Prostate Segmentation by Sparse Representation Based Classification

Yaozong Gao[1,2], **Shu Liao**[2], and **Dinggang Shen**[2]

[1]Department of Computer Science, University of North Carolina at Chapel Hill yzgao@cs.unc.edu

[2]Department of Radiology and BRIC, University of North Carolina at Chapel Hill liaoshu.cse@gmail.com, dgshen@med.unc.edu

Abstract

Accurate segmentation of prostate in CT images is important in image-guided radiotherapy. However, it is difficult to localize the prostate in CT images due to low image contrast, unpredicted motion and large appearance variations across different treatment days. To address these issues, we propose a sparse representation based classification method to accurately segment the prostate. The main contributions of this paper are: (1) A discriminant dictionary learning technique is proposed to overcome the limitation of the traditional Sparse Representation based Classifier (SRC). (2) Context features are incorporated into SRC to refine the prostate boundary in an iterative scheme. (3) A residue-based linear regression model is trained to increase the classification performance of SRC and extend it from hard classification to soft classification. To segment the prostate, the new treatment image is first rigidly aligned to the planning image space based on the pelvic bones. Then two sets of location-adaptive SRCs along two coordinate directions are applied on the aligned treatment image to produce a probability map, based on which all previously segmented images of the same patient are rigidly aligned onto the new treatment image and majority voting strategy is further adopted to finally segment the prostate in the new treatment image. The proposed method has been evaluated on a CT dataset consisting of 15 patients and 230 CT images. Promising results have been achieved.

1 Introduction

Prostate cancer is the second-leading cancer for American men. Currently one of the major treatment methods is external beam radiation therapy, which basically has two stages, namely *the planning stage* and *the treatment stage*. In the planning stage, *a planning image* is scanned from the patient and a dose plan is designed. During the treatment stage, a CT image is acquired at each treatment day for the same patient, which could be repeated for up to 40 times, each with a CT image acquired. The prostate in each treatment image needs to be accurately localized so that the dose plan made in the planning image can be adjusted to the current treatment image. Therefore, the success of external beam radiation therapy highly depends on the accurate localization of the prostate.

However, there are three main challenges to accurately segment the prostate. First, prostate boundary is of extremely low contrast with its surrounding tissues in the CT images as shown in Fig. 1. Second, prostate motion is unpredictable due to the uncertain existence of bowel gas in different treatment days. Third, the bowel gas can significantly alter the image appearance and makes it inconsistent across different treatment days as illustrated in Fig. 1(a) and 1(b). In order to address these challenges, many novel methods have been proposed these years. The first category of methods is deformable model [1, 2]. The second category is registration-based method [3, 4]. Recently, Li et al. [5] incorporated context features into prostate segmentation and achieved promising results.

On the other hand, Sparse Representation based Classifier (SRC) [6] has achieved the state-of-the-art results in face recognition. It represents a testing sample as a sparse linear combination with respect to an over-complete dictionary, which consists of training samples from all classes. Representation residue with respect to each class is used to determine the class label of a testing sample. However, the good performance of the traditional SRC depends on the assumption that the training samples in each class are distinct from those in other classes. In practice, especially in pixel-wise classification, different classes may include very similar samples. To overcome this limitation, we propose a discriminant dictionary learning technique to enhance the dissimilarity between classes. Moreover, context features are incorporated into SRC to refine the prostate boundary in an iterative scheme. Finally, a linear regression model is further trained to predict the class probability based on the representation residues. The proposed method has been evaluated on 230 CT images from 15 patients. The experimental results show that our method can achieve promising results and outperform other state-of-the-art prostate segmentation methods.

2 Methodology

2.1 Sparse Representation Based Classifier (SRC)

Given a dictionary $D \in \mathbb{R}^{n \times N}$ and a sample $\mathbf{y} \in \mathbb{R}^n$, sparse representation aims to find a sparse linear combination of dictionary elements in D for best Prostate Segmentation by Sparse Representation Based Classification representing $\mathbf{y}$. Mathematically, the problem can be formulated as the following minimization:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0, \text{ subject to} \|\mathbf{y} - D\mathbf{x}\|_2 \leq \varepsilon, \quad (1)$$

where $\mathbf{x}$ contains the sparse linear coefficients and is usually called as sparse code in the literature, and ε is the maximum allowable representation error. Although solving (1) is a NP-hard problem, the solution can be well approximated by many pursuit algorithms such as Basis Pursuit (BP) and Orthogonal Matching Pursuit (OMP) [7]. In consideration of both efficiency and performance, we use OMP to solve the sparse coding problem in this paper.

