Published in final edited form as:

IEEE Trans Biomed Eng. 2013 February ; 60(2): 479–488. doi:10.1109/TBME.2012.2228644.

Prostate Segmentation in MR Images Using Discriminant Boundary Features

Meijuan Yang [Fellow, IEEE],

Center for OPTical IMagery Analysis and Learning, State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, Shaanxi, China.

Xuelong Li [Fellow, IEEE],

Center for OPTical IMagery Analysis and Learning, State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, Shaanxi, China.

Baris Turkbey [Senior Member, IEEE],

National Institutes of Health, National Cancer Institute, Bethesda, MD 20892 USA.

Peter L. Choyke [Senior Member, IEEE], and

National Institutes of Health, National Cancer Institute, Bethesda, MD 20892 USA.

Pingkun Yan [Senior Member, IEEE]

Center for OPTical IMagery Analysis and Learning, State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, Shaanxi, China.

Abstract

Segmentation of the prostate in magnetic resonance image has become more in need for its assistance to diagnosis and surgical planning of prostate carcinoma. Due to the natural variability of anatomical structures, statistical shape model has been widely applied in medical image segmentation. Robust and distinctive local features are critical for statistical shape model to achieve accurate segmentation results. The scale invariant feature transformation (SIFT) has been employed to capture the information of the local patch surrounding the boundary. However, when SIFT feature being used for segmentation, the scale and variance are not specified with the location of the point of interest. To deal with it, the discriminant analysis in machine learning is introduced to measure the distinctiveness of the learned SIFT features for each landmark directly and to make the scale and variance adaptive to the locations. As the gray values and gradients vary significantly over the boundary of the prostate, separate appearance descriptors are built for each landmark and then optimized. After that, a two stage coarse-to-fine segmentation approach is carried out by incorporating the local shape variations. Finally, the experiments on prostate segmentation from MR image are conducted to verify the efficiency of the proposed algorithms.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier [10.1109/TBME.2012.2228644](https://doi.org/10.1109/TBME.2012.2228644)

Authors' photographs and biographies not available at the time of publication.

Index Terms:

Discriminant analysis; image feature; prostate segmentation; statistical shape model (SSM)

I. Introduction

SEGMENTATION of organs from medical images is a very important task, which is, however, challenging for the reason that the organs may have low contrast with their adjacent structures, resulting in misleading boundary detection. Among the existing segmentation methods, the statistical shape model (SSM) [1]–[3] has been successfully applied for medical image segmentation, since it can fit the learned shape to new image using the statistical analysis from the training data. It means that SSM can constrain the new shape to the space modeled by the training shapes [4]. For its modules (including the shape representation, shape model construction, image appearance model, and so on), much improvement has been made aiming at increasing the segmentation accuracy in certain applications [5]–[8]. In shape representation, the simplest and most generic method is the point distribution model (PDM), where sequences of points named as landmarks are used to represent shapes [1], [6].

In the deformation step of PDM, the image appearance model is used to generate image force to guide the shape to latch on the boundary of the organ [1], [5]. The feature employed to build the image appearance model should be able to characterize and detect the boundaries of the organ. The features to capture the saliency of the boundary drive the deformation of the shape to the true boundary. Subsequently, robustness and accuracy of the segmentation algorithm are significantly influenced by the feature descriptors. This highlights the need for designing the distinctive feature according to the property of the boundary of the organ and the imaging modality.

A. Image Features

The features used in SSM can be divided into two categories, namely, boundary-based feature and region-based feature. The former is effective in the case that the boundary of the anatomical structure is notable [5]. The latter shows robustness to noise [9]. The gradient information, belonging to the boundary-based feature, is widely used in SSM to define the boundary, but it can be trapped by spurious edges and noise points. Especially in medical images, the adjacent structures may have similar intensity distributions with the boundary of the target organ, which makes the gradient information become weak. Kirschner *et al.* [10] considered gradient in three orthogonal directions besides the intensity and the magnitude of the gradient. Another approach to enhance the distinctiveness is to add the statistical analysis, such as the mean, standard deviation, range, skewness, and kurtosis in the neighborhood of the point presented by Toth *et al.* [7]. The multifeature can alleviate the problem of the weak edges and boundaries. To deal with the inefficiency of the 1-D gradient information, Zheng *et al.* [11] extended 1-D profile to 2-D and extracted the gray values in a rectangle to describe the landmark to localize the facial feature. In this case, the target point, which is not located in the normal direction of current landmark, can be located. The 2-D

profile increases the accuracy by adding the information in the other dimension in 2-D image segmentation.

