Psi}$. For other landmarks, which are not previously included in subsets $\mathbf{\Omega}$ and $\mathbf{\Psi}$, their positions in the shape atlas will be transformed by T to compare with the corresponding detected positions in the new treatment image. If the distance is below a certain threshold for one landmark, which means that this unselected landmark agrees with the estimated transformation T, it will be regarded as inlier and, thus, included into both subset $\mathbf{\Omega}$ and $\mathbf{\Psi}$ to re-estimate T. This procedure is iterated until no further inlier is found. In different iterations, different random landmark subsets will be initially selected. Thus, a set of different estimated transformations T can be derived. Finally, we pick the transformation T that obtains the smallest sum of distances between the aligned inlier landmarks of subset $\mathbf{\Omega}$ and their corresponding detected landmarks of subset $\mathbf{\Psi}$.

Although the single-atlas RANSAC method can localize the prostate shape in the new treatment image, it merely uses the prostate shape of the planning image and does not fully utilize the patient-specific prostate shape information contained in all the previous treatment images. Accordingly, its accuracy of prostate localization could be affected by shape variations between the planning day and the current treatment day. Generally, the longer the interval between two random days, the greater the prostate shape variations exist in the images of these treatment days. Thus, we expect more serious performance degradation if the current treatment day is far away from the planning day.

To address the above issue, we propose a new multiatlas RANSAC algorithm. Given the newly detected landmarks of the current treatment image, our multiatlas RANSAC algorithm first uses the RANSAC algorithm[41] to get rid of the wrongly detected landmarks. Since the landmarks and prostate shapes of the planning and previous treatment images are known, the RANSAC algorithm is utilized again to estimate the optimal rigid transformations between the newly detected landmarks of the current treatment image and the landmarks in each of the planning and previous treatment images.

Based on the estimated rigid transformations, multiple prostate shapes from the planning image and previous treatment images of the current patient can be independently aligned onto the space of the current treatment image. In this paper, prostate shapes of the planning image and previous treatment images are used as prostate shape atlases. Afterward, majority voting is adopted to fuse the labels of aligned multiple patient-specific prostate shapes, which finally gives the prostate shape in the new treatment image. By using multiple patient-specific prostate shapes, the negative influence brought by daily shape variations between treatment images can be better reduced.

It is worth noting that the majority voting used in multiatlas RANSAC is different from the regression voting used in landmark detection (Subsection 2.A.2). In multiatlas RANSAC, multiple prostate shapes are aligned onto the current treatment image. Then, each voxel of the current treatment image would receive votes, indicating whether this voxel belongs to the prostate or background. The label of this voxel is finally determined by the label with the maximum votes. For example, if we have N previous prostate shapes, each voxel of the current treatment image will receive N votes, belonging to the prostate or background. Then, this voxel is labeled as prostate if it receives more than $N/2$ votes from prostate. Otherwise, it is labeled as background.

Again, the newly localized prostate shape will be reviewed or even adjusted by a clinician. Then, we can update the patient-specific prostate shape information by adding the reviewed or adjusted prostate shape of the new treatment image into the training pool of the current patient. For the next treatment day, the updated pool of the landmarks and prostate shape atlases will be used for prostate localization in the new treatment image. Since more and more prostate shapes of the current patient are localized and collected during the serial treatment days, the patient-specific prostate shape information will be successively updated.

3. EXPERIMENTAL RESULTS

Our experimental data were acquired at the North Carolina Cancer Hospital and consist of 330 images from 24 different patients. All the images were scanned from a Siemens Somatom CT scanner, and the treatment images were acquired in every treatment day when patients received radiation therapy. The typical dose was 2–3 cGy for the planning imaging and less than 2 cGy for the treatment imaging. The field of view (FOV) is 50 cm for the planning images and 40 cm for the treatment images. The resolution and the image size of the acquired images are $0.98\times0.98\times3$ mm^3 and $512\times512\times30 \sim 120$, respectively. The prostates and their corresponding six anatomical landmarks in all CT images have been manually segmented and annotated by an expert, which serve as the ground-truth.

The results presented in this paper were generated using a system of Windows 7 with a 2.4 GHz Intel CPU and 8G memory.

For prostate localization, the Dice ratio (Dice) and the average surface distance (ASD) are adopted to evaluate the performance of the proposed method. Dice ratio measures the overlap ratio between the automatically segmented prostate volume by the proposed method (V_c) and the manually segmented prostate volume (V_g). Its definition is given as $2\times(V_c\cap V_g)/(V_c+V_g)$. ASD measures the average surface distance between V_c and V_g. By using surface U and surface Q to represent the boundaries of V_c and V_g, respectively, ASD is mathematically defined as

$$\mathrm{ASD}(U,Q)=\frac{1}{2}\left\{\frac{1}{|U|}\sum\nolimits_{t_u\in U}\min\nolimits_{t_q\in Q}\mathrm{dist}(t_u,t_q)+\frac{1}{|Q|}\sum\nolimits_{t_q\in Q}\min\nolimits_{t_u\in U}\mathrm{dist}(t_u,t_q)\right\}, \quad (2)$$

where $\mathrm{dist}(t_u,t_q)$ denotes the distance between vertex t_u from surface U and vertex t_q from surface Q and $|U|$ is the number of vertices in surface U.