In the traditional SRC, the dictionary D is formed as a collection of training samples from all classes:

$$D = [D_1, D_2, \ldots, D_i, \ldots, D_K] = \left[\mathbf{d}_{1,1}, \mathbf{d}_{1,2}, \ldots, \mathbf{d}_{i,j}, \ldots, \mathbf{d}_{K,N_K} \right], \quad (2)$$

where D_i is a sub-dictionary of class i that only contains training samples from class i, $\mathbf{d}_{i,j}$ is the j-th training sample of class i, K is the number of classes and N_K is the number of training samples in class K. To classify a new sample $\mathbf{y}$, its sparse code $\mathbf{x}$ is first computed according to (1). Then the representation residual vector $\mathbf{r}_i$ with respect to class i is computed as:

$$\mathbf{r}_i = \mathbf{y} - D_i \mathbf{x}_i, \quad (3)$$

where $\mathbf{x}_i$ carries entries of $\mathbf{x}$ indexed by an index set, which contains indices of columns in D belonging to D_i. Finally, the new sample $\mathbf{y}$ is classified to the class with the minimum $\|\mathbf{r}_i\|_2$.

2.2 Discriminant Sub-dictionary Learning

The traditional SRC method works well when there are no similar elements between sub-dictionaries. However, in many cases this assumption doesn't hold. In order to overcome this limitation, we need to build discriminant sub-dictionaries whose elements are distinct from those in other sub-dictionaries. In this paper, we combine feature selection with dictionary

learning technique to learn a discriminant sub-dictionary for each class. Here only the case of two classes is illustrated since a voxel is either classified to object (prostate) or background in the case of prostate segmentation. But this idea can be readily extended to multi-class cases when combined with multi-class feature selection techniques.

Given both background and object training samples, we want to select the topmost discriminant features that can enlarge the dissimilarities between training samples of two classes. Feature ranking based on Fisher's Separation Criteria (FSC) [10] is adopted in this paper to select the most discriminant features while eliminate features that are similar in both classes.

After feature selection, background and object samples can be directly used to form two sub-dictionaries. However, in practice the number of samples is usually large. In consideration of both sparse coding efficiency and dictionary storage, we need to use dictionary learning technique to learn a compact representation of training samples. In this paper, we adopt K-means as a way to learn sub-dictionaries. Compared with many reconstruction-oriented dictionary learning methods such as K-SVD [8] which don't consider discriminability during dictionary optimization, K-means can identify the individual clusters of different classes and thus can better preserve the dissimilarity between background and object class during dictionary learning. Therefore, it is more suitable when combined with SRC in classification.

2.3 Boundary Refinement by Context Features

In order to accurately localize the prostate boundary, it is necessary to draw more background and object training samples near the prostate boundary. However, these background and object samples are quite similar even after feature selection basically for two reasons: First, these samples are spatially close and sometimes next to each other. Second, prostate boundary in CT images is of extremely low contrast. To the best of our knowledge, no effective features which can accurately localize the prostate boundary have been identified. Therefore, even after performing discriminant sub-dictionary learning strategy, the SRC method can still produce many classification errors along the prostate boundary, which results in a zigzag boundary, as shown in Fig. 2(a).

Motivated by [5], we incorporate context features into SRC and propose an iterative SRC classification scheme. For each voxel, its features include not only local features but also context features taken at context locations as illustrated in Fig. 1(c). Previous classification results at context locations are used as context probability features which help guide the boundary refinement in the next classification iteration. Assume no prior information is available, we start with an uniform probability map. These context features don't help in the first iteration since they are filtered out by feature selection. However, in the later classification iterations, when the probability map becomes clearer and clearer, more context features will be selected to guide the classification refinement. Usually after several iterations, the prostate boundary becomes more refined as shown in Fig. 2(c).

2.4 Prediction by Residue-Based Linear Regression Model

The traditional SRC compares residual norms of different classes to determine the class label of a testing sample. In such case, residues of different features are equally treated. Usually a voxel is represented by the combination of different types of features, the discriminabilities of individual features are different and their contributions to classification are also different. Therefore, equally weighting them in determining the class label limits the classification performance. Besides, the traditional SRC is a hard classification method, which only assigns class label to a new sample. In contrast, soft classification provides more

quantitative informaton, especially in the decision margin where the class membership is unclear.

