Learning Image Context for Segmentation of Prostate in CT-Guided Radiotherapy

Wei Li[1,2], Shu Liao[1], Qianjin Feng[2], Wufan Chen[2], and Dinggang Shen[1,*]

[1] IDEA Lab, Department of Radiology and BRIC,
University of North Carolina at Chapel Hill, USA
[2] Biomedical Engineering College, Southern Medical University, Guangzhou, China
dgshen@med.unc.edu

Abstract. Segmentation of prostate is highly important in the external beam radiotherapy of prostate cancer. However, it is challenging to localize prostate in the CT images due to low image contrast, prostate motion, and both intensity and shape changes of bladder and rectum around the prostate. In this paper, an online learning and patient-specific classification method based on location-adaptive image context is proposed to precisely segment prostate in the CT image. Specifically, two sets of position-adaptive classifiers are respectively placed along the two coordinate directions, and further trained with the previous segmented treatment images to jointly perform the prostate segmentation. In particular, each location-adaptive classifier is recursively trained with different image context collected at different scales and orientations for better identification of each prostate region. The proposed learning-based prostate segmentation method has been extensively evaluated on a large set of patients, achieving very promising results.

Keywords: Radiotherapy, Prostate segmentation, Classification, Image context.

1 Introduction

Prostate cancer is one of the most common cancers in males and is a leading cause of male cancer death in US [1]. The external beam radiation treatment can provide a non-invasive and effective therapy for prostate cancer. It is usually planned on a planning CT image, namely *planning image*, on which the prostate and nearby critical structures are manually contoured. Then, the treatment is delivered in daily fractions during a period of eight to nine weeks. At each treatment day, a new CT image, namely *treatment image*, is acquired to guide the dose delivery, with goal of maximizing the dose delivered to the prostate and minimizing the dose delivered to healthy tissue such as bladder and rectum. Therefore, the success of image-guided radiotherapy (IGRT) highly depends on the accurate localization of prostate.

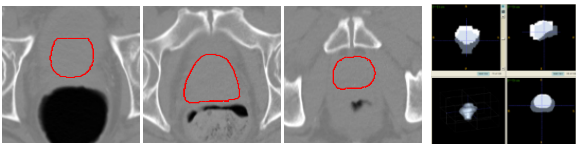
However, the prostate segmentation from CT images is challenging mainly due to three factors. *First*, prostate has low intensity contrast in the CT images. *Second*, the prostate motion/deformation (mainly caused by both water in the bladder and air in

* Corresponding author.

G. Fichtinger, A. Martel, and T. Peters (Eds.): MICCAI 2011, Part III, LNCS 6893, pp. 570–578, 2011.

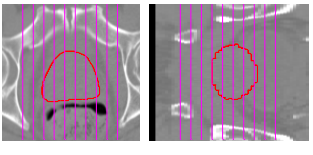
the rectum) is complex and unpredictable, even after rigid alignment of treatment images to the planning image based on the bone structures. *Third*, the uncertain presence of rectum gas makes the appearance of treatment images inconsistent, as demonstrated in Fig.1.

So far, several categories of methods have been proposed for prostate segmentation from CT images. The first category of segmentation methods is the deformable-model based segmentation [2-4]. The second category of segmentation methods is the statistical-atlas based segmentation with registration techniques [5, 6]. Also, other methods have been proposed based on different optimization strategies [7, 8], or working on different imaging modalities, e.g., MR and US [9]. In this paper, we focus on developing a learning-based method to segment prostate from CT images.



(a) (b)

Fig. 1. (a) Slices showing the prostate and surrounding organs (i.e., bladder, rectum). The black parts are the air in the rectum, and the red curves are the boundaries of prostate manually delineated by a radiologist. (b) 3D-display showing large motion/deformation of prostate in the two treatment images of the same patient, even after bone-based rigid alignment. Here, white and grey represent the two prostates after bone-based registration.



(a) (b)

Fig. 2. Illustration of location-adaptive classifiers placed along the two coordinate directions around prostate region. Each red straight line represents the place where we will place a location-adaptive classifier. The red curves represent the boundaries of prostate in (a) axial and (b) sagittal slices.

It is well known that the context and high-level information plays an important role in object recognition and scene understanding, especially for medical image analysis [10, 11]. Recently, a supervised machine learning method with context information has been proposed in [12]. In this work, the discriminative probability maps created by the learned classifier are used as context information recursively. Motivated by [12], we propose a patient-specific classification method for segmentation of prostate using context and appearance features under the online-learning framework. In particular, we improve the segmentation of prostate by learning both context and appearance information from the previous treatment images, thus achieving accurate segmentation for new treatment images of the same patient.