Combining the features of different categories is another alternative approach. For example, Feng *et al.* [12] employed the gradient feature and probability distribution feature to segment the prostate from CT images. At each vertex, the feature to be used depends on the consistency of the gradient feature at that vertex in the training procedure [13].

The existing features have three shortcomings when being employed for segmenting the prostate from MR images. First, the region inside the prostate is inhomogeneous and does not follow the pathognomonic texture (see Fig. 1). In another words, the region-based feature cannot provide some prior information and is not suitable for prostate segmentation from MR images. Second, when the image texture becomes complex and the boundary is not in accordance with the edges or ridges, the organ is difficult to be distinguished from the neighboring structure (also see Fig. 1). Third, the features have substantial influence on the shape model-based segmentation. For the previous reasons, a robust and distinctive feature to capture the information of the boundary is highly desired.

B. Assessment of Image Features

When designing the boundary features, the pattern of the gray values and gradients may vary over the boundary of the prostate. Accordingly, the parameters in the boundary feature significantly affecting the distinctiveness should be optimized. However, how to evaluate the distinctiveness still needs to be addressed. The traditional evaluation is achieved by using the final segmentation results. Nevertheless, it cannot evaluate the performance for each landmark individually. The other disadvantage is that it cannot evaluate the performance of the features directly. The accuracy of the segmentation is related to the integrated algorithm including several modules, such as the shape model and image appearance model. It is important to note that the shape needs to be smoothed by the internal energies and constrained by the shape model after the feature matching, where the contribution of image feature may be disguised. Thus, a new measurement is desired to directly evaluate the distinctiveness of the features locally.

C. Brief Outline of Our Method

In this paper, we propose a framework that includes both the boundary-based and patch-based features. For the entire region inside the prostate showing large variation, the information is captured in a local fashion, i.e., by a patch surrounding a boundary point [14]. The scale and variance are not specified with the location of the point of interest when SIFT feature is employed in segmentation. However, the scale and variance significantly affect the distinctiveness of the SIFT feature. To make the scale and the variance adaptive for each point of interest, the discriminant analysis in machine learning is introduced. The local and immediate assessment of the distinctiveness of the boundary feature is available simultaneously with the discriminant analysis. One doesn't have to wait after the segmentation is completed. The adaptive SIFT feature is denoted as scale and variance adaptive SIFT (SVA-SIFT). As the gray values and the gradients may vary over the entire boundary, separate image appearance descriptors, normal vector feature profile (NVFP), are

built based on SVA-SIFT for each landmark. The anatomical constraint for each landmark is incorporated by learning the training samples. In this paper, we also propose to use the learned image appearance and the special structure of the prostate to divide the entire shape into several segments.

The rest of the paper is organized as follows. In Section II, the details of the discriminant analysis in machine learning to define the distinctive function of the features are given, followed by extracting the SVA-SIFT feature and constructing the image appearance description. In Section III, the proposed local shape model is described based on the learned features. In Section IV, both the qualitative and quantitative experimental results are given, followed by the discussion. In Section V, the conclusion and future work are presented.

II. Extracting Descriptive Image Features

In this section, the method for image feature extraction and landmark appearance description is illustrated in detail. Tuning the parameters for the image features and choosing the optimized appearance descriptor for each landmark are achieved in the learning stage. They are both based on the proposed principles of the discriminant analysis, which is also presented in this section.