Motivated by these observations, a residue-based linear regression model is trained to learn the contributions of different features in class probability prediction. For each training sample, its background residual vector $\mathbf{r}_0$ and object residual vector $\mathbf{r}_1$ are computed and stacked into a single vector $\mathbf{r}=\left[\mathbf{r}_0^{\mathrm{T}}\mathbf{r}_1^{\mathrm{T}}\right]^{\mathrm{T}}$, which is used together with its class label $l \in \{-1, 1\}$ to train a linear regression model $\mathbf{m} \in \mathbb{R}^{2t}$, where t is the number of selected features for each sample. For a new testing sample $\mathbf{y}_{new}$, its object (prostate) class probability is computed as:

$$p=g\left(\frac{\mathbf{m}^{\mathrm{T}}\mathbf{r}_{new}+1}{2}\right), \quad (4)$$

where $\mathbf{r}_{new}$ is the stacked residual vector of $\mathbf{y}_{new}$ and $g(\cdot)$ is defined as a piece-wise function that maps any value outside [0, 1] to its nearest boundary value in order to keep the predicted probability between [0, 1].

2.5 Iterative Prostate Segmentation by SRC

We believe patches repeat not only spatially but also longitudinally. Therefore, in prostate segmentation, patches in the new treatment image likely have appeared in the previous treatment images or the planning image. If we build two discriminant patch-based sub-dictionaries for prostate and background using previous images, for a new patch in the new treatment image, it tends to draw more supports from the respective sub-dictionary in the sparse representation. Based on the representation residues corresponding to each class, we can estimate class probability of the voxel associated with this patch.

Our segmentation method consists of two stages, namely *training stage* and *classification stage*. **In the training stage**, two sets of location-adaptive [5] SRCs along two coordinate directions are learned using previous images of the same patient, which take the variability of different prostate regions into account. For each location-adaptive SRC, it only draws training samples from slices that it is responsible for. Then, based on the training samples, discriminant features are selected, two discriminant sub-dictionaries are constructed, and a residue-based linear regression model is finally learned. All these three steps are used together to classify all training slices and the class probabilities after classification are used to update the corresponding context probability features of the training samples. After the training samples are updated, we can learn a new SRC for the next classification iteration. The process is repeated until a specified number of iterations have been reached.

In the classification stage, the middle slice along each of two coordinate directions needs to be manually specified by users in order to shift the learned SRCs to the new treatment image space for classification (Note that the automatic middle slice identification method will be developed in our future work). The classification results along two coordinate directions are fused to form a final probability map. After classification is done, all previously segmented prostate images of the same patient are rigidly aligned to the probability map of the new treatment image and then majority voting strategy is adopted to segment the prostate finally.

3 Experimental Results

Our dataset consists of 15 patients, each with more than 11 CT images, with total 230 CT images. The resolution of each CT image is 1mm×1mm×3mm. The expert manual segmentation results are available for each image to serve as the ground truth. We use the

first 3 images including the planning image to initialize our method. As more treatment images are collected, only the latest 5 images are used as training images, which account for the tissue appearance change under radiation treatment. Two sets of location-adaptive SRCs are placed along anterior-posterior (y) direction and superior-inferior (z) direction, respectively, because slices along these two directions contain richer context information (e.g., pelvic bones) than slices along lateral (x) direction.

Before any operation is applied in the training stage, all previous treatment images are rigidly aligned to the planning image based on the pelvic bone structures in order to remove the irrelevant whole-body motion. The same preprocessing is also applied to the new treatment image before classification.

For each voxel, its features include both local appearance features and context features. Context features have two types of features, namely context probability features and context appearance features. Context probability features have been introduced in the previous section. They are used to refine classification results and updated in each iteration. In the experiment we only use 3 classification iterations. Context appearance features are the same kinds of features as local appearance features, but taken at context locations. 9 dimensional Histogram of Oriented Gradient (HOG) [9] and 23 Haar features computed in a 21mm×21mm local window are used as appearance features in this paper.

The box-and-whisker plot of the DICE measures and centroid distances along three coordinate directions of our method are shown in Fig. 3. Fig. 4 visually compares the segmentation results using residual norm comparison and residue-based linear regression. It can be seen that in the beginning and ending slices where the prostate is relatively small and difficult to accurately localize, the proposed linear regression model performs better than the traditional residual norm comparison. Four existing state-of-the-art prostate segmentation methods [1–3, 5] are compared with our method. The mean and standard deviation of DICE measures of our method is 0.912 ± 0.044 based on Fig. 3, which is better than 0.820 ± 0.060 in [3], 0.893 ± 0.050 in [2] and 0.908 in [5]. The median DICE measure of our method is 0.918, which is also better than 0.840 in [3] and 0.906 in [2]. The median probability of detection and false alarm of our method are 0.913 and 0.072, respectively, which are better than 0.840 and 0.130 reported in [1], and 0.900 and 0.100 reported in [5]. Besides, we also compared the centroid distances. The mean centroid distances along lateral (x), anterior-posterior (y) and superior-inferior (z) direction of our method in Fig. 3 are 0.06 mm, −0.07 mm and 0.19 mm, respectively, which are much better than the respective centroid distances of −0.26 mm, 0.35 mm and 0.22 mm reported in [3], and comparable to the result of 0.18 mm, −0.02 mm and 0.57 mm in [5].

4 Conclusion

We have proposed a sparse representation based classification method for segmentation of prostate in CT images. Feature selection is combined with dictionary learning technique to learn two discriminant sub-dictionaries which overcome the limitation of the traditional SRC. Context features are further incorporated into SRC to refine the classification results (especially the prostate boundary) in an iterative scheme. A residue-based linear regression model is finally learned to increase the classification performance and extend the traditional SRC from hard classification to soft classification. Experimental results show that our proposed method can achieve more accurate prostate segmentation results than other state-of-the-art segmentation methods under comparison.

References

1. Chen S, Lovelock DM, Radke RJ. Segmenting the prostate and rectum in CT imagery using anatomical constraints. Med. Image Anal. 2011; 15:1–11. [PubMed: 20634121]

2. Feng Q, Foskey M, Chen W, Shen D. Segmenting CT prostate images using population and patient-specific statistics for radiotherapy. Med. Phys. 2010; 37:4121–4132. [PubMed: 20879572]

3. Davis, BC.; Foskey, M.; Rosenman, J.; Goyal, L.; Chang, S.; Joshi, S. Automatic Segmentation of Intra-treatment CT Images for Adaptive Radiation Therapy of the Prostate. In: Duncan, JS.; Gerig, G., editors. MICCAI 2005. LNCS. Vol. vol. 3749. Springer; Heidelberg: 2005. p. 442-450.

4. Liao, S.; Shen, D. A Learning Based Hierarchical Framework for Automatic Prostate Localization in CT Images. In: Madabhushi, A.; Dowling, J.; Huisman, H.; Barratt, D., editors. Prostate Cancer Imaging 2011. LNCS. Vol. vol. 6963. Springer; Heidelberg: 2011. p. 1-9.

5. Li, W.; Liao, S.; Feng, Q.; Chen, W.; Shen, D. Learning Image Context for Segmentation of Prostate in CT-Guided Radiotherapy. In: Fichtinger, G.; Martel, A.; Peters, T., editors. MICCAI 2011, Part III. LNCS. Vol. vol. 6893. Springer; Heidelberg: 2011. p. 570-578.

6. Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y. Robust Face Recognition via Sparse Representation. PAMI. 2009; 31:210–227.

7. Tropp JA. Greed is good: algorithmic results for sparse approximation. IEEE Transactions on Information Theory. 2004; 50:2231–2242.

8. Aharon M, Elad M, Bruckstein A. K-SVD: An Algorithm for Designing Over-complete Dictionaries for Sparse Representation. IEEE Transactions on Signal Processing. 2006; 54:4311–4322.

9. Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection. Computer Vision and Pattern Recognition. 2005; vol. 1:886–893.

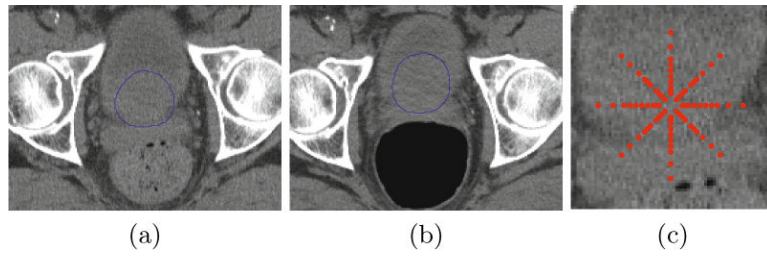10. Guyon I, Elisseeff A. An introduction to variable and feature selection. J. Mach. Learn. Res. 2003; 3:1157–1182.

Fig. 1.
(a) and (b) are two axial slices from different treatment images of the same patient. Blue contours are the prostate contours manually delineated by experts. (c) is an illustration of context locations of the center pixel by red points.
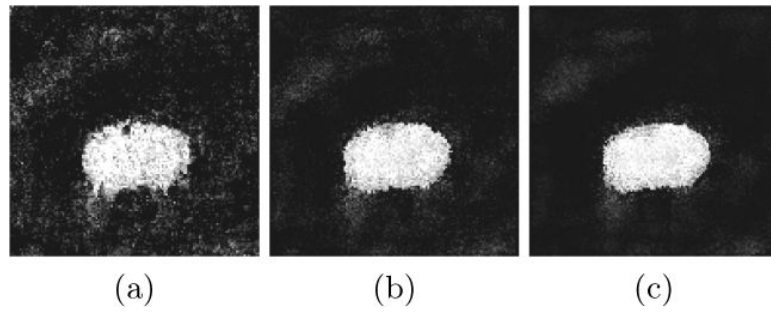
(a)   (b)   (c)

Fig. 2.
The first, second and third column represents the results of the first, second and third classification iteration, respectively
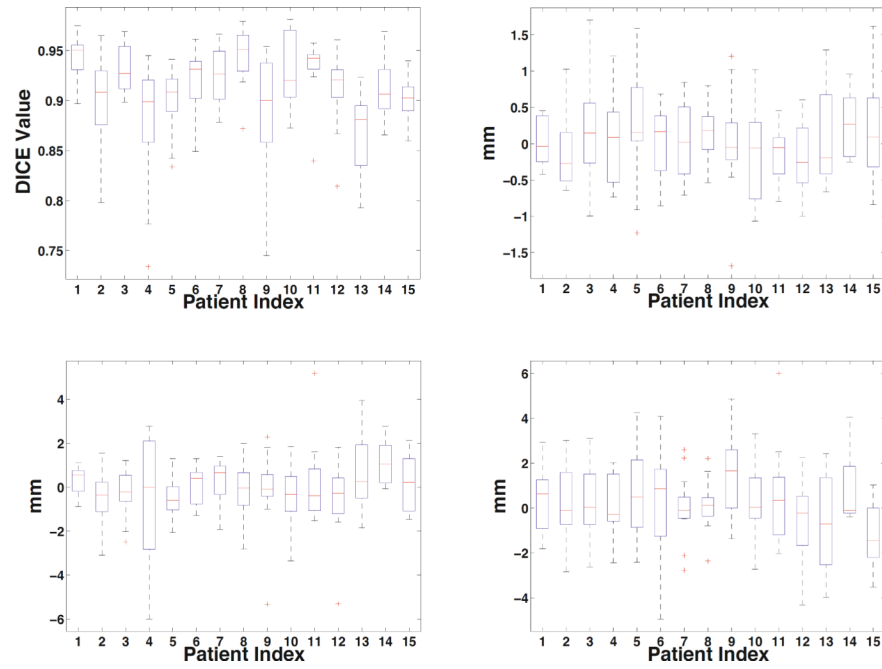
Fig. 3.
Left-top figure shows the DICE measures of our method. Right-top, left-bottom and right-bottom figures are centroid distances in lateral (x), anterior-posterior (y) and superior-inferior (z) directions, respectively.
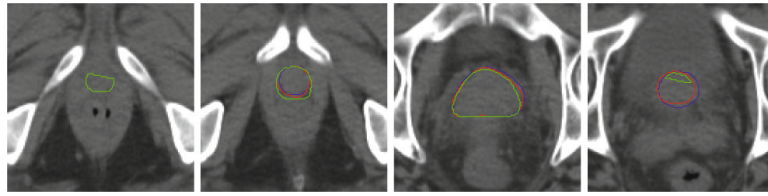
Fig. 4.
Comparison of the segmentation results between linear regression and residual norm comparison. Blue contours are the prostate boundaries manually delineated by experts. Red and green contours are the segmentation results of the proposed method with linear regression and residual norm comparison, respectively. This indicates that our proposed method with linear regression achieves better results, especially in the beginning and ending slices of the prostate.