Moreover, we further propose a location-adaptive classification method to improve the classification accuracy of prostate. This is motivated by the fact that a single learned global classifier and its model parameters (trained globally through the entire volume) may not properly reflect the features of the local regions. Some similar ideas of employing distributed local classifiers have been used for medical image segmentation [13]. Specifically, we will train two sets of location-adaptive classifiers along the two coordinate directions for a 3D region of interest (ROI) around prostate, as shown in Fig. 2 (a) and (b). Each location-adaptive classifier is trained only for some 2D slices within in the ROI. The classification results with the two sets of location-adaptive classifiers are fused into a final 3D classification result and further segmented by a level-set based method.

2 Method

Our approach consists of an online training stage and a classification stage in the course of radiotherapy for the patient. **In the training stage**, two sets of location-adaptive classifiers are initially trained by the patient's first 3 daily scans with manual segmentations, which can be further online-updated by the latest acquired scans (e.g., 4 scans) of the same patient during the radiotherapy. With more and more subsequent treatment images are acquired and segmented, our classifiers can be updated with the automatic segmentation results of the latest-acquired treatment images (e.g., 4 images), along with the first 3 daily scans with manual segmentations. In this way, our trained classifiers can track the up-to-date patient-specific image context and appearance to improve the classification of prostate in the next daily treatment image. Note that our approach can work fully automatically after the initialization.

Specifically, to allow our classifiers to learn the image context and appearance in the training images, besides normalizing their intensities, the patient's pose in these images need to be aligned. We can use the pelvic bones extracted (with simple thresholding) from the training images to rigidly align all the training images onto the planning image space. For saving the time for both the training and testing stages, we further extract a ROI around the prostate, i.e., shown in Fig. 2. Then, to train each location-adaptive classifier, we just collect image features from the corresponding slices that this classifier is responsible for, and then use a machine learning technique [12] to build this classifier, along with their respective segmentation labels. Finally, these trained classifiers can be used to provide a probability segmentation map for prostate in the new treatment image, which can be further transformed into a binary segmentation by a level-set based technique. Note that both the treatment images and their respective segmentation results by our method can be used to update the training dataset for online refinement of our classifiers.

In the segmentation stage, each new treatment image is first aligned onto the planning image space. Then, the ROI used in the training stage will be used to extract the target region in the current aligned treatment image, which will then be classified/labeled by the two sets of our trained location-adaptive classifiers. The two prostate probability maps produced by the two sets of classifiers can be fused into a single probability map and further transformed into a binary segmentation by a level-set based technique [14]. Finally, the binary prostate segmentation result, along with its corresponding intensity image, can be used to update the classifiers.

2.1 Auto-context Classification Formulation

In this section, we briefly present the formulation of auto-context classification and notations. More details can be referred to [12]. We view an image as a function X from the spatial domain $\grave{U} \subset \mathbb{R}^2$ to an intensity value in $\mathbb{R}$. For each image X, its corresponding ground-truth segmentation/label image Y is also available (i.e., manual prostate segmentation), with value of each pixel belonging to $\{0,1\}$. There are M images used for training. And the j-th pixel of image X_i $(i = 1 \cdots M)$ is represented by x_{ij}, where $j \in \{1 \cdots N\}$ and N is the number of pixels in image X_i. Then, the training set can be represented by $S = \{(x_{ij}, y_{ij}), i = 1 \cdots M, j = 1 \cdots N\}$, where y_{ij} is the corresponding label in the label image Y_i.

We aim to learn a classifier to find an optimal classification result through Maximum a Posteriori (MAP) given a new image X, i.e. $Y^* = argmax(p(Y|X))$. Here, the marginal distribution is employed to solve $p(Y|X)$. To better approximate the marginal distribution, an auto-context model is proposed, in which the classification map produced by a traditional classifier, such as SVM or AdaBoost, is integrated into the training data set recursively to input context information. The iterative algorithm updates $p^{(t+1)}(y_j|B_j(X), C^t)$ to approach $p(y_j|X)$ asymptotically, where C^t is the context features at the t-th iteration and $B_j(X)$ denotes all pixels in an image patch centered at the j-th pixel.