A. Distinctiveness of Features

When training image features, the designed descriptors of corresponding landmarks in different images should be as similar as possible. For example, Feng *et al.* [12] employed the consistency to describe the similarity of the features or descriptors of the corresponding landmarks, which are of the same anatomical location in different images. At the same time, the selected features can also help to differentiate the target landmark from the points in its neighborhood. For instance, Zhan *et al.* [15] used a redundant feature pool and selected the distinctive features for local boundary via a cascaded Adaboost classifier. The selected distinctive features are most powerful to distinguish the landmarks from the candidates along their normal directions. Two principles in discriminant analysis can be followed as the work in [16] to integrate the two requirements. The first principle requires the feature to capture the common characteristics of the similar texture, namely the learning principle. The second one is that the features of the objects of two classes should be as different as possible. The second principle can be denoted as distinctive principle. The principles can be formulated as follows.

Let F be the features or descriptors of all the landmarks and points in their neighborhood. Then, the two principles can be integrated as

$$S(F) = \frac{\sum_{t=1}^c n_t (m_t - m)(m_t - m)^T}{\sum_{t=1}^c \sum_{f \in F_t} (f - m_t)(f - m_t)^T} \quad (1)$$

where c is the number of points in the neighborhood of each landmark including the landmark itself. n_t equals to the number of training samples. F_t denotes the set of features for

the points in the neighborhood of the t th landmark from the training images. In (1), m_t and m are the mean vectors of F_t and F , respectively.

In the distinctiveness criterion (1), the numerator term represents the variance of the features between the landmark and its surroundings points. This variance should be large, following the distinctive principle. The denominator term denotes the variance of the features or descriptors of the corresponding points for each landmark. The smaller this variance is, the more similar the features or descriptors of corresponding points are, following the learning principle. Both of them are essential for describing the distinctiveness. The most distinctive feature is with the largest variance of the numerator term and the smallest variance of the denominator term. That is, the larger $S(F)$ is, the more distinctive the feature descriptors are.

B. SVA-SIFT

In medical images, object boundary is not always in accordance with edges, which makes it difficult to define the boundary feature. In addition, the texture in medical images is quite complex that the existing boundary features could not differentiate the target landmarks from the points with similar local information.

SIFT descriptor is desirable to be used in segmentation, for its ability to capture the scale and rotation invariant feature of the landmark. However, it faces challenges when SIFT descriptor is employed in the segmentation of medical images. One challenge is that the textures vary over the boundary of the anatomical structures. In addition, the scale and variance are not specified with the location of each landmark. The other challenge is that the assignment of the principal orientation of SIFT descriptor is affected easily by the strong edges and the existence of the weak boundary, as shown in Fig. 2. The misalignment of the principal orientations may degrade the learning ability of the feature training method. In this paper, the SVA-SIFT feature is proposed to make the scale and variance adaptive to the location for each landmark.

The SVA-SIFT feature, employed in this paper, is computed as follows.

1. The scale space $L(x, y, \sigma)$ is computed by smoothing the input image $I(x, y)$ with the Gaussian function $G(x, y, \sigma)$

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y). \quad (2)$$

When SIFT feature was proposed by Lowe [17], it was initially used for an object recognition and later for an image registration. In those applications, the point of interest is detected at a local extremum in Difference-of-Gaussian with the optimized attributes of image scale and variance. The SIFT feature is the most distinctive and robust when being computed at those locations. While the SIFT feature is applied to segmentation, it should be directly computed at the landmarks. As a result, the optimized scale and variance are not available. Additionally, the result of feature matching is related to the learning ability of the features. Therefore, the two parameters can be optimized by $S(F)$ in (1)

indicating the distinctiveness of the features, and are adaptive to the locations of the points of interest, forming the SVA-SIFT feature.

2. The landmarks representing the shape of a target organ are determined as the points of interest in SVA-SIFT descriptor.
3. The gradient magnitude $m(x, y)$ and the orientation $\theta(x, y)$ are computed in the scale space $L(x, y, \sigma)$ in (2). Then, they are used to assign the principal orientations for the points of interest

$$m(x, y) = ((L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2)^{1/2} \quad (3)$$

$$\theta(x, y) = \arctan \frac{(L(x+1, y) - L(x-1, y))}{(L(x, y+1) - L(x, y-1))}. \quad (4)$$

In this step, to enable the feature consistency of corresponding landmarks in different images, the main orientation of the feature descriptor is assigned as the normal direction of the contour at that point in SVA-SIFT computation.