All the following experiments use the same parameters. As for random patch-based Haar-like features, Z is limited to $\{1, 2\}$ and s_h is limited to $\{3, 5\}$. The patch size is set to $30\times30\times30$ voxels. We do not limit a_h, which can have arbitrary values as long as the 3D cubic function stays inside the local patch. The parameters used in the multiatlas RANSAC method can be referred to the Ref. 6.

In the training of the landmark detectors for a specific patient, we choose the first two treatment images of other patients as the population training images. As described in Subsection 2.B.1, for the treatment day M, we use 1 planning image and $M-2$ previous treatment images as patient-specific training images. After reviewing or even adjusting all automatically detected landmarks of the Mth treatment image by the clinician, both the Mth treatment image and its landmarks will be added into the patient-specific training pool to update the landmark detectors for the next treatment day $M+1$. This updating process continues until the end of the radiation treatment course. In many cases, after several days of updating, the landmark detectors have already captured sufficient intrapatient appearance variation. Further updating with more patient-specific

training images may only have negligible improvement. In such a case, we will choose to stop the online updating procedure. In the following experiments, when we mention that the training uses *up to* M patient-specific training images (i.e., 1 planning and $M-1$ treatment images), we mean that the online updating process stops after treatment day $M-1$. For the later treatment days, we just use the previously trained landmark detectors for prostate localization. Note that the experimental results reported below are computed on all treatment images.

3.A. Two-layer regression forest vs online-updated two-layer regression forest

In this subsection, we conduct the experiments to validate whether incorporating the online update scheme into the two-layer regression forest would lead to a performance improvement in landmark detection of the treatment images. Since there are two different layers of regression forests, we will carry out two separate experiments to verify the contribution of the online update scheme to each layer, respectively. In the following experiments, we use *up to* six patient-specific training images. We will show later that it is sufficient to use *up to* six patient-specific training images to capture the intrapatient appearance variations.

The first experiment aims to prove that the online update scheme can increase the accuracy of landmark detection in the treatment images for the first-layer regression forests. To accomplish that, we train the first-layer regression forests *with* or *without* the online update scheme, respectively, for each patient. The mean distance errors of the detected landmarks, obtained from the first-layer regression forests (denoted by the blue bars) and the *online-updated* first-layer regression forests (denoted by red bars) for all patients, are compared in Fig. 6. In this paper, the mean distance error of a landmark denotes the mean Euclidean distance between the detected landmarks and their corresponding ground-truths for all treatment images of all patients. From Fig. 6, we can see that the mean distance errors of all six landmarks by the online-updated first-layer regression forests are lower than those by the first-layer regression forests *without* the online update scheme. This indicates the advantage of incorporating the online update scheme into the first-layer regression forests for more accurate landmark detection. A two sample t-test for all six landmarks is conducted, which shows that the p-value is 8.32×10^{-14}. This low p-value implies the statistical difference between first-layer regression forests and online-updated first-layer regression forests.

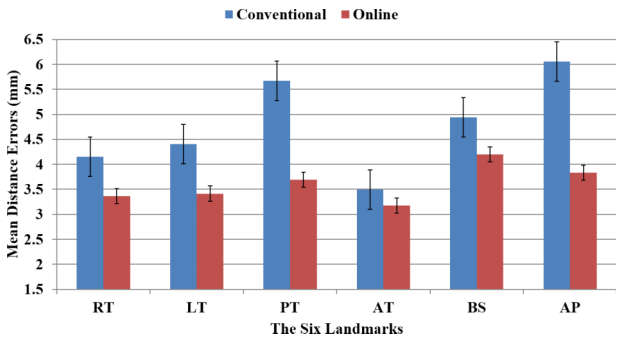

FIG. 6. Comparison between the first-layer regression forests and the *online-updated* first-layer regression forests using mean distance errors (mm) of the detected landmarks for all treatment images of 24 patients. Here, "conventional" and "online" denote the conventional first-layer regression forests and the *online-updated* first-layer regression forests, respectively. The error bars show the standard deviations.

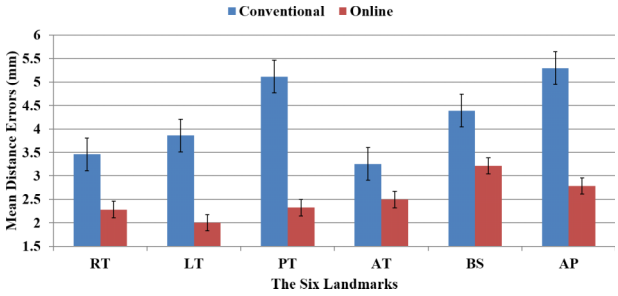

FIG. 7. Comparison of the second-layer regression forests without and with the online update scheme using the mean distance errors (mm) of six landmarks for all treatment images of 24 patients. Here, conventional and online denote the conventional second-layer regression forests and the *online-updated* second-layer regression forests, respectively. The error bars show the standard deviations.

The second experiment is designed to show that the online update scheme also improves the accuracy of landmark detection for the second-layer regression forests. Figure 7 compares the second-layer regression forests (denoted by the blue bars) and the *online-updated* second-layer regression forests (denoted by red bars) using the mean distance errors of the detected landmarks for all patients. Figure 7 shows that, for all six landmarks, the mean distance errors decline after using the online update scheme. These results prove that landmark detection by the second-layer regression forests can benefit from our online update scheme. The p-value between detection results of second-layer regression forests with and without online update scheme is 8.17×10^{-20}, indicating the statistical significance.

Thus, from above two experiments, we can conclude that combining the online update scheme with the two-layer regression forest can significantly improve the performance in landmark detection of serial treatment images.

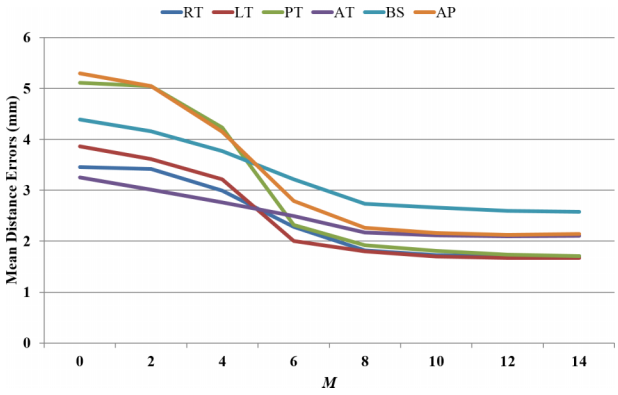

FIG. 8. The mean distance errors of six detected landmarks for 24 patients, by using *up to* M patient-specific training images. $M=0$ indicates the results obtained from the population-based landmark detectors.

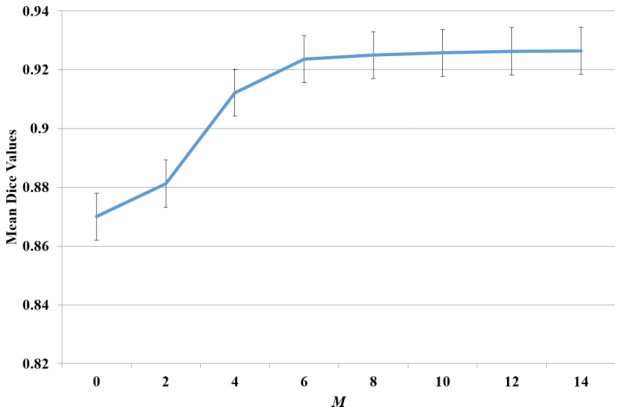

FIG. 9. The mean Dice curve of prostate localization for 24 patients, by using *up to* M patient-specific training images. $M = 0$ indicates the results obtained from the population-based landmark detectors. The error bars show the standard deviations.

3.B. The effect of the number of patient-specific training images

In this subsection, we investigate the performance of landmark detection and prostate localization with respect to the number of patient-specific training images used, and also how many patient-specific training images are sufficient for the online update process.

Figure 8 shows the mean distance errors of six target landmarks detected by our method *with respect to* different number M of patient-specific training images (i.e., M varies from 0 to 14). From Fig. 8, it can be observed that, as expected, the mean distance errors of all six landmarks keep decreasing as M increases from 0 to 14. However, when $M \geq 6$, the accuracy of landmark detection becomes saturated.

Figure 9 provides the mean Dice values with different M. Here, all the localization accuracies are obtained using up to five previous prostate shape atlases in multiatlas RANSAC. We can see that the mean Dice values ascend greatly with M changing from 0 to 6 but are saturated after that.

Both Figs. 8 and 9 show that the accuracies of landmark detection and prostate localization are low when no patient-specific images participate in the training process ($M = 0$). This confirms that the population images, which already exclude the images of the current patient under testing, cannot well capture image appearances of the current patient when used alone in training. However, as more and more treatment images of the current patient contribute to the training ($M = 2$–14), the results become better and better as shown in Figs. 8 and 9, indicating that the usage of more patient-specific treatment images as training images can raise the accuracies of both landmark detection and prostate localization.

Nevertheless, the improvement will be slight after using more than six treatment images of the same patient as training images. Further, adding more patient-specific images to the training pool will not bring significant improvement in both landmark detection and prostate localization. Thus, considering the trade-off between computational time and prostate localization performance, it is sufficient to use the population image and *up to* six patient-specific treatment images as training images.

3.C. Single-atlas RANSAC vs multiatlas RANSAC

We also compare the performances of the single-atlas RANSAC and the multiatlas RANSAC in localizing the prostates of 24 patients in Fig. 10. As suggested by the previous experiment, we use *up to* six patient-specific training images, which are sufficient for capturing the intrapatient appearance variations. In the multiatlas RANSAC, we use the prostate shapes from both the planning image and the latest five treatment images (before the current treatment image) in alignment. If no more than five patient-specific prostate shapes are available, all of the prostate shapes delineated in the planning day and previous treatment days will be used.

As depicted in Fig. 10, the mean Dice values obtained by the multiatlas RANSAC for all the patients are higher than those obtained by the single-atlas RANSAC. The two sample t-test is conducted for comparing the difference between single-atlas RANSAC and multiatlas RANSAC, which gives p-value 1.76×10^{-18} and shows the significant difference of two methods. This validates that our multiatlas RANSAC method is superior to the single-atlas RANSAC method in prostate localization.

3.D. The effect of the number of prostate shapes used in the multiatlas RANSAC

In this subsection, we investigate the performance of our multiatlas RANSAC in prostate localization *with respect to* the use of different number K of patient-specific prostate shapes (i.e., using *up to* K prostate shapes of the planning image and

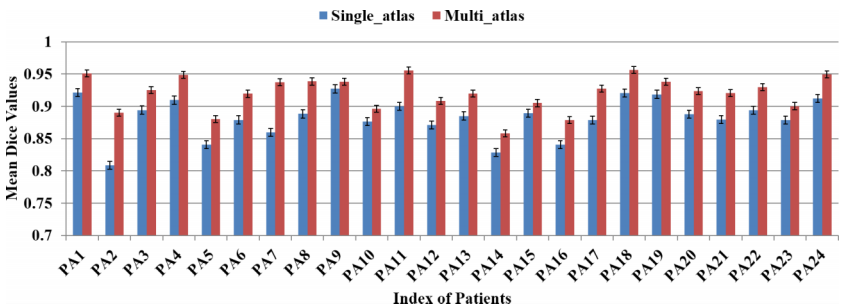

FIG. 10. Comparison between the single-atlas RANSAC and the multiatlas RANSAC on the mean Dice values of 24 patients. The error bars show the standard deviations.

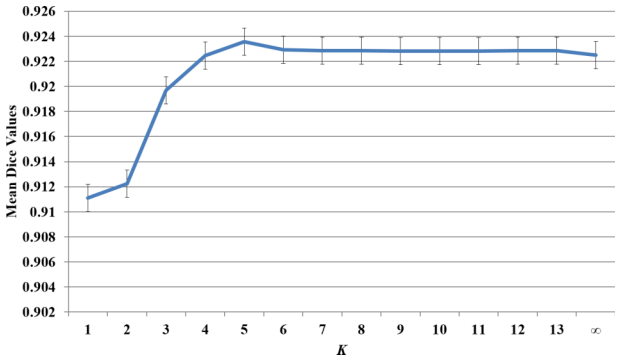

FIG. 11. The curve of mean Dice values for all patients when using *up to K* prostate shapes of the planning image and the latest treatment images for the multiatlas RANSAC. $K = \infty$ indicates that all previous prostate shapes will be used in the multiatlas RANSAC. The error bars show the standard deviations.

the latest treatment images). Figure 11 shows the respective mean Dice values of all patients. Here, $K = \infty$ indicates that all previous patient-specific prostate shapes are used. For the treatment day with less than K previous prostate shapes, all previous shapes are used as shape atlases in the online multiatlas RANSAC. Thus, the localized results when using *up to K* prostate shapes are composed of both the results obtained after the treatment day $K-1$ and the results of the first $K-1$ treatment days, in which there are less than K previous prostate shapes.

It can be seen from Fig. 11 that the mean Dice value with $K = 5$ is higher than others. That is, it is generally better to use *up to* five prostate shapes of the planning image and the latest treatment images for the multiatlas RANSAC. The reason using more prostate shapes leads to slightly worse results, which is possibly because of the daily shape variation between the prostates in the previous treatment images and the current treatment image. Generally, if a treatment day is far away from the current treatment day, the prostate shape variation between these two days could be large. Hence, compared to the use of only the latest prostate shapes, employing all the previous patient-specific prostate shapes (which could include more prostate shapes with large variations from the current treatment image) degrades the localization performance.

3.E. Comparison with other prostate segmentation methods on the same dataset

To further evaluate our method in prostate localization, we compare it with other state-of-the-art prostate localization methods, such as Feng,[10] Liao,[7,14] Li,[23] Shi,[24] Gao,[22] and ILSM,[6] by using the same dataset.

The comparison among different methods is listed in Table I. Table I shows that our method is better, with higher mean Dice and lower ASD measurements, than all other comparison methods, except for Shi's method. Although the mean Dice value of Shi's method is 1% higher than our method, it is worth noting that Shi's method is semiautomatic, which requires clinicians to manually specify the beginning and end slices. Besides, similar to most of the previous works (i.e., Liao,[7,14] Li,[23] and Gao[22]), Shi's method requires manual segmentation for at least three patient-specific images for initializing the algorithms. This means that one cannot adopt Shi's method to segment the initial treatment images. In contrast, by utilizing both the population training images and patient-specific training images, our method does not have such a limitation and, thus, can be applied to segment the prostate in any treatment day.

Notably, despite the fact that our method is directly related to Li's method,[23] there are three major differences between Li's method and our method. (1) Li's work learns an Adaboost classifier to voxel-wisely classify each voxel in a new treatment CT for obtaining a prostate likelihood map and then uses level set to segment the prostate from the likelihood map. In our method, we first localize six landmarks around the prostate by landmark detectors and then segment the prostate by aligning the previously segmented prostates onto the new treatment CT based on the automatically detected landmarks. (2) From Table I, it can be easily seen that our method achieves much better performance than Li's work. (3) Li's method cannot be used for localizing the prostate in the beginning treatment days, since it requires *at least* three manually segmented prostate shapes from the current patient for learning the classifier.

Table I also shows that although the computational time of our method is less than 1.5 min, it is still longer than that of ILSM.[6] However, in terms of segmentation accuracy, our method is much better than ILSM.[6]

For the purpose of testing whether our method achieves statistically significant improvement over the previous methods, we also list p-values in Table I. To compute the p-value between our method and any other method, the segmentation results of all subjects at all time points should be available. However, we can only get the segmentation results from Gao[22] and ILSM.[6] Thus, only the p-values with respect to these two methods (by comparison with our method) are reported in Table I. For other methods, we just provide their standard

TABLE I. Comparison between our method and other state-of-the-art prostate segmentation methods on the same dataset. Here, n/a means that the particular value was not reported in publications. The numbers following "±" is the standard deviations of Dice values and ASD values.

Method	Feng (Ref. 10)	Liao (Ref. 14)	Liao (Ref. 7)	Li (Ref. 23)	Shi (Ref. 24)	Gao (Ref. 22)	ILSM (Ref. 6)	Ours
Automaticity	Fully	Fully	Fully	Fully	Semi	Semi	Fully	Fully
Mean Dice	0.89 ± 0.05	0.90 ± n/a	0.91 ± n/a	0.91 ± n/a	0.94 ± 0.05	0.91 ± 0.06	0.88 ± 0.06	0.93 ± 0.06
Mean ASD (mm)	2.08 ± 0.79	1.08 ± n/a	0.97 ± n/a	1.40 ± n/a	n/a	1.24 ± 1.00	1.89 ± 0.98	0.97 ± 0.85
Speed (s)	96	228	156	180	n/a	600	4	78
P-value	n/a	n/a	n/a	n/a	n/a	4.51×10^{-4}	1.45×10^{-14}	n/a

TABLE II. Comparison between our method and some state-of-the-art prostate segmentation methods using different datasets. Here, n/a means that the particular value was not reported in publications.

Method	Davis (Ref. 18)	Foskey (Ref. 13)	Chen (Ref. 9)	Ours
Image no.	40	65	185	330
Automaticity	Fully	Fully	Fully	Fully
Mean Dice	0.82	0.84	n/a	0.93
Mean ASD (mm)	n/a	n/a	1.10	0.97
Speed (s)	n/a	750	60	78

deviations of segmentation results (if available in their papers) in Table I for reference. It can be concluded from Table I that our method is statistically ($p < 0.05$) better than Gao[22] and ILSM,[6] respectively.

3.F. Comparison with other prostate segmentation methods on different datasets

In this subsection, we compare the performance of our method with several other CT prostate segmentation methods on different datasets. Because these CT prostate segmentation methods are not using the same dataset as ours and also their source codes are not publicly available, it is difficult for us to directly compare our method with theirs. Consequently, Table II just lists the measurements reported in their publications for reference. One can observe from Table II that our method achieves higher segmentation accuracy and is evaluated on a larger dataset.

4. CONCLUSION AND DISCUSSION

In this paper, we propose an online-updated landmark-guided method for localizing the prostate in the treatment CT images. In our method, a two-layer regression forest is trained to detect each of the target landmarks in the new treatment image and further online-updated with the patient-specific information collected from the planning image and previous treatment images. After detecting all the target landmarks of the new treatment image, a multiatlas RANSAC method is adopted to localize the prostate by aligning the latest patient-specific prostate shapes onto the space of the new treatment image. By using the online update scheme in both landmark detection and prostate localization, the abundant patient-specific information can be continuously updated and further utilized to enhance the accuracies of both landmark detection and prostate localization. The experimental results have validated the effectiveness of the proposed method in landmark detection and prostate localization for the treatment CT images, as well as its better performance in comparison with the other state-of-the-art methods.

4.A. Qualitative comparison with other related works

The group led by Lei Dong applied an intensity-based deformable registration algorithm[21] to register the treatment CT with the planning CT and subsequently transformed the manual contouring result from the planning CT onto the treatment CT. They evaluated their method on one prostate patient with 24 repeat CT scans.[42] The average DSC (volume overlap index) and ASD (mean absolute surface-to-surface distance) of their method are 0.83 and 1.3 mm, respectively. The typical runtime for their method is 2 min/case. By comparison with their method, our method achieves an average of DSC 0.93 and ASD 0.97 mm for 24 patients with 330 images in total. Among 24 patients, the minimum DSC achieved by our method is 0.89, which is still better than their reported performance. With regard to runtime, the computational time required by our method (1 min 18 s) is shorter than their method.

In the work by Jan-Jakob Sonke's group, they developed a rigid intensity-based registration method to transform the contouring result from the planning CT to the treatment CT/cone-beam CT (CBCT). For treatment CT,[43] they achieved a 91% successful rate in 45 s. The average volume overlap ratio for those successful cases is 0.76 ± 0.05. For treatment CBCT,[12] they achieved a 65% successful rate with their regular registration method and a 83% successful rate when combined with the fixed apex registration method. It is worth noting that visual inspection was required in their algorithm to decide whether the fixed apex registration method should be used. The registration error obtained by their final method on those successful cases (9 patients 71 cases) is 0.1 ± 1.0, 0.5 ± 2.0, and -0.2 ± 1.7 mm along left–right, cranial–caudal, and anterior–posterior directions, respectively. The runtime is 1 min 30 s if only regular registration is used (with three start positions), and 2 min 30 s when fixed apex registration is additionally used (with three start positions). In contrast, our method is evaluated with a 100% successful rate in 24 patients of 330 cases. The average volume ratio is 0.91 ± 0.06, which is much higher than their method. The landmark-guided registration error by our method is 0.03 ± 0.64, 0.13 ± 1.94, -0.10 ± 1.81 mm along left–right, cranial–caudal, and anterior–posterior directions, respectively, which is also better.

The group of Steve Jiang recently implemented many popular deformable image registration methods, such as demon algorithm and its variants,[44–47] on the GPU to accelerate the registration for image guided radiation therapy.[48] By the acceleration of the GPU, registering images of size $256 \times 256 \times 100$ only needs 7–11 s. They also successfully applied the demons algorithm to register 4D CT images of different phases. Even though they did not apply their method for prostate localization, their method can certainly be borrowed to localize the prostate, as the above two registration methods did. Actually, similar works have been done by Foskey *et al.*,[13] who also applied an intensity-based deformable registration method to localize the prostate in treatment CT. For the registration-based methods, the main difficulty/challenge comes from the possible dramatic change of bowel gas in the rectum, which will negatively affect the registration algorithm.[12,18,23,43] Therefore, various strategies were developed in many works[12,13,18,43] to fill up the bowel gas before registration. However, when the change of bowel gas becomes large between the planning CT and the treatment CT, the registration algorithm would fail in most cases.[12,43] In our method, we localize the prostate

by automatically detecting several key landmarks around the prostate. Compared with registration methods, our method does not need to globally match the intensity appearance, thus, is free from the bowel gas problem. As shown in the above two paragraphs, our method achieves more accurate results of prostate localization than intensity-based registration methods.

4.B. Applicability to CBCT treatment images

Among various prostate segmentation methods, our method belongs to the machine learning method. In our method, we did not handcraft features that are specific to CT modality. Useful features are automatically learned by the random forest from the previously segmented treatment and planning CT images. Random forest as a general machine learning method has been applied to various image modalities. For example, Cheng *et al.*[26] used the random forest as a classification method for CBCT dental landmark detection and achieved good results. Criminisi *et al.*[38] reported the application of the random forest as a regression way for efficient CT anatomy detection. In Maglietta's work,[49] the random forest is employed for automatic hippocampal segmentation in MR images. Therefore, on the basis of the data-driven nature of a random forest, we are sure that our method can automatically identify useful features for landmark detection in CBCT prostate images if enough treatment CBCTs are provided. In the future, we will test our algorithm on CBCT treatment images.

4.C. Limitations for clinical use

Although the performance of our method in prostate localization is satisfactory, it still has some limitations for clinical use. One limitation is the computational time of the current method. Even though the computational time of our method is comparable with most of the other prostate localization algorithms, such as works from Dong,[21] Sonke,[12,43] Foskey,[13] and Chen,[26] there is still a room for further improvement. To further reduce the computational time, we consider implementing the random forest on the GPU,[50] similar to Steve Jiang's work. It will significantly reduce the time. Besides, now we detect each landmark sequentially in the same level. Since the detections of different landmarks in the same level are independent, given the displacement fields of the previous iteration, this parallelism can be further exploited to reduce the computational time. Another limitation is the requirement of at least one patient-specific image for training. Since there are no images available for the patient in the very beginning of treatment, we have to use population images from other patients as the training data. As we discussed in Subsection 3.B, population training images cannot well capture the appearance of the current patient. Therefore, localizing the prostate in the planning image without any patient-specific information would be a challenging task, which will be our future research direction.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. CA140413, the National Natural Science Foundation of China (Grants No. 31200747 and 61473190), the Natural Science Foundation of Jiangsu Province (Grant No. BK2012437), and the Starting Scientific Research Foundation of Nanjing University of Posts and Telecommunications for New Teachers (Grant No. NY211030).

[a]Author to whom correspondence should be addressed. Electronic mail: dgshen@med.unc.edu; Telephone: 1-919-966-4397.

[1]A. C. Society, "Cancer facts and figures 2012," J. Consum. Health Internet **16**, 366–367 (2012).
[2]D. Shen, Z. Lao, J. Zeng, W. Zhang, I. A. Sesterhenn, L. Sun, J. W. Moul, E. H. Herskovits, G. Fichtinger, and C. Davatzikos, "Optimized prostate biopsy via a statistical atlas of cancer spatial distribution," Med. Image Anal. **8**, 139–150 (2004).
[3]L. Xing, "Overview of image-guided radiation therapy," Med. Dosim. **31**, 91–112 (2006).
[4]L. Dawson and D. Jaffray, "Advances in image-guided radiation therapy," J. Clin. Oncol. **25**, 938–946 (2007).
[5]W. Liu, J. Qian, S. L. Hancock, L. Xing, and G. Luxton, "Clinical development of a failure detection-based online repositioning strategy for prostate IMRT—Experiments, simulation, and dosimetry study," Med. Phys. **37**, 5287–5297 (2010).
[6]Y. Gao, Y. Zhan, and D. Shen, "Incremental learning with selective memory (ILSM): Towards fast prostate localization for image guided radiotherapy," IEEE Trans. Med. Imaging **33**, 518–534 (2014).
[7]S. Liao, Y. Gao, J. Lian, and D. Shen, "Sparse patch-based label propagation for accurate prostate localization in CT images," IEEE Trans. Med. Imaging **32**, 419–434 (2013).
[8]M. Costa, H. Delingette, S. Novellas, and N. Ayache, "Automatic segmentation of bladder and prostate using coupled 3D deformable models," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, edited by N. Ayache, S. Ourselin, and A. Maeder (Springer, Berlin, Heidelberg, 2007), Vol. 4791, pp. 252–260.
[9]S. Chen, D. M. Lovelock, and R. J. Radke, "Segmenting the prostate and rectum in CT imagery using anatomical constraints," Med. Image Anal. **15**, 1–11 (2011).
[10]Q. Feng, M. Foskey, W. Chen, and D. Shen, "Segmenting CT prostate images using population and patient-specific statistics for radiotherapy," Med. Phys. **37**, 4121–4132 (2010).
[11]D. Freedman, R. J. Radke, Z. Tao, J. Yongwon, D. M. Lovelock, and G. T. Y. Chen, "Model-based segmentation of medical imagery by matching distributions," IEEE Trans. Med. Imaging **24**, 281–292 (2005).
[12]M. H. P. Smitsmans, J. de Bois, J.-J. Sonke, A. Betgen, L. J. Zijp, D. A. Jaffray, J. V. Lebesque, and M. van Herk, "Automatic prostate localization on cone-beam CT scans for high precision image-guided radiotherapy," Int. J. Radiat. Oncol. **63**, 975–984 (2005).
[13]M. Foskey, B. Davis, L. Goyal, S. Chang, E. Chaney, N. Strehl, S. Tomei, J. Rosenman, and S. Joshi, "Large deformation three-dimensional image registration in image-guided radiation therapy," Phys. Med. Biol. **50**, 5869–5892 (2005).
[14]S. Liao and D. Shen, "A feature-based learning framework for accurate prostate localization in CT images," IEEE Trans. Image Process. **21**, 3546–3559 (2012).
[15]G. Wu, F. Qi, and D. Shen, "Learning-based deformable registration of MR brain images," IEEE Trans. Med. Imaging **25**, 1145–1157 (2006).
[16]J. Yang, D. Shen, C. Davatzikos, and R. Verma, "Diffusion tensor image registration using tensor geometry and orientation features," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, edited by D. Metaxas, L. Axel, G. Fichtinger, and G. Székely (Springer, Berlin, Heidelberg, 2008), Vol. 5242, pp. 905–913.
[17]Y. Zhan, Y. Ou, M. Feldman, J. Tomaszeweski, C. Davatzikos, and D. Shen, "Registering histologic and MR images of prostate for image-based cancer detection," Acad. Radiol. **14**, 1367–1381 (2007).
[18]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," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, edited by J. Duncan and G. Gerig (Springer, Berlin, Heidelberg, 2005), Vol. 3749, pp. 442–450.
[19]L. Zhou, S. Liao, W. Li, and S. Dinggang, "Learning-based prostate localization for image guided radiation therapy," in *2011 IEEE International*

Symposium on Biomedical Imaging: From Nano to Macro (IEEE, Chicago, IL, 2011), pp. 2103–2106.
[20]S. Martin, V. Daanen, and J. Troccaz, "Atlas-based prostate segmentation using an hybrid registration," Int. J. Comput. Assisted Radiol. Surg. **3**, 485–492 (2008).
[21]H. Wang, L. Dong, M. F. Lii, A. L. Lee, R. de Crevoisier, R. Mohan, J. D. Cox, D. A. Kuban, and R. Cheung, "Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy," Int. J. Radiat. Oncol. **61**, 725–735 (2005).
[22]Y. Gao, S. Liao, and D. Shen, "Prostate segmentation by sparse representation based classification," Med. Phys. **39**, 6372–6387 (2012).
[23]W. Li, S. Liao, Q. Feng, W. Chen, and D. Shen, "Learning image context for segmentation of the prostate in CT-guided radiotherapy," Phys. Med. Biol. **57**, 1283–1308 (2012).
[24]Y. Shi, S. Liao, Y. Gao, D. Zhang, Y. Gao, and D. Shen, "Prostate segmentation in CT images via spatial-constrained transductive Lasso," in *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, Portland, OR, 2013), pp. 2227–2234.
[25]T. Zhuowen and B. Xiang, "Auto-context and its application to high-level vision tasks and 3D brain image segmentation," IEEE Trans. Pattern Anal. Mach. Intell. **32**, 1744–1757 (2010).
[26]C. Erkang, C. Jinwu, Y. Jie, D. Huiyang, W. Yi, V. Megalooikonomou, B. Gable, and L. Haibin, "Automatic Dent-landmark detection in 3-D CBCT dental volumes," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC* (IEEE, Boston, MA, 2011), pp. 6204–6207.
[27]A. Criminisi, J. Shotton, and S. Bucciarelli, "Decision forests with long-range spatial context for organ localization in CT volumes," in *MICCAI Workshop on Probabilistic Models for Medical Image Analysis* (Springer, London, UK, 2009).
[28]Y. Zhan, M. Dewan, M. Harder, A. Krishnan, and Z. X. Sean, "Robust automatic knee MR slice positioning through redundant and hierarchical anatomy detection," IEEE Trans. Med. Imaging **30**, 2087–2100 (2011).
[29]R. Bahbibi, "Image analysis using machine learning: Anatomical landmarks detection in fetal ultrasound images," in *2012 IEEE 36th Annual Computer Software and Applications Conference* (IEEE, Izmir, 2012), pp. 354–355.
[30]X. Zhao, X. Chai, Z. Niu, C. Heng, and S. Shan, "Context modeling for facial landmark detection based on Non-Adjacent Rectangle (NAR) Haar-like feature," Image Vision Comput. **30**, 136–146 (2012).
[31]J. Sivic, M. Everingham, and A. Zisserman, "'Who are you?'–Learning person specific classifiers from video," in *IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009* (IEEE, Miami, FL, 2009), pp. 1145–1152.
[32]H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," IEEE Trans. Med. Imaging **21**, 450–461 (2002).
[33]S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, and X. S. Zhou, "Towards robust and effective shape modeling: Sparse shape composition," Med. Image Anal. **16**, 265–277 (2012).
[34]T. Cootes, M. Ionita, C. Lindner, and P. Sauer, "Robust and accurate shape model fitting using random forest regression voting," in *Computer Vision–ECCV 2012*, edited by A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid (Springer, Berlin, Heidelberg, 2012), Vol. 7578, pp. 278–291.
[35]M. Valstar, B. Martinez, X. Binefa, and M. Pantic, "Facial point detection using boosted regression and graph models," in *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, San Francisco, CA, 2010), pp. 2729–2736.
[36]P. N. Belhumeur, D. W. Jacobs, D. Kriegman, and N. Kumar, "Localizing parts of faces using a consensus of exemplars," in *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, Providence, RI, 2011), pp. 545–552.
[37]M. Dantone, J. Gall, G. Fanelli, and L. Van Gool, "Real-time facial feature detection using conditional regression forests," in *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, Providence, RI, 2012), pp. 2578–2585.
[38]A. Criminisi, J. Shotton, D. Robertson, and E. Konukoglu, "Regression forests for efficient anatomy detection and localization in CT studies," in *Medical Computer Vision*, Recognition Techniques and Applications in Medical Imaging Vol. 6533, edited by B. Menze, G. Langs, Z. Tu, and A. Criminisi (Springer, Berlin, Heidelberg, 2011), pp. 106–117.
[39]Y. Gao, L. Wang, Y. Shao, and D. Shen, "Learning distance transform for boundary detection and deformable segmentation in CT prostate images," in *5th International Workshop on Machine Learning in Medical Imaging* (Springer, Boston, MA, 2014).
[40]L. Breiman, "Random forests," Mach. Learn. **45**, 5–32 (2001).
[41]M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Commun. ACM **24**, 381–395 (1981).
[42]H. Wang, A. S. Garden, L. Zhang, X. Wei, A. Ahamad, D. A. Kuban, R. Komaki, J. O'Daniel, Y. Zhang, R. Mohan, and L. Dong, "Performance evaluation of automatic anatomy segmentation algorithm on repeat or four-dimensional computed tomography images using deformable image registration method," Int. J. Radiat. Oncol. **72**, 210–219 (2008).
[43]M. H. P. Smitsmans, J. W. H. Wolthaus, X. Artignan, J. de Bois, D. A. Jaffray, J. V. Lebesque, and M. van Herk, "Automatic localization of the prostate for on-line or off-line image-guided radiotherapy," Int. J. Radiat. Oncol. **60**, 623–635 (2004).
[44]J. P. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," Med. Image Anal. **2**, 243–260 (1998).
[45]P. Rogelj and S. Kovačič, "Symmetric image registration," Med. Image Anal. **10**, 484–493 (2006).
[46]H. Wang, L. Dong, J. O'Daniel, R. Mohan, A. S. Garden, K. K. Ang, A. K. Deborah, B. Mark, J. Y. Chang, and R. Cheung, "Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy," Phys. Med. Biol. **50**, 2887–2905 (2005).
[47]D. Yang, H. Li, A. L. Daniel, O. D. Joseph, and I. E. Naqa, "A fast inverse consistent deformable image registration method based on symmetric optical flow computation," Phys. Med. Biol. **53**, 6143–6165 (2008).
[48]X. Gu, H. Pan, Y. Liang, R. Castillo, D. Yang, D. Choi, E. Castillo, A. Majumdar, T. Guerrero, and S. B. Jiang, "Implementation and evaluation of various demons deformable image registration algorithms on a GPU," Phys. Med. Biol. **55**, 207–219 (2010).
[49]R. Maglietta, N. Amoroso, S. Bruno, A. Chincarini, G. Frisoni, P. Inglese, S. Tangaro, A. Tateo, and R. Bellotti, "Random forest classification for hippocampal segmentation in 3D MR images," in *2013 12th International Conference on Machine Learning and Applications (ICMLA)* (IEEE, Miami, FL, 2013), pp. 264–267.
[50]B. Van Essen, C. Macaraeg, M. Gokhale, and R. Prenger, "Accelerating a random forest classifier: Multi-core, GP-GPU, or FPGA?," in *2012 IEEE 20th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)* (IEEE, Toronto, ON, 2012), pp. 232–239.