2.2 Location-Adaptive Classifier

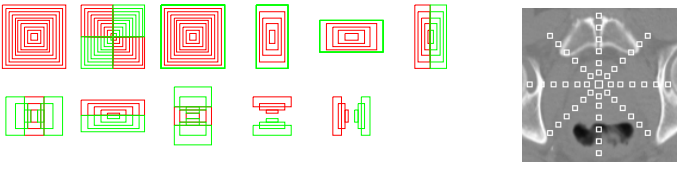
The identification of prostate, bone, air, and other tissue regions is important for training the location-adaptive classifiers. To achieve this, we use the segmentation labels of training images to locate the prostate, and identify (i.e., with simple thresholding) bone, air and other tissue regions on the training images. After that, we randomly sample pixels from each of the above regions as training instances. Then, the image appearance features of these training instances are computed and denoted as A_{kl} $(k = 1 \cdots K, l = 1 \cdots L)$, where K is the number of slices used for training a specific location-adaptive classifier, and L is the number of training instances. Context features are computed from the classification map by using a pattern that includes a large number of locations, i.e., shown in Fig. 3(b). These context features are denoted by $C_{kl}^t (k = 1 \cdots K, l = 1 \cdots L, t = 0 \cdots T)$. Finally, we can iteratively train our location-adaptive classifier. At the end of the training procedure, we get a sequence of classifiers L_t $(t = 0 \cdots T)$ for a specific location-adaptive classifier.

2.3 Feature Extraction

Two types of features are employed in our learning-based segmentation algorithm. The first one is the appearance feature, calculated from the original image. The second one is the context feature, computed from the classification map. The appearance features include a set of rectangular Haar features as used in [15]. Haar features are widely used in object recognition, which can be computed at different scales and with high speed by using integral images. We extract Haar features from different scales to facilitate our prostate classification. Fig. 3(a) shows some 2D Haar-like features used in our algorithm. Other appearance features adopted are the

histogram of oriented gradient (HOG) features[16]. The patch size for computing the above appearance features is 21×21. Finally, the coordinate of each pixel is also included as feature for learning the movement information of prostate in the pelvic.

Context features are updated recursively based on the classification maps which are produced iteratively. For each pixel of interest, a number of rays in equal-degree intervals are extended out from the current pixel and we sparsely sample the context locations on these rays, as shown in Fig. 3(b). For each context location, their classification probabilities and the mean probability within 3×3 windows are included as context features. We consider that the context information is also from the intensity variation in these selected context locations. Thus, the absolute intensity variation between the context location and current pixel of interest is also included as context features in our algorithm. Then, the classifiers can learn location information from the air or bone regions (which often have large intensity variation from the current pixel of interest) through our context features. These context features implicitly represent the shape and configuration information of the prostate.



(a) (b)

Fig. 3. (a) Examples of some 2D Haar-like features. (b) Context sample pattern

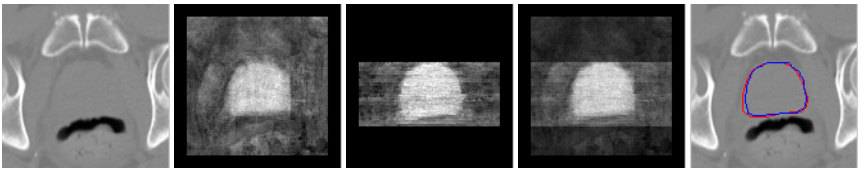
3 Experiments

Our data consists of 11 patients, each with more than 9 daily CT scans, with total image count of 161. Each image has resolution of 1×1×3 mm^3. The first 3 images of each patient are used to initialize the patient-specific learning process. The expert manual segmentation results are also available for each image. Since the classifiers are designed to learn the patient-specific information, all training and testing experiments are done on the images from the same patient. Hereafter, we index all images starting from number 1, i.e., the planning image will be image 1 and the first treatment image will be image 2, and so on. Three quantitative measures are used to evaluate the performance of our algorithm by comparing the automated segmentation results with the manual segmentations, and they are: Dice similarity coefficient (DICE), average surface distance (ASD), and centroid distance (CD).

Fusion of results along different coordinate directions. To validate the effectiveness of the fusion on the two results obtained from two classifier sets along two coordinate directions, we carried out a visual comparison between the results obtained from classifiers along each direction and the results obtained by the fusion. The fusion result was obtained by averaging the two classification results; then, a level-set based technique was used to extract smooth prostate boundary surface from the fused result.

Fig. 4 shows a sample fusion result. From Fig. 4, it can be observed that the fusion result is better than the results obtained along two directions separately.

Evaluation of performance on different number of training images. Here, we tested the performance of our proposed method under different number of training images. As examples, we selected 3 patients (patient ID 3, 5 and 10) with larger inter-treatment segmentation performance, e.g., these three patients have large variations of Dice measures compared to other patients (see Fig. 6(a)). Next, we selected one particular image from each of these three selected patients, i.e., image 11 of patient 3, image 12 of patient 5, and image 14 of patient 10. Then, we tested on each of these 3 images by using our location-adaptive classifiers trained by the latest 2 to 6 treatment images of the same patient. The Dice measures are shown in Fig. 6(c). It can be observed that the Dice measures become steady when the number of training images is 4 or larger. Therefore, the number of the training images is set to 3~7 for the tradeoff between training time and segmentation accuracy in all our experiments.



(a) (b) (c) (d) (e)

Fig. 4. Demonstration on the fusion of classification results along two coordinate directions. (a) Original image; (b) Classification result along z coordinate; (c) Classification result along y coordinate; (d) Fusion of (b) and (c); (e) Ground truth (red) and our estimated result (blue).

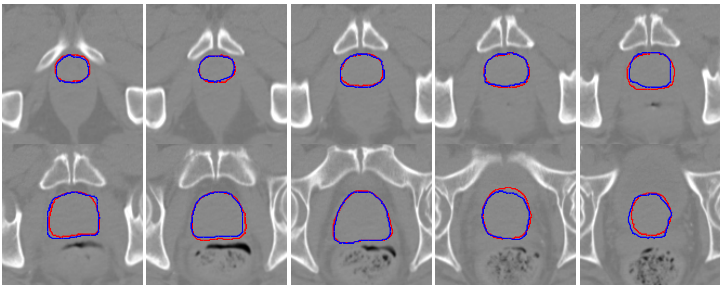
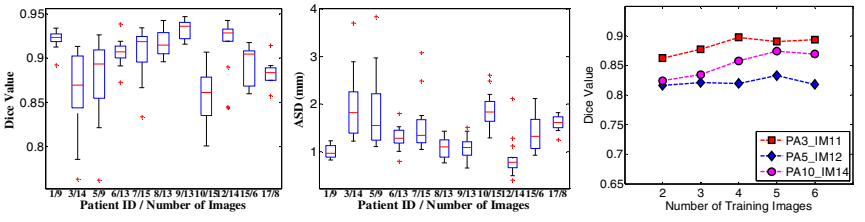


Fig. 5. Segmentation results on selected slices of image 11 of patient 5 (Dice 0.906). The red contours show the manually delineated boundaries of prostate by a radiation oncologist, while the blue contours show the results of the proposed method.

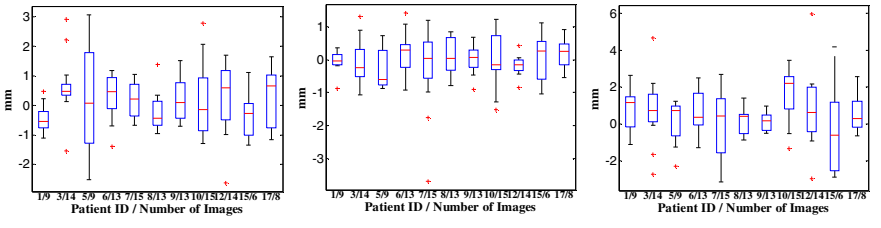
Evaluation of online learning mechanism. Since our method is learning-based, the number of training data cannot be too small. Thus, we use the first 3 images with manual segmentation and the latest 4 images with automatic segmentations to train our classifiers. In this way, our location-adaptive classifiers can learn the up-to-date information and reserve the information from manual segmentations of radiologist as

well. Fig. 5 shows an example on segmentation results for a patient image. The results on all images of all 11 patients are shown in Fig. 6(a) for Dice measures and Fig. 6(b) for ASD measures. Also, the centroid differences in lateral (x), anterior-posterior (y), and superior-inferior (z) directions are shown in Fig. 7. Three existing state-of-the-art prostate segmentation algorithms [3], [5] and [2] are compared with our method. Our mean Dice measure and mean ASD on all patients are 0.908 and 1.40 mm, respectively, which are better than the mean Dice measure of 0.893 and the mean ASD of 2.08 mm reported in [3]. Our mean centroid differences are 0.18 mm, -0.02 mm and 0.57 mm along x, y and z directions, which are comparable to the best result of -0.26 mm, 0.35 mm, and 0.22 mm reported in [5], while our mean Dice measure 0.908 is significantly higher than theirs 0.82. Our median probability of detection and false alarm are 0.90 and 0.10, respectively, which are much better than 0.84 and 0.13 reported in [2].