4. The 128-D descriptor is computed as the original SIFT.

In SVA-SIFT feature, the image scale and variance need to be tuned. This is a special case when SIFT feature is used in segmentation, not in the recognition or registration [18], [19], where the two parameters are available during the detection of the points of interest. In our work, the adaptation to image scale and variance is achieved by using the distinctive function $\mathcal{S}(F)$ in (1). The scale and variance adaption are illustrated in Fig. 3. The statistical results of the distinctive function for SVA-SIFT, $\mathcal{S}(\text{SVA-SIFT})$, are provided at each landmark. It is shown that SVA-SIFT is more susceptible to variation of the image scale than variance.

To show the performance of SVA-SIFT feature-based matching, a Simulated Search [20] was performed. The aim is to detect the points on the boundary of the prostate from one MR image using the corresponding landmarks in another prostate image based on the boundary feature. Features extracted from the landmarks shown in Fig. 4(a) were used to search for matched points around the landmarks in (b) and (c). It can be seen that the image appearance descriptor using SVA-SIFT features in Fig. 4(b) is more distinctive than that using SIFT features shown in Fig. 4(c), since the yellow dots are closer to red crosses in (b) than in (c).

C. NVFP

Since the intensity values and gradients on the boundary of the prostate are similar with that of the neighboring regions of the prostate, especially in MR images with complex textures, it is difficult to detect the boundary of the prostate. The SVA-SIFT feature of each landmark includes the information of pixels in its neighborhood, which is still not enough to describe the neighboring pixels. To further differentiate the true boundary from the spurious edges, the regional image appearance descriptor is built along the normal direction.

In SSM, a set of landmarks are used to represent a shape. Let x_i be a 1-D vector, denoting the location of the i th landmark, then $X = (x_1, x_2, \dots, x_M)$ represents the shape of the organ, and M is the number of landmarks. The normal direction of x_i is n_i . In order to improve the learning ability of the descriptors, the features of the points that are inconsistent across images are removed from NVFP.

In general medical image segmentation, features of the same anatomical structure in different images with the same imaging protocol can be mapped to a similar range. However, for the prostate, the neighborhood region of a landmark outside of the boundary usually shows larger variation between different images, compared with that of a landmark inside of the prostate boundary. For example, the structure may be the bladder close to the base of the prostate or the rectum close to the bottom of the prostate. Bowel gas may exist in the rectum [21]. Those factors create inconsistency from one image to another. The inconsistent anatomical structure degrades the learning ability of the algorithms. Taking all the possible cases into consideration, two patterns of image appearance descriptors are proposed. The first is to truncate the features in the outside of the prostate according to the outside structures. This pattern is related to the cases, where the inside region is more homogeneous than the outside region. The second pattern is to truncate the features on both side of the prostate at the same time, assuming that both the inside and the outside of the organ are not homogeneous. In total, $(k-1)/2$ cases are designed to capture the local characteristic of the landmarks. In Fig. 5(a) and (b), the two patterns of NVFPt are shown. The algorithm for generating the truncated NVFP (NVFPt) features are provided as follows.

1. For each point x_i , k number of points are sampled from both sides along the normal direction to the contour. Those k points are distributed evenly in the outside and inside of the normal direction. k is selected as an odd number. The spacing between the adjacent points is δ pixels. The coordinate of the sample point x_i^j is determined by

$$x_i^j = x_i + j \cdot \delta \cdot n_i \quad (5)$$

where $-(k-1)/2 \leq j \leq (k-1)/2$. The parameters of k and δ will be chosen experimentally.

2. Compute the SVA-SIFT features for all k sample points, then concatenate them as $g(x_i) = (f_1, f_2, \dots, f_k)$. f_j to represent the NVFP feature of x_i^j . The length of the initial image appearance descriptor $g(x_i)$ is $128 \times k$.
3. Truncate the inhomogeneous structures from NVFP to get the truncated version as

$$\text{NVFPt}(x_i, l) = g(x_i, \text{loc}(l), \text{num}(l)) \quad (6)$$

where g is the initial image appearance descriptor and $\text{NVFPt}(x_i, l)$ is the l th truncated descriptor for x_i . $\text{loc}(l)$ and $\text{num}(l)$ are the corresponding number and locations of the sample points.

4. 4) Choose the optimized image appearance descriptor for each landmark. The selection of image appearance descriptors is based on the distinctiveness function $S(\text{NVFPt}(x_i, l))$. The NVFPt with the largest distinctiveness is selected as the optimized NVFPt for x_i

$$\arg \max_{1 \leq l \leq k} S(\text{NVFPt}(x_i, l)). \quad (7)$$

In Fig. 6, the number and the locations of the sample points building the image appearance descriptors are presented for each landmark. In Fig. 6(a), the SVA-SIFT features of the most points located outside the prostate close to the bottom are truncated, which validates the inconsistency between the prostate and the rectum.

III. Coarse-to-Fine Image Segmentation

In this section, a coarse-to-fine segmentation strategy is presented to segment the anatomical structures from medical images. The coarse-to-fine segmentation method is comprised of two phase, i.e., the global deformation and the local deformation. During the global deformation, the segmentation proceeds as in the classic active shape model (ASM). However, the shapes of the anatomical structures vary greatly from one patient to another. As a result, the excessive shape constraint is imposed on the shape of the anatomical structure being segmented when principal component analysis (PCA) is employed to capture the most important variations of the global shape. The shape is located around the anatomical structure, but some segment of the shape is not delineated accurately. For example, the top of the prostate is often segmented inaccurately.

A coarse shape of the structure is acquired from the global contour deformation

$$E = E_{\text{ext}} + E_{\text{gpca}} \quad (8)$$

where E_{ext} represents the image energy that is defined by the matching of the boundary feature. The first term E_{ext} is used to guide the point along the contour to a target position. The second term E_{gpca} is the global shape constraint imposed by PCA of the training samples. In the fine stage, we define a new energy function our model

$$E = E_{\text{ext}} + E_{\text{lpca}} + E_{\text{int}} \quad (9)$$

where E_{lpca} is used to preserve the delicate information about the local shape of the anatomical structure. The third term E_{int} is the shape smoothing energy, which is employed by dynamic programming. After testing on different values, the shape is divided into four segments to achieve better segmentation results.

Our work mainly focuses on the shape model and the image appearance model. The flowchart of the proposed method is shown in Fig. 7. In the image appearance model, we first aim to learn the most distinctive boundary feature for each landmark by SVA-SIFT. Then, taking the anatomical constraints of each landmark into consideration, the truncated NVFP is developed based on the number and location of the sample points, of which the SVA-SIFT is used to form the truncated NVFP. In selecting the most distinctive boundary feature and learning the truncated NVFP for each landmark, the discriminant analysis is introduced including both the in-class consistency and cross-class differences. In shape modeling, the delicate information about the local shape is derived from the training shapes as well as the global shape model. To segment a test image, the model initialization is achieved by the global shape model. Then, the learning of truncated NVFPs produces the image force to guide the shape deformation constrained by the global and local shape model. The shape deformation proceeds iteratively until it converges.

IV. Experimental Results

A. Evaluation Measures and Experimental Data

In our work, the prostate data of MR derives from 52 patients on a 3.0T Philips Gyroscan (Philips Medical Systems). Each patient has one axial scan and the MR protocol is TSE (T2-w) acquisition (TR = 8868.72 ms or 4434.36 ms, TE = 120 ms, angle = 90°, FOV=100 mm) with 3-mm axial slice thickness. The size of each image is 512×512 and the pixel resolution is $0.3 \text{ mm} \times 0.3 \text{ mm}$. To each image, the boundary was manually outlined by a clinical expert, which is provided as the ground truth. Given that the number of the medical images is generally small, the leave-one-out method is employed to show a statistical segmentation results on 52 images. To compare the performance of the proposed algorithm with that of other methods, the statistical segmentation results were used.

To evaluate the performance of the proposed algorithm, the manual segmentation result is often considered as the gold standard. Generally, the similarity or the distance between the automatic segmentation and the manual segmentation is used to compare the accuracy of different algorithms. Most evaluation measures are based on shape distance and area overlap.

In the shape distance-based metrics, the average surface distance (ASD) is used in this paper. Let p denote a point and $X(A)$ be a shape. The distance from p to $X(A)$ is defined as

$$d(p, X(A)) = \min_{p_A \in X(A)} \|p - p_A\| \quad (10)$$

where p_A is a landmark on the shape $X(A)$. Given that, ASD between $X(A)$ and $X(B)$ is defined as

$$ASD(A, B) = \frac{1}{|X(A)| + |X(B)|} \left(\sum_{p_A \in X(A)} d(p_A, X(B)) + \sum_{p_B \in X(B)} d(p_B, X(A)) \right). \quad (11)$$

ASD is computed symmetrically. The reason for ASD being selected from the available shape distance-based metrics is that the gross errors caused by the micro misalignment can be avoided. For example, the first landmark on one shape may correspond to the second landmark on the other shape, which may cause misalignment error. In this case, ASD can be more objective.

In the area overlap-based metrics, the dice similarity coefficient (DSC) is chosen. Let $S(A)$ and $S(B)$ represent the regions inside of the shapes of $X(A)$ and $X(B)$, respectively. DSC is computed as

$$DSC(S(A), S(B)) = \frac{2(S(A) \cap S(B))}{S(A) + S(B)}. \quad (12)$$

B. Parameter Settings

1) Parameters in the Shape Model:

M: 64, the number of points representing the shape.

T: 6, maximum displacement of the point on each side along the normal direction at each iteration.

The aforesaid parameters were set by balancing the segmentation accuracy and the computational time. The larger M and T may give better segmentation results but will take longer time to compute. These two parameters are set globally for all the segmentation tasks, which do not depend on training data.

2) Parameters in SVA-SIFT Features: The image scale is crucial for the distinctiveness of the SVA-SIFT local descriptor. In this paper, the image scales are 1024×1024 (upsampled), 512×512 (original), 256×256 (downsampled once), and 128×128 (downsampled twice). The distinctiveness of the SVA-SIFT feature can be assessed by the discriminant analysis for each landmark. In traditional methods, the performance of the features is assessed as an integral part of the segmentation algorithm. On one hand, with the discriminant analysis of the features for each landmark, the assessment can be done for each landmark locally. On the other hand, the performance of the features can be assessed independently on other factors, such as the searching algorithms and the shape model. The prostate is segmented in these four scales. The accuracy of the segmentation validates the performance of the proposed assessing measure.

Fig. 8 shows the segmentation results with the four image scales. It can be seen that the shape from the automatic segmentation is closest to the ground truth when the image scale is 1024×1024 (upsampled). The SVA-SIFT feature for one landmark includes the information

of 16 neighborhood in the image with the current scale. That means, when the image is upsampled, a 16×16 patch contains information not only from the original 8×8 patch but also the information from the interpolation.

In theory, the more distinctive the boundary feature is, the more accurate the segmentation result is. The distinctiveness of the SVA-SIFT feature in Fig. 3(a) is in accordance with the segmentation results in Fig. 8 with four different image scales. The measurement of assessing the distinctiveness of the boundary feature can be validated via this experiment.

3) Parameters in NVFP: In building NVFPt, the initial feature is truncated to adapt to the anatomical structure for each landmark. The two parameters in the initial feature k and δ are discussed in this experiment. The values of k and δ are shown in Fig. 9. The values of k and δ represent the number of sample points and the spacing. The two parameters also can indicate the range, within which the information is captured along the normal direction. Within the proper range, the least redundant information is included in the initial NVFP. The values of k and δ are determined according to the segmentation results. Each bin of the histogram in Fig. 9 represents the accuracy of the segmentation with one set of parameters. In Fig. 9, the shape from the automatic segmentation is closest to the ground truth when the values of k and δ are set to 9 and 1, respectively. In fact, the accuracy of the segