

Automatic multiorgan segmentation in CT images of the male pelvis using region-specific hierarchical appearance cluster models

Dengwang Li^{a)}

Shandong Province Key Laboratory of Medical Physics and Image Processing Technology, Institute of Biomedical Sciences, School of Physics and Electronics, Shandong Normal University, Jinan 250014, China and Medical Physics Division, Department of Radiation Oncology, Stanford University, Stanford, California 94305

Pengxiao Zang

Shandong Province Key Laboratory of Medical Physics and Image Processing Technology, Institute of Biomedical Sciences, School of Physics and Electronics, Shandong Normal University, Jinan 250014, China

Xiangfei Chai, Yi Cui, Ruijiang Li, and Lei Xing

Medical Physics Division, Department of Radiation Oncology, Stanford University, Stanford, California 94305

(Received 24 December 2015; revised 16 August 2016; accepted for publication 19 August 2016; published 13 September 2016)

Purpose: Accurate segmentation of pelvic organs in CT images is of great importance in external beam radiotherapy for prostate cancer. The aim of this study is to develop a novel method for automatic, multiorgan segmentation of the male pelvis.

Methods: The authors' segmentation method consists of several stages. First, a pretreatment includes parameterization, principal component analysis (PCA), and an established process of region-specific hierarchical appearance cluster (RSHAC) model which was executed on the training dataset. After the preprocessing, online automatic segmentation of new CT images is achieved by combining the RSHAC model with the PCA-based point distribution model. Fifty pelvic CT from eight prostate cancer patients were used as the training dataset. From another 20 prostate cancer patients, 210 CT images were used for independent validation of the segmentation method.

Results: In the training dataset, 15 PCA modes were needed to represent 95% of shape variations of pelvic organs. When tested on the validation dataset, the authors' segmentation method had an average Dice similarity coefficient and mean absolute distance of 0.751 and 0.371 cm, 0.783 and 0.303 cm, 0.573 and 0.604 cm for prostate, bladder, and rectum, respectively. The automated segmentation process took on average 5 min on a personal computer equipped with Core 2 Duo CPU of 2.8 GHz and 8 GB RAM.

Conclusions: The authors have developed an efficient and reliable method for automatic segmentation of multiple organs in the male pelvis. This method should be useful for treatment planning and adaptive replanning for prostate cancer radiotherapy. With this method, the physicist can improve the work efficiency and stability. © 2016 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4962468>]

Key words: automatic multiorgan segmentation, radiotherapy treatment planning, image segmentation, cancer treatment

1. INTRODUCTION

Prostate and bladder cancers are two common cancers in the male pelvis. In particular, prostate cancer is the most common cancer and the second leading cause of cancer death among men in the US.^{1,2} Radiotherapy is a major treatment modality for prostate cancer.³ The accuracy of the segmentation for pelvic organs plays a critical role in maximizing the efficacy of radiotherapy.

In current practice, the gross tumor volume and surrounding critical organs are manually delineated on planning CT images. However, such a manual segmentation process is both time-consuming and subjective, leading to large intraobserver and interobserver variations.⁴ Moreover, in adaptive radiotherapy, manual segmentation is very time consuming and instability on daily cone beam computed

tomography (CBCT). For example, in online adaptive schemes, Ahunbay *et al.* proposed an online adaptive replanning method for prostate radiotherapy.⁵ In light of these, an automatic or semiautomatic segmentation method of pelvic organs on CT image is needed to improve the efficiency and accuracy of the delineation process.

There are two major challenges for the automatic segmentation of pelvic organs. First, CT images have a low soft tissue contrast in the pelvis due to the small intensity differences between organs (see Fig. 1). Therefore, segmentation solely based on the intensity of the image is difficult to obtain a reliable result. In the past, several methods have been proposed to address this challenge for segmentation in the lung,^{6–12} heart,^{13–17} and liver.^{18–22}

The second challenge relates to the changing morphology of pelvic organs during the treatment course, due to normal

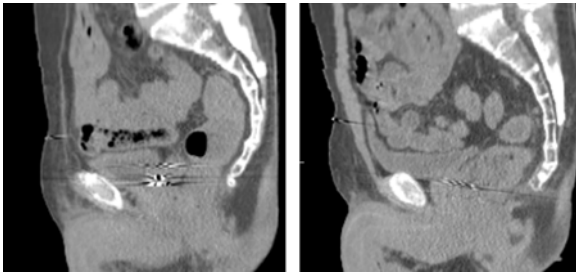


FIG. 1. Examples of the pelvic organs show low soft-tissue contrast on CT image.

physiological activities such as bowel filling and bowel gas, both on an interfractional basis and intrafractional basis. Several segmentation methods for the pelvic organs have been proposed to address the variability of the shape and appearance of the pelvic organs.^{23–26} The most popular category of these methods is the segmentation algorithm based on the deformation model.^{27–34} For example, Broadhurst *et al.* proposed a deformable model based on the histogram-statistics appearance model for prostate segmentation.³² Pizer *et al.* proposed a medial shape model to segment bladder, prostate, and rectum from CT images.²⁷ Methods using deformable models based on the statistical information extracted from both organ shape and image appearance have been studied in recent years and these methods showed good performance for pelvic organs. However, segmentation results using these methods depend on a good initial model. Another class of methods is based on the deformable registration.^{35–37} Liao and Shen *et al.* presented a method to guide the deformable registration of prostate CT with an evaluation function.³⁷ However, as mentioned earlier, the accuracy of registration-based methods can be unreliable because of the complex physiological activities in the pelvis.

In recent years, the shape model based on prior knowledge and machine learning has been widely used in automatic segmentation of CT images. This type of methods uses a training dataset as the prior knowledge. The training dataset is divided into two kinds: population-based and patient-specific. The method based on the patient-specific model often has higher segmentation accuracy because the patient-specific dataset consists of several CT images from the same patients, and there is a large shape variation between different patients (see Fig. 2). For example, Stough *et al.* proposed a segmentation method based on regional appearance within a deformable model.³⁸ Feng *et al.* proposed a segmentation method based on population and patient-specific statistics for radiotherapy.³⁰ However, the patient-specific training dataset requires at least four CT scans from the same patient, prohibiting its use for the first few daily CT scans. On the other hand, population-based method can be used for every CT image during treatment. Nevertheless, these methods have limited segmentation accuracy.

In this paper, we develop a novel automatic, multiorgan segmentation method based on CT images of prostate cancer patients using region-specific shape information. A statistical point distribution model using principal component analysis (PCA) is first built based on the training dataset to grossly

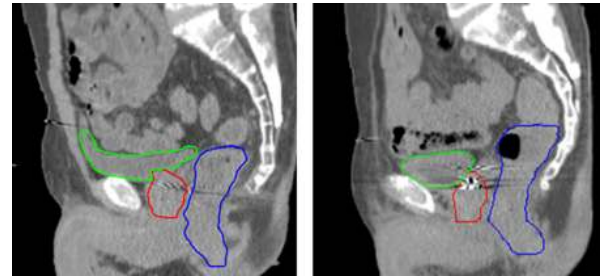


FIG. 2. Sagittal slices from two patients manually contoured by the same radiation oncologist. Prostate, bladder, and rectum were, respectively, contoured by red, green, and blue thick solid curves. The images show that the organ shapes and appearances are very different among patients. (See color online version.)

capture the major variation of pelvic organs. A novel region-specific agglomerative hierarchical cluster model is established to further improve the segmentation accuracy. The aim of this study is to present a novel segmentation method and to evaluate its accuracy on a large number of CT images.

2. METHODS AND MATERIALS

The method is divided into two main parts. The first part is the training process which needed to be performed before the automatic segmentation. And this training process is just needed once to calculate the point distribution model and the region-specific hierarchical appearance cluster (RSHAC) model. The second part is the automatic segmentation part which mainly includes a coarse and refined segmentation process. The two models obtained in the training process will be used to guide the segmentation in this part. The whole flow chart of this method was shown in Fig. 3.

2.A. Patient data for training and testing

The training dataset included a total of 50 planning CT scans with a *slice* size of 512×512 and 200–400 slices (average 256) at a voxel resolution of $1.0 \times 1.0 \times 1.5$ mm. The contours of the pelvic organs in the training dataset were manually delineated by an experienced medical physicist and reviewed by a radiation oncologist.

The CT images used for validation of the proposed segmentation method were from another twenty male patients with prostate cancer: twelve of them were in the early diagnosis of prostate cancer and the other eight patients were in the middle stage of the prostate cancer. The planning CT of these twenty prostate patients was used as the test data. The manual contour of these planning CTs will serve as references to test the segmentation accuracy.

2.B. Training process

In the training process, the shape model and appearance model of the multiple pelvic organs were calculated, respectively.

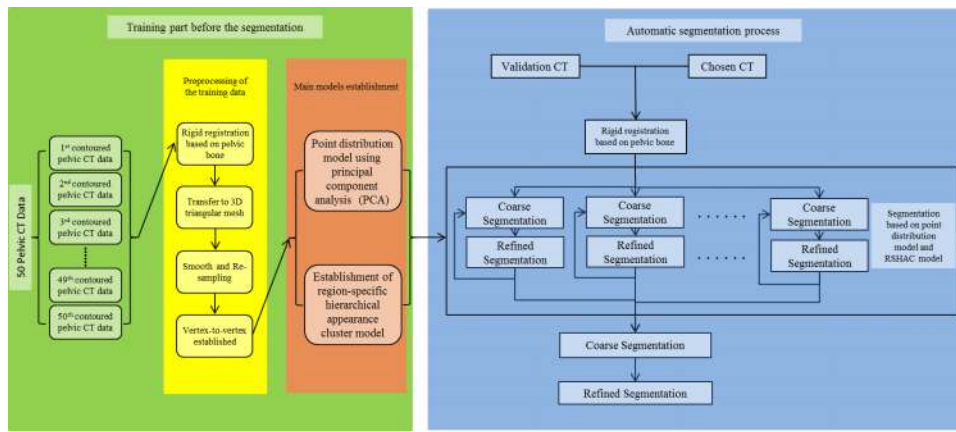


FIG. 3. The flow chart of the whole method.

2.B.1. Preprocessing of the training data

First, the prostate, bladder, and rectum were manually delineated slice by slice for the 50 training datasets. And then one CT image of the 50 training datasets was randomly selected as the reference image. The rest 49 CT images from the training dataset were registered to the reference image using the rigid registration based on the pelvic bone (Fig. 4). And then the 2D slice-by-slice mesh was then connected to 3D triangular mesh based on “Marching cubes.”³⁹ After the 3D triangular mesh was obtained, the jagged contours were smoothed using a simple Laplacian smoothing operation by moving each vertex to the average position with neighboring vertices.⁴⁰ The smoothing method was performed with five iterations. Second, the number of triangles was reduced by a down-sampling process. The triangle meshes of the pelvic organs were reduced to 10% of the original mesh size. After the down-sampling, the vertices’ number of the contour of each organ was about 3000. Therefore, each contour with about 3000 points was adjusted to 3000 points by an up-sample or down-sample. This up- or down-sample was averaged distributed in the whole contour according to the number of vertices which were needed to be increased or decreased. Third, the contour was smoothed again by the same method of the first step. Finally, due to the volume shrinkage caused by smoothing, the contour meshes were magnified to the

same volume as the original ones. After all contours of the data sets were confirmed, each contour was mapped to binary image after the second rigid registration based on the pelvic bone between each CT image. The binary images will be well registered by the free-form deformation registration based on B-spine.⁴¹ The vertex-to-vertex between each two contours mesh will be found by applying the calculated deformable field to the contour vertex using tri-linear interpolation (Fig. 5). And the whole preprocessing was shown in Fig. 6.

The shapes of the pelvic organs were parameterized by the set of positions of L ($L = 3000$) contour vertex of each organ. If we denote $\vec{x}_j(i)$ as the position of the j th point in the i th CT in the training dataset, then the surface shape vector can be represented as

$$p_i = \left(\vec{x}_1(i), \dots, \vec{x}_L(i) \right) \in \mathcal{R}^{3D}. \tag{1}$$

2.B.2. Establishment of a point distribution model of pelvic organs

In this process, it was assumed that the set of surface shape vector could be seen as samples from a random process. For anatomical reasons, displacements of L vertices due to the deformation of pelvic organs were highly correlated, which implied that the dimensionality of the multivariate statistical

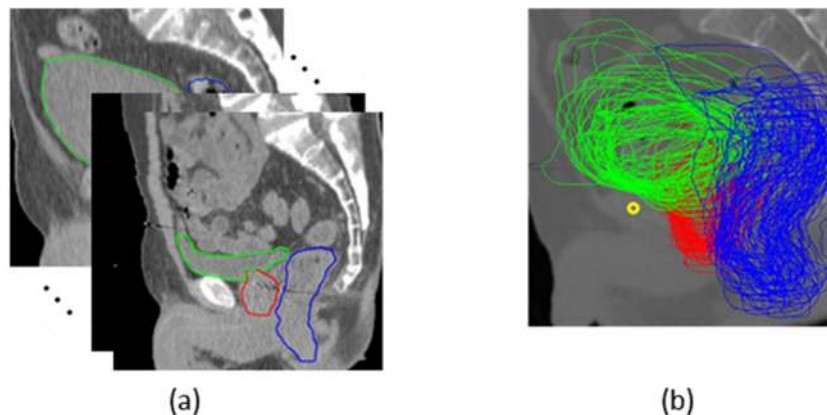


FIG. 4. (a) The training dataset containing 50 CT images of pelvic organs. (b) All the pelvic organs were aligned by pelvic bone, and the outline of each color represents one kind of organs (prostate, bladder, and rectum were contoured by red, green, and blue solid curves, respectively). (See color online version.)

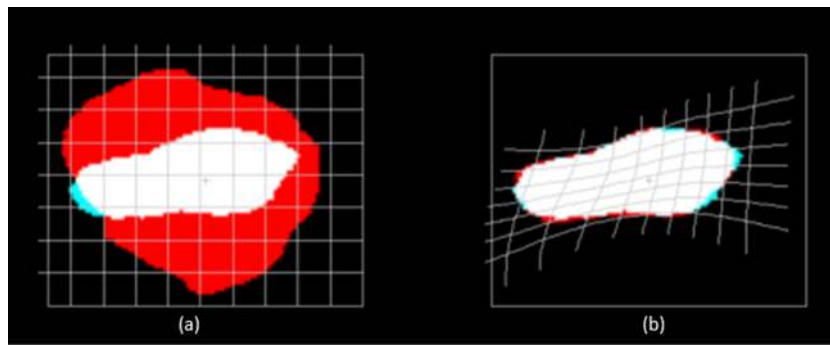


FIG. 5. (a) The bladder contours from two patients were mapped to two binary images. The cyan and red represent each image and the white represents the overlap of these two images. (b) Applying the B-spline free-form deformable registration, the two binary images were well registered. This deformation field will be applied to contour vertex using tri-linear interpolation. (See color online version.)

problem was actually much smaller than $3 \times L$. For a $3 \times L$ dimensional problem with N ($N = 50$) samples, we used a method from multivariate statistics, PCA,^{42,43} to model the anatomical shape variation of pelvic organs. The eigenmodes which represent the shape variation of the pelvic organs will be used to guide the segmentation of the validation images.

After the parameterization of the pelvic organs in the training dataset, the geometry of the pelvic organs in image i ($i = 1, \dots, N$) was represented by the shape vector p_i . According to the shape vector, the mean vector p_0 of the N vector was first computed by

$$p_0 = \frac{1}{N} \sum_{i=1}^N p_i. \tag{2}$$

Next, a covariance matrix C with dimensions $3L \times 3L$ which represents the shape variation of the pelvic organs was calculated,

$$C = \frac{1}{N-1} \sum_{i=1}^N (p_i - p_0) \cdot (p_i - p_0)^T. \tag{3}$$

After the matrix C was obtained, we decomposed the shapes of the deformed organ by the following formula:

$$p = p_0 + \sum_{n=1}^N \psi_n q_n, \|q_n\| = 1, \tag{4}$$

where q_n represent the eigenvectors of the covariance matrix of the input data p , corresponding to the N eigenvalues λ_n . The coefficients ψ_n were normally distributed random variables with zero means and the corresponding eigenvalues λ_n as variances.

In order to reduce the number of point distribution models, we need to choose M ($M < N$) eigenmodes to represent the deformation pattern of the pelvic organs. The mean residual error $R^{[M]}$ was calculated as follows:

$$P^{[M]} = p_0 + \sum_{n=1}^M \psi_n q_n, \|q_n\| = 1, \tag{5}$$

$$e^{[M]} = p^{[M]} - p, \tag{6}$$

$$R^{[M]} = \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{L} \sum_{l=1}^L \|e_{i,j}^{[M]}\| \right), \tag{7}$$

where $e_{i,j}^{[M]}$ is the 3D vector of local representation error of the j th vertex and i th contour.

Finally, the number of eigenmodes $M = 15$ was chosen such that 95% of the pelvic organs' shape variation was described and the mean residual $R^{[M]}$ close to 0. The eigenmodes $\{P_j\}_{j=1, \dots, M}$ and the corresponding weight matrix $\{T_j = q_j\}_{j=1, \dots, M}$ will be used for the segmentation process on validation CT images.

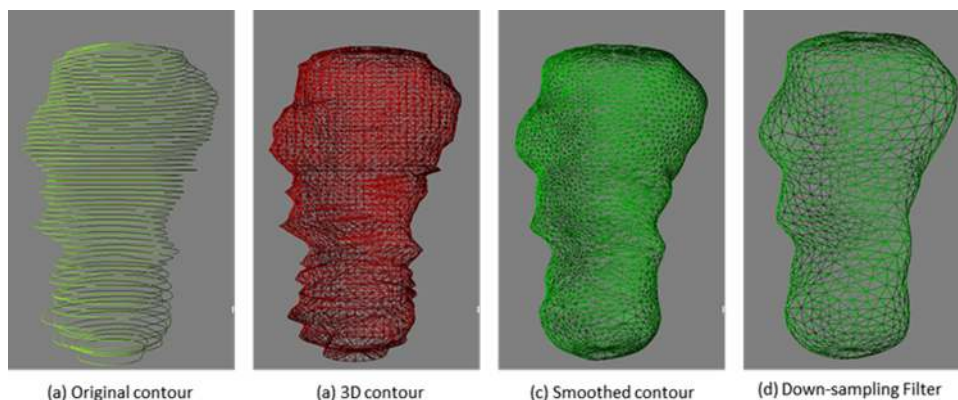


FIG. 6. The preprocessing of modeling with (a) the original 2D slice-by-slice mesh, (b) initial 3D triangular mesh, (c) contour with only one smooth process, and (d) the final contour with down-sampling and second smoothing.

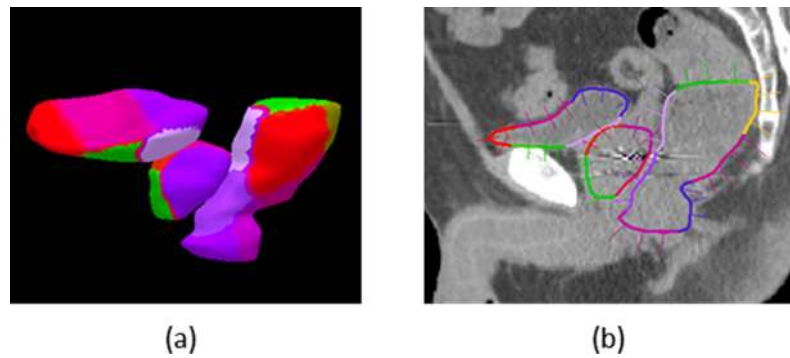


FIG. 7. (a) Pelvic organs were divided into five surface patches on the chosen shape. (b) The patch instance in one slice. In (a) and (b), the five surface patches were marked by green, yellow, red, purple, and pink, respectively. The overlap areas between each patch were marked by the average color of the corresponding patch. (See color online version.)

2.B.3. Establishment of an appearance model of pelvic organs

In order to improve the segmentation accuracy and efficiency, a reliable, region-specific appearance model of pelvic organs was established, which is the core part of the proposed method. In this process, each contour of pelvic organ was divided into five patch surfaces according to the suggestion of the chosen experienced radiation oncologist. At first, three radiation oncologists divided each pelvic organ of two chosen CT volume data into five patches independently. In this chosen process, the mutual information between two patch areas was calculated. And then, the mutual information between each pair of corresponding patch from the two CT data was calculated and averaged. At last, the radiation oncologist with the largest average mutual information was chosen to decide the patch pattern. After the patch pattern was obtained, the patch pattern was transformed to the 50 CT images of the training dataset based on the surface mapping vectors obtained in the preprocessing stage (the patch pattern on one of the CT images was shown in Fig. 7). For each surface patch, 50 CT images of the dataset which have been divided were grouped together for hierarchical clustering calculation using similarity based on the cross-correlation between the different patches as the figure of merit. In the hierarchical clustering process, N ($N = 50$) surfaces were first traded as N clusters. The distance between each two clusters was calculated using cross-correlation. The two clusters with the smallest distance

were combined to a new cluster. After then, the number of the clusters was decreased to $N - 1$. The average surface of this new cluster will be calculated as the representative of this cluster. And then the two clusters with the smallest distance were combined to a new cluster in the same way. This clustering calculation will be performed until all the N original clusters were combined (the number of cluster decreased to 1). After the clustering calculation was finished, each cluster was transformed to its corresponding profile by sampling voxel values along the normal direction of the contour surface. Finally, an agglomerative hierarchical cluster tree was established (see Fig. 8). The automatic segmentation will be performed from the first layer to the last layer of the agglomerative hierarchical cluster tree. After the segmentation in first layer, the next layer was chosen based on the similarity respective between the next two profiles. And the final surface of this patch will be obtained after the segmentation on the last layer of the agglomerative hierarchical cluster tree. In this process, the number of the profiles in the tree was a key influence of this algorithm. The segmentation results of each patch will be highly improved when the number of the profiles was huge enough to include every form of this patch.

2.C. Automatic segmentation on new images

This segmentation method uses principal component analysis to model organ shapes and hierarchical clustering to model the appearance of each organ surface patch. Coarse

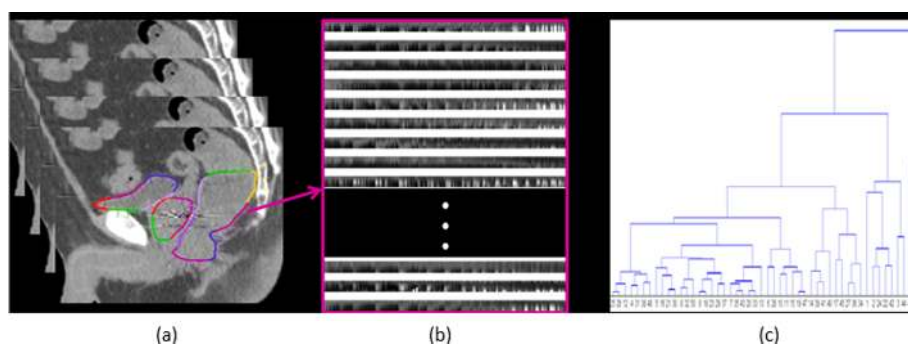


FIG. 8. For each surface patch, an agglomerative hierarchical cluster tree for the 50 profiles is calculated using similarity between the profiles as the figure of merit. (a) The patch pattern reflected by one slice in one of the 50 profiles. (b) The 50 patch surface of one patch region. (c) The agglomerative hierarchical cluster tree with 50 nodes in the last layer corresponding to the 50 profiles.

and refined segmentations are done iteratively for each surface patch to find the best reference patch profile via a hierarchical cluster decision tree. Final segmentation is obtained using the generated global reference profile.

The segmentation process on the validation CT images was divided into three stages. First, a rigid registration was performed between the validation CT and the chosen CT image of the training dataset. Second, each patch in the validation CT was segmented with a coarse segmentation and a refined segmentation based on the RSHAC and PCA eigenmodes. Last, the segmentation results of the five patches were combined to form a new segmented CT image as the initial segmentation of the whole validation CT [see Fig. 9(a)], and the final segmentation image was obtained with a coarse segmentation followed by a refined segmentation process on the new validation CT just acquired.

The pelvic organ contour of each patch on the validation CT was initialized after the rigid registration based on the pelvic bone with one of the CT images in the training dataset. The initial surface patch of the validation CT will be the surface patch of the chosen training CT. After the registration, the contour of the pelvic organs on the validation CT was initialized to five patch $c_{i=1,2,3,4,5}^* (\{P_j = 0\}_{j=1,\dots,M}) = c_{i=1,2,3,4,5}$. For each patch of the validation CT, the most similar patch in the RSHAC model will be found through several segmentation processes with two steps of each level of the tree model [Fig. 9(b)].

During the segmentation process on the RSHAC model, PCA modes were used as the guidance information in each segmentation calculation. Besides the deformation pattern stipulated by the eigenmodes, a cost function E_c based on the directional gradient field was also needed to measure how well the deformed contour matched the underlying edge of the

pelvic organs on the validation CT image. For this purpose, the extra interior and exterior contours of each patch $c_{i,int}$ and $c_{i,ext}$ were created by shrinking and expanding the vertices of the initial contour $c_{i=1,2,3,4,5}$ by 5 mm along the normal direction to each vertex.

The segmentation of patch No. 1 was chosen as an example of the segmentation process on the RSHAC model. (The segmentation of all the five patches was performed in the same time.) The segmentation contour $c_1^*(P)$ was deformed starting from the initial contour c_1 by changing $\{T_j = q_j\}_{j=1,\dots,M}$ which was the weight matrix of the eigenmodes $\{P_j\}_{j=1,\dots,M}$ based on the PCA modes,

$$c_1^*(P) = c_1 + T \cdot P. \tag{8}$$

At the same time, interior contour candidate $c_{1,int}^*(P)$ and exterior contour candidate $c_{1,ext}^*(P)$ were also deformed in the same way.

The cost function E_c was used to measure how well the deformed contour matched the underlying edge of the pelvic organs on the validation CT image. The cost function was evaluated on the interior, central, and exterior contours to capture information of the neighborhood of the edge of the pelvic organs,

$$E_c = \sum (|I_{ref}(C_{ref}) - I_{val}(C_{val})|) = \sum_{j=1}^{Num} (|(I_R(c_{1,j}) - I_R(c_{1,int,j})) - (I_V(c_{1,j}^*(P)) - I_V(c_{1,int,j}^*(P)))| + |(I_R(c_{1,j}) - I_R(c_{1,ext,j})) - (I_V(c_{1,j}^*(P)) - I_V(c_{1,ext,j}^*(P)))|), \tag{9}$$

where Num is the number vertex in this patch. $I_R(c)$ and $I_V(c)$ denote the intensity of the reference image and

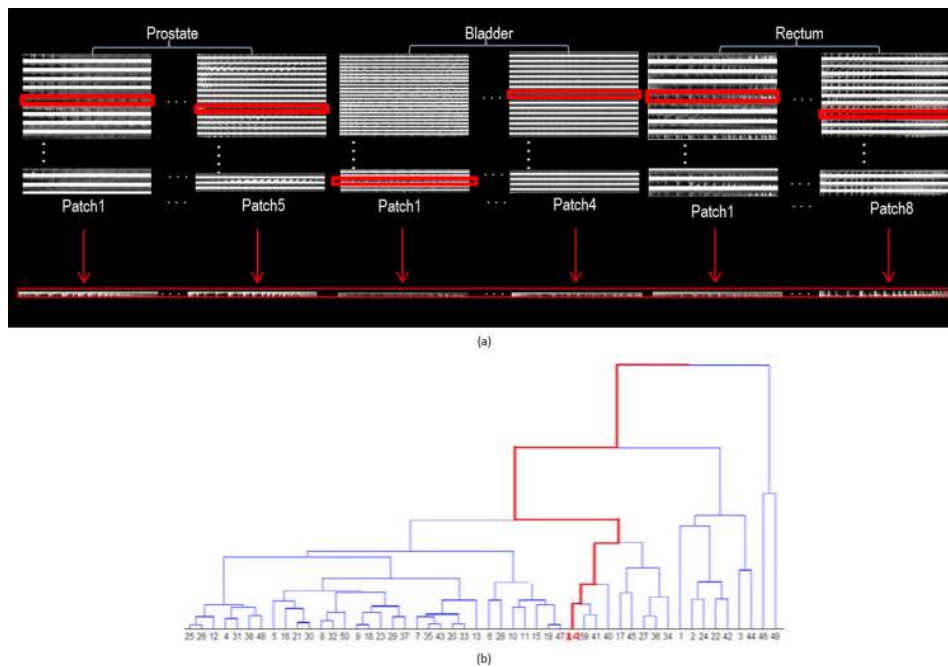


FIG. 9. Flow chart of the automatic segmentation process on the validation CT images. (a) The flow chart of the segmentation process of each patch and five patches made up the initial segmentation of the pelvic organs. (b) The process of finding the best patch in the agglomerative hierarchical cluster tree. The x-axis was the sequence number of 50 profiles. The red line is the example of a search path to the most similar patch of the 50 profiles. (See color online version.)

validation image sampled by the vertex of contour c . $(I_R(c_{1,j}) - I_R(c_{1,int,j}))$ and $(I_R(c_{1,j}) - I_R(c_{1,ext,j}))$ are the directional gradient values on the j th vertex of the interior and exterior reference contours. $(I_V(c_{1,j}^*(P)) - I_V(c_{1,int,j}^*(P)))$ and $(I_V(c_{1,j}^*(P)) - I_V(c_{1,ext,j}^*(P)))$ are the directional gradients on the j th vertex of the interior and exterior segmentation contours.

The optimum shape vector P^* was selected when the cost function E_c achieves the smallest value,

$$P^* = \arg \min(E|P) \quad P_i \in [-3\sigma_i, 3\sigma_i], \tag{10}$$

where σ_i is the square root of the eigen value λ_i . The search for the optimum shape parameter vector was carried out by a simplex optimizer within the interval $[-3\sigma_i, 3\sigma_i]$ to minimize the cost function of E_c .

After the optimum shape vector was obtained, the coarse segmentation process on this layer of the RSHAC model was finished. The most similar patch in this layer was found based on the similarity measure between the segmented validation patch and each patch of this layer.

However, the RSHAC model was established based on the training dataset. The similarity patch needs an adjustment to match the real situation of the validation CT. So a new cost function of the refined segmentation process was defined and used to adjust the chosen patch to match the underlying edge of this patch on the validation CT,

$$E_r = E_{\text{Similarity}} + \alpha E_{\text{Curvature}} + \beta E_{\text{Overlap}}, \tag{11}$$

where $E_{\text{Similarity}}$ denotes the level of similarity between the validation image and reference image,

$$E_{\text{Similarity}} = H(V) + H(R) - H(V,R), \tag{12}$$

where $H(V)$ and $H(R)$ denote the marginal entropies of the surface area between the interior and exterior contours of the validation image and the reference image; $H(V,R)$ denotes their joint entropy, which is calculated from the joint histogram of the two surface areas.

In order to constrain the segmentation to be smooth, the penalty term $E_{\text{Curvature}}$ was introduced to regularize the refined segmentation. The general form of such a penalty term has been described by Wahba.⁴⁴ In 3D, the penalty term takes the following form:

$$E_{\text{Curvature}} = \frac{1}{\text{Volume}} \int_0^X \int_0^Y \int_0^Z \left[\left(\frac{\partial^2 \text{TP}}{\partial x^2} \right)^2 + \left(\frac{\partial^2 \text{TP}}{\partial y^2} \right)^2 + \left(\frac{\partial^2 \text{TP}}{\partial z^2} \right)^2 + 2 \left(\frac{\partial^2 \text{TP}}{\partial xy} \right)^2 + 2 \left(\frac{\partial^2 \text{TP}}{\partial xz} \right)^2 + 2 \left(\frac{\partial^2 \text{TP}}{\partial yz} \right)^2 \right] d_x d_y d_z, \tag{13}$$

where Volume denotes the volume of the image domain. This quantity is the 3D counterpart of the 2D bending energy of a thin-plate of metal and defines a cost function which is associated with the smoothness of the segmentation.

In order to prevent the possibility of overlapping between the pelvic organs, a penalty term E_{Overlap} is defined. For two pelvic organs PO_1 and PO_2 , the E_{Overlap} was defined as the

ratio of the volume of intersection to the volume of union between PO_1 and PO_2 ,

$$E_{\text{Overlap}} = \frac{\text{Volume}(PO_1 \cap PO_2)}{\text{Volume}(PO_1 \cup PO_2)}, \tag{14}$$

where E_{Overlap} represent the degree of overlapping between each two pelvic organs.

After the refined segmentation of patch 1, the segmentation of one layer in the RSHAC model was finished. This segmentation process will continue in the next layer until the most similar patch in the last layer (one of the 50 profiles in this tree) was found and adjusted. The segmentation of the five patches was finished at the same time. In the last, preliminary contour of the whole three pelvic organs $c^*(P^*)$ was obtained by making the union of the five segmented surface of corresponding patch. And the overlap areas and holes between each two patches were averaged and interpolated. The pelvic organs of the validation CT were accurately localized and contoured after an adjustment based on $c^*(P^*)$ using $E_{\text{Curvature}}$ and E_{Overlap} .

2.D. Validation of segmentation results

In order to evaluate the performance of this automatic segmentation, the 210 validation CT images were manually delineated by an experienced medical physicist. The Dice similarity coefficient (DSC) between the automated and manually segmented validation CT images was first chosen as the evaluation criterion.⁴⁵ Specifically, DSC is defined as

$$\text{DSC} = \frac{2 * V_{a \cap m}}{V_a + V_m}, \tag{15}$$

where $V_{a \cap m}$ denotes the number of voxels in the overlapping area between the validation image segmented by the automatic (a) and the manual (m) methods. Similarly, V_a and V_m denote the number of voxels in the validation image, respectively, segmented by the automatic (a) and the manual (m) methods.

In order to further access the performance of this method, the mean absolute distance (MAD) was also calculated. The MAD is the average of absolute differences between each value in a dataset and the average value of this dataset. In the

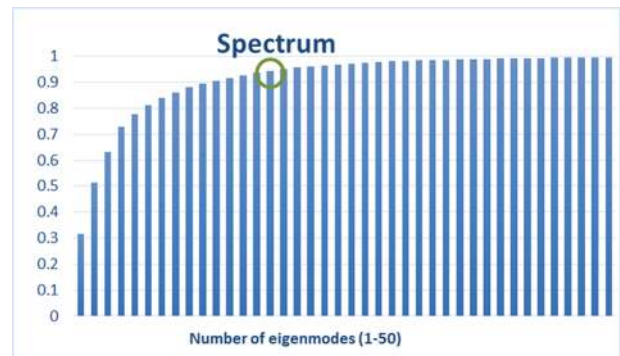


Fig. 10. The integral of the eigenvalues for the training dataset as a function of the number of PCA modes. The sum of all eigenvalues was normalized to 100%. The square column which represents the 15th eigenmodes was marked by a green circle.

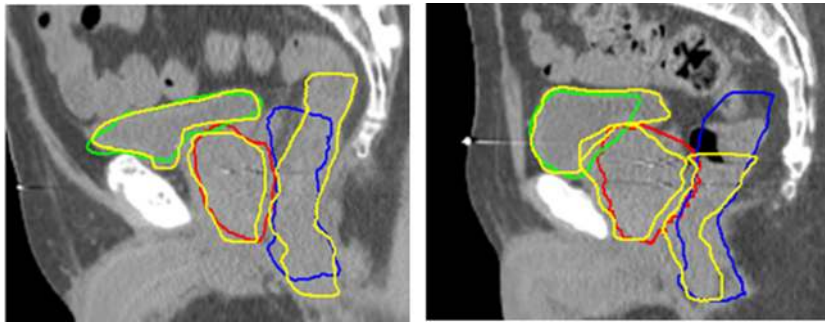


FIG. 11. The automatically segmented pelvic organ contour (red, green, and blue thick solid curves) and manually delineated contours of the pelvic organs (yellow thick solid curves) were overlaid on a sagittal section of a validation CT image. (See color online version.)

validation process, the dataset used for the calculation of MAD consists of the surface distance (SD) between the automated and manual segmentation results. The SD is calculated by the distance between the surfaces of each pelvic organ of the manual and automated segmentation images along 1000 rays directed to the spherical surface from the centroid of the manually contoured validation CT. The results of the evaluation will be introduced in Sec. 3.

3. EXPERIMENT RESULTS

We first determined the number of eigenmodes required in the PCA point distribution model. In order for the PCA model to accurately represent the deformation pattern of the pelvic organs, the average residual errors need to be controlled within 1 mm. At the same time, the number of eigenmodes needs to be as small as possible to minimize overfitting and reduce computational complexity. In Fig. 10, we can see 15 eigenmodes represented more than 95% (95.2%) of the pelvic organs' shape variation. So we chose 15 eigenmodes for the segmentation process to balance the accuracy and efficiency.

The whole process of the automatic segmentation process took on average 5 min on a personal computer equipped with Core 2 Duo CPU of 2.8 GHz and 8 GB RAM. An example was shown in Fig. 11, the comparison between automatically and manually segmented contours.

In order to quantify the accuracy of the automatic segmentation method on the validation dataset, we showed the average DSC value of each patient. The average test values of these

patients were shown by histogram in Fig. 12 (after ordering). And the average MAD of each patient was shown in Fig. 13.

The automatic segmentation results of the prostate and bladder agreed well with those based on manual segmentation. The average DSC of the prostate is 0.751 (range: 0.697 and 0.834). Our segmentation results of the prostate were consistent across patients with the maximum difference of 0.137 in terms of DSC, suggesting its robustness against differences of the prostate anatomy. The performance in the case bladder was slightly better, with an average DSC of 0.783, and sixty percent of the DSC value of the bladder was over 0.80. However, the segmentation of the rectum was not as good, with the average DSC being only 0.573, and the highest DSC value being less than 0.70. The spatial discrepancy based on MAD showed a similar trend. As mentioned earlier, many methods have been proposed about the automatic segmentation of the pelvic organs in the CT of prostate cancer. In order to intuitively contrast these methods, the parameters of these methods were shown in Table I.

In Table I, the results of our method were not the best because our method is very dependent on the size of the training dataset. But the increase of the size of the training dataset also results to the biggest improvement on our method. And our method is very suitable for the medical big data which is very popular in recent years. Our method can get huge promotion based on this research. Furthermore, to achieve better contrast effect, 52 CT images from five patients were manually delineated by another physicist. The average MAD of each patient between the pelvic contours from two physicists was calculated and shown in Fig. 14.

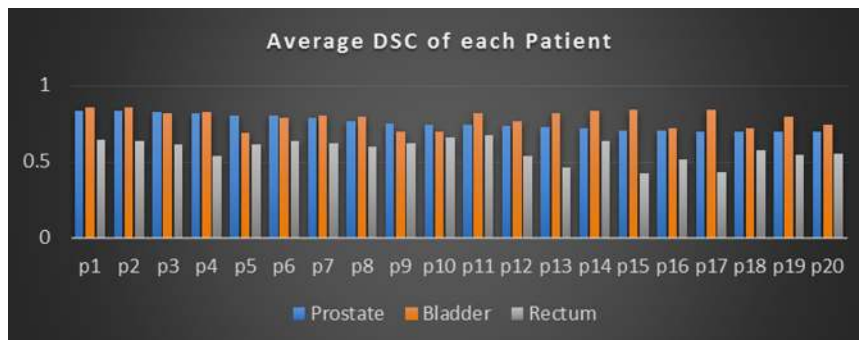


FIG. 12. The DSC of prostate (blue), bladder (orange), and rectum (gray) for twenty patients. (See color online version.)

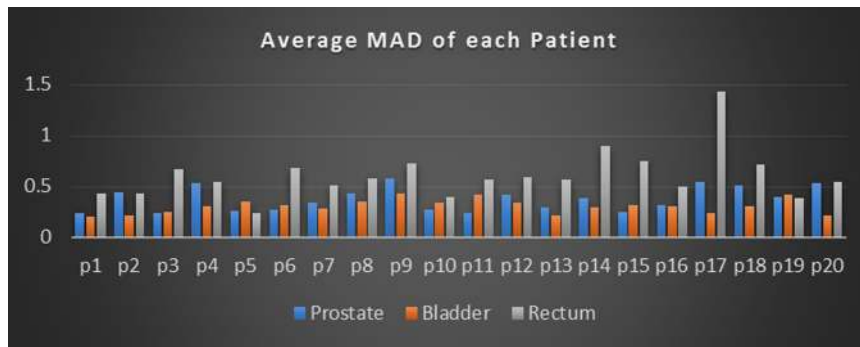


FIG. 13. The average MAD of prostate (blue), bladder (orange), and rectum (gray) for twenty patients. (See color online version.)

Through Fig. 14, we can see that the average MAD of prostate and bladder was little better than us. But due to the large physiological changes of the rectum, the MAD values of rectum were also unsatisfactory. The overall average MAD of the 52 CT was 0.601 cm which was almost the same with us.

4. DISCUSSION

In this paper, we presented a novel automatic, multiorgan segmentation method based on CT images of prostate cancer patients. We validated the accuracy and robustness of the method using a large number of CT images from a new group of patients independent of the training dataset. The key contributing factor to the success of our method is the region-specific agglomerative hierarchical cluster model for refining the result on top of the PCA-based gross segmentation. The segmentation accuracy without this fine-tuning process decreased sharply with the average DSC being only 0.567.

The relatively lower segmentation accuracy for rectum compared with prostate and bladder is likely due to the large variations between each situation of the rectum (mainly bowl filling and bowl gas). The relatively small size of our training dataset (50 CT images of eight patients) may not be sufficient to capture all the possible changes of the rectum. Also, the manual contour of rectum is highly variable between different observers given a decreased soft-tissue contrast near the colon and rectum interface, which may contribute to the lower accuracy.

In previous work, Davis *et al.* proposed a method based on patient-specific registration,³⁵ with a mean DSC slightly better than ours. However, this method was based on the patient-specific database, so it cannot segment the first few CT scans

during treatment. Further, like most previous methods, their method was used for segmentation of prostate only. On the contrary, our method does not have the imposed restriction of training dataset and we can segment the prostate, bladder, and rectum at the same time.

This method was based on a hypothesis that the surface contours of each patch area in the validation CT image could be found in the RSHAC model when the training data set was large enough. So the segmentation accuracy in this study was limited by two main points. The first point is the rationality of the patch pattern based on the experience of doctors. The patch pattern in this study was limited by the opinions from doctors and the performance of the work station. So the patch pattern in this study was just the optimal pattern in these limitations. And that was why the $E_{Overlap}$ was used in the cost function while the patch in this patch pattern had overlap areas because the optimal segmentation results should have little overlap areas between each patch contour. The second point is the size of the training data set. The training data with 50 CT images in this study were far not large enough to include all the shapes of each patch area. Even so, the DSC values in this study were still around 80%. The segmentation result could already use as reference samples to help the physicists to speed up their work.

So there are mainly three ways to improve the accuracy of the segmentation results. The first way is increasing the diversity of the alternative patch pattern and improving the performance of the workstation to unfreeze the limitation of the patch number. The second way is expanding the training dataset. The diversity of the pelvic organ variations can be better represented by the RSHAC model with the expanded training dataset. For a single patient, the third way is changing the region-specific model from population based to specific

TABLE I. The comparison of different methods.

Method		Gao <i>et al.</i> (Ref. 24)	Gao <i>et al.</i> (Ref. 28)	Shao <i>et al.</i> (Ref. 29)	Feng <i>et al.</i> (Ref. 30)	Our method
Prostate	Mean DSC	0.87 ± 0.04	0.86 ± 0.05	0.88 ± 0.02	0.89 ± 0.05	0.75 ± 0.07
	Mean MAD (cm)	NA	NA	NA	NA	0.37
Bladder	Mean DSC	0.92 ± 0.05	0.91 ± 0.10	0.86 ± 0.08	NA	0.78 ± 0.08
	Mean MAD (cm)	NA	NA	NA	NA	0.303
Rectum	Mean DSC	0.88 ± 0.05	0.79 ± 0.20	0.84 ± 0.05	NA	0.57 ± 0.08
	Mean MAD (cm)	NA	NA	NA	NA	0.604

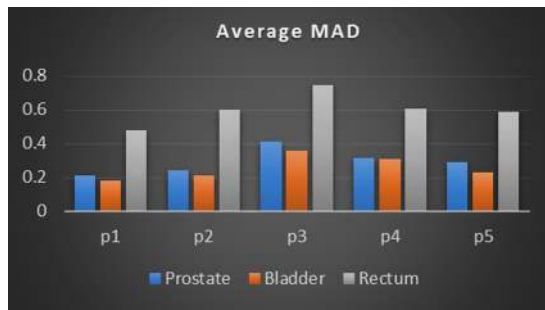


FIG. 14. The average MAD of five patients between the pelvic contours from two physicists.

patient based. Along with the increased number of daily CT images which from the same patient become available in fractionated radiotherapy, a new RSHAC model can be established, which will be only used for the segmentation of the specific patient.

For future work, we plan to develop a web-based platform to streamline image visualization, segmentation, and validation on a large database. Finally, we note that even with suboptimal results in certain patients, the automatic segmentation of the prostate, bladder, and rectum can be a good starting point for the subsequent manual delineation.

5. CONCLUSION

We have presented a novel method for automatic, multi-organ segmentation based on CT images of the male pelvis. We combined the PCA-based point distribution model with RSHAC model to achieve accurate, robust, and efficient automatic segmentation of multiple pelvic organs on CT images. This method should be useful in future radiation treatment planning and adaptive radiotherapy for prostate cancer.

ACKNOWLEDGMENTS

Dengwang Li and Pengxiao Zang contributed to this work equally. The author would like to express great thanks to the staff in the Shandong Cancer Hospital and Institute for their valuable suggestions to this work. This work is supported by NIH/NIBIB (Grant Nos. 1R01-EB016777 and 1R00CA166186), NSFC (Grant Nos. 61471226 and 61201441), and research funding from Shandong Province (Grant No. JQ201516). Thanks to the support of Taishan scholar project of Shandong Province.

CONFLICT OF INTEREST DISCLOSURE

The authors have no COI to report.

^{a)} Author to whom correspondence should be addressed. Electronic mail: lidengwang@sdu.edu.cn

¹ A. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, C. Smigal, and M. J. Thun, "Cancer statistics 2006," *CA-Cancer J. Clin.* **56**, 106–130 (2006).

- ² T. Tarver, "Cancer facts and figures 2012. American Cancer Society (ACS)," *J. Consum. Health Internet* **16**, 366–367 (2012).
- ³ 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).
- ⁴ D. C. Collier, S. S. Burnett, M. Amin, S. Bilton, C. Brooks, A. Ryan, D. Roniger, D. Tran, and G. Starkschall, "Assessment of consistency in contouring of normal-tissue anatomic structures," *J. Appl. Clin. Med. Phys.* **4**, 17–24 (2003).
- ⁵ E. E. Ahunbay, C. Peng, S. Holmes, A. Godley, C. Lawton, and X. A. Li, "Online adaptive replanning method for prostate radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **77**, 1561–1572 (2010).
- ⁶ S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric x-ray CT images," *IEEE Trans. Med. Imaging* **20**, 490–498 (2001).
- ⁷ J. Tschirren, E. A. Hoffman, G. McLennan, and M. Sonka, "Intrathoracic airway trees: Segmentation and airway morphology analysis from low-dose CT scans," *IEEE Trans. Med. Imaging* **24**, 1529–1539 (2005).
- ⁸ Y. Gu, V. Kumar, L. O. Hall, D. B. Goldhof, C. Y. Li, R. Korn, C. Bendtsen, E. R. Velazquez, A. Dekker, H. Aerts, and P. Lambin, "Automated delineation of lung tumors from CT images using a single click ensemble segmentation approach," *Pattern Recognit.* **46**, 692–702 (2013).
- ⁹ D. Aykac, E. A. Hoffman, G. McLennan, and J. M. Reinhardt, "Segmentation and analysis of the human airway tree from three-dimensional x-ray CT images," *IEEE Trans. Med. Imaging* **22**, 940–950 (2003).
- ¹⁰ S. G. Armato and W. F. Sensakovic, "Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis," *Acad. Radiol.* **11**, 1011–1021 (2004).
- ¹¹ J. M. Kuhnigk, V. Dicken, L. Bornemann, A. Bakai, D. Wormanns, S. Krass, and H. O. Peitgen, "Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans," *IEEE Trans. Med. Imaging* **25**, 417–434 (2006).
- ¹² P. Korfiatis, S. Skiadopoulos, P. Sakellaropoulos, C. Kalogeropoulou, and L. Costaridou, "Combining 2D wavelet edge highlighting and 3D thresholding for lung segmentation in thin-slice CT," *Br. J. Radiol.* **80**, 996–1005 (2014).
- ¹³ Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features," *IEEE Trans. Med. Imaging* **27**, 1668–1681 (2008).
- ¹⁴ O. Ecabert, J. Peters, H. Schramm, C. Lorenz, J. Von Berg, M. J. Walker, M. Vembar, M. E. Olszewski, K. Subramanian, G. Lavi, and J. Weese, "Automatic model-based segmentation of the heart in CT images," *IEEE Trans. Med. Imaging* **27**, 1189–1201 (2008).
- ¹⁵ I. Isgum, M. Staring, A. Rutten, M. Prokop, M. A. Viergever, and B. Van Ginneken, "Multi-Atlas-based segmentation with local decision fusion—Application to cardiac and aortic segmentation in CT scans," *IEEE Trans. Med. Imaging* **28**, 1000–1010 (2009).
- ¹⁶ Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features," in *IEEE Conference on Computer Vision (IEEE, Rio de Janeiro, 2007)*, pp. 1–8.
- ¹⁷ H. A. Kirisli, M. Schaap, C. T. Metz, A. S. Dharampal, W. B. Meijboom, S. L. Papadopoulou, A. Dedic, K. Nieman, M. A. De Graaf, M. F. L. Meijs, and M. J. Cramer, "Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography," *Med. Image Anal.* **17**, 859–876 (2013).
- ¹⁸ K. T. Bae, M. L. Giger, C. T. Chen, and C. E. Kahn, Jr., "Automatic-segmentation of liver structure in CT images," *Med. Phys.* **20**, 71–78 (1993).
- ¹⁹ L. Soler, H. Delingette, G. Malandain, J. Montagnat, N. Ayache, C. Koehl, O. Dourthe, B. Malassagne, M. Smith, D. Mutter, and J. Marescaux, "Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery," *Comput. Aided Surgery* **6**, 131–142 (2001).
- ²⁰ 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).
- ²¹ C. Chu, M. Oda, T. Kitasaka, K. Misawa, M. Fujiwara, Y. Hayashi, Y. Nimura, D. Rueckert, and K. Mori, "Multi-organ segmentation based on spatially-divided probabilistic atlas from 3D abdominal CT images," in

- Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (Springer, Berlin, Heidelberg, 2013), Vol. 8150, pp. 165–172.
- ²²D. Li, L. Liu, D. S. Kapp, and L. Xing, “Automatic liver contouring for radiotherapy treatment planning,” *Phys. Med. Biol.* **60**, 7461–7483 (2015).
- ²³D. Freedman and T. Zhang, “Interactive graph cut based segmentation with shape priors,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (IEEE, Los Alamitos, CA, 2005), Vol. 1, pp. 755–762.
- ²⁴Y. Gao, Y. Shao, J. Lian, A. Z. Wang, R. C. Chen, and D. Shen, “Accurate segmentation of CT male pelvic organs via regression-based deformable models and multi-task random forests,” *IEEE Trans. Med. Imaging* **35**, 1532–1543 (2016).
- ²⁵W. 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).
- ²⁶Y. Gao, S. Liao, and D. Shen, “Prostate segmentation by sparse representation based classification,” *Med. Phys.* **39**, 6372–6387 (2012).
- ²⁷S. 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).
- ²⁸Y. Gao, J. Lian, and D. Shen, “Joint learning of image regressor and classifier for deformable segmentation of CT pelvic organs,” *Medical Image Computing and Computer-Assisted Intervention-MICCAI* (Springer, New York, NY, 2015), Vol. 9351, pp. 114–122.
- ²⁹Y. Shao, Y. Gao, Q. Wang, X. Yang, and D. Shen, “Locally-constrained boundary regression for segmentation of prostate and rectum in the planning CT images,” *Med. Image Anal.* **26**, 345–356 (2015).
- ³⁰Q. 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).
- ³¹S. 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).
- ³²R. E. Broadhurst, J. Stough, S. M. Pizer, and E. L. Chaney, “Histogram statistics of local model-relative image regions,” in *Deep Structure, Singularities, and Computer Vision* (Springer, Berlin, Heidelberg, 2005), Vol. 3753, pp. 72–83.
- ³³N. Lay, N. Birkbeck, J. Zhang, and S. Zhou, “Rapid multi-organ segmentation using context integration and discriminative models,” in *Information Processing in Medical Imaging* (Springer, Berlin, Heidelberg, 2013), Vol. 7917, pp. 450–462.
- ³⁴C. Lu, Y. Zheng, N. Birkbeck, J. Zhang, T. Kohlberger, C. Tietjen, T. Boettger, J. Duncan, and S. Zhou, “Precise segmentation of multiple organs in CT volumes using learning-based approach and information theory,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (Springer, Berlin, Heidelberg, 2012), Vol. 7511, pp. 462–469.
- ³⁵B. 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)* (Springer, Berlin, Heidelberg, 2005), Vol. 3749, pp. 442–450.
- ³⁶U. Malsch, C. Thieke, P. E. Huber, and R. Bendl, “An enhanced block matching algorithm for fast elastic registration in adaptive radiotherapy,” *Phys. Med. Biol.* **51**, 4789–4806 (2006).
- ³⁷S. Liao and D. Shen, “A learning based hierarchical framework for automatic prostate localization in CT images,” in *Prostate Cancer Imaging, Image Analysis and Image-Guided Interventions* (Springer, Berlin, Heidelberg, 2005), Vol. 6963, pp. 1–9.
- ³⁸J. V. Stough, R. E. Broadhurst, S. M. Pizer, and E. L. Chaney, “Regional appearance in deformable model segmentation,” in *Information Processing in Medical Imaging* (Springer, Berlin, Heidelberg, 2007), Vol. 4584, pp. 532–547.
- ³⁹W. E. Lorensen and H. E. Cline, “Marching cubes: A high-resolution 3D surface construction algorithm,” *Comput. Graphics* **21**, 163–196 (1987).
- ⁴⁰A. Nealen, T. Igarashi, O. Sorkine, and M. Alexa, “Laplacian mesh optimization,” in *Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques, Australasia and Southeast Asia* (ACM GRAPHITE, New York, NY, 2006), pp. 381–389.
- ⁴¹D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, “Nonrigid registration using free-form deformations: Application to breast MR images,” *IEEE Trans. Med. Imaging* **18**, 712–721 (1999).
- ⁴²M. Söhn, M. Birkner, D. Yan, and M. Alber, “Modelling individual geometric variation based on dominant eigenmodes of organ deformation: Implementation and evaluation,” *Phys. Med. Biol.* **50**, 5893–5908 (2005).
- ⁴³I. Jolliffe, *Principal Component Analysis* (John Wiley & Sons, Ltd., New York, NY, 2002).
- ⁴⁴G. Wahba, *Spline Models for Observational Data* (Siam, Philadelphia, PA, 1990).
- ⁴⁵L. R. Dice, “Measures of the amount of ecologic association between species,” *Ecology* **26**, 297–302 (1945).