(a) (b) (c)

Fig. 6. (a) Dice measure of results and (b) average surface distance (ASD). (c) Change of the performance of our proposed method w.r.t. the different number of images used for training. Here, "PA3_IM11" means the image 11 of patient 3.



(a) (b) (c)

Fig. 7. (a), (b) and (c) are the centroid differences in lateral (x), anterior-posterior (y), and superior-inferior (z) directions, respectively. The symbols in the figure are the same as Fig. 6.

4 Conclusion

We have proposed a patient-specific online-learning classification method based on image context information, for segmenting prostate in the CT scans during radiotherapy. Our method uses only the training samples from the previous treatment

images of the same patient to specifically learn the patient-specific information for enhancing the prostate segmentation result, which is different from the conventional deformable-model or image-registration based methods that generally use the population information to guide the segmentation. Our trained classifiers can be updated online by the latest-acquired treatment images, thus better adapting to the prostate change of the current patient during the treatment. Experimental results show that our proposed method can produce prostate segmentations more accurately from clinical CT images, than the other state-of-the-art segmentation algorithms.

Acknowledgments. This work was supported by NIH grant R01 CA140413 and the National Basic Research Program of China (973 Program) (No. 2010CB732505). The image data and expert contours used in this study were provided Dr. E.L. Chaney in the Department of Radiation Oncology, UNC-Chapel Hill.

References

1. American Cancer Society.: Cancer facts and figures 2006. In: American Cancer Society (ed.). American Cancer Society, Atlanta (2006)
2. Chen, S., Lovelock, D.M., Radke, R.J.: Segmenting the prostate and rectum in CT imagery using anatomical constraints. Med. Image Anal. 15, 1–11 (2011)
3. Feng, Q., Foskey, M., Chen, W., Shen, D.: Segmenting CT prostate images using population and patient-specific statistics for radiotherapy. Med. Phys. 37, 4121–4132 (2010)
4. Freedman, D., Radke, R.J., Zhang, T., Jeong, Y., Lovelock, D.M., Chen, G.T.: Model-based segmentation of medical imagery by matching distributions. IEEE Trans. Med. Imag. 24, 281–292 (2005)
5. Davis, B.C., Foskey, M., Rosenman, J., Goyal, L., Chang, S., Joshi, S.: Automatic segmentation of intra-treatment CT images for adaptive radiation therapy of the prostate. Med. Image Comput. Comput. Assist. Interv. 8, 442–450 (2005)
6. Wang, H., Dong, L., Lii, M.F., Lee, A.L., de Crevoisier, R., Mohan, R., Cox, J.D., Kuban, D.A., Cheung, R.: Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy. Int. J. Radiat. Oncol. Biol. Phys. 61, 725–735 (2005)
7. Ghosh, P., Mitchell, M.: Segmentation of medical images using a genetic algorithm. In: Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, pp. 1171–1178. ACM, Seattle (2006)
8. Song, Q., Wu, X., Liu, Y., Smith, M., Buatti, J., Sonka, M.: Optimal graph search segmentation using arc-weighted graph for simultaneous surface detection of bladder and prostate. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 827–835. Springer, Heidelberg (2009)
9. Zhan, Y., Shen, D.: Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method. IEEE Trans. Med. Imag. 25, 256–272 (2006)
10. Hupse, R., Karssemeijer, N.: Use of normal tissue context in computer-aided detection of masses in mammograms. IEEE Trans. Med. Imaging 28, 2033–2041 (2009)
11. Oliva, A., Torralba, A.: The role of context in object recognition. Trends Cogn. Sci. 11, 520–527 (2007)

12. Tu, Z., Bai, X.: 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)
13. Richard, N., Dojat, M., Garbay, C.: Distributed Markovian segmentation: Application to MR brain scans. Pattern Recogn. 40, 3467–3480 (2007)
14. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Transactions on Image Processing 10, 266–277 (2001)
15. Viola, P., Jones, M.J.: Robust Real-Time Face Detection. Int. J. Comput. Vision 57, 137–154 (2004)
16. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), vol. 1, pp. 886–893. IEEE Computer Society, Los Alamitos (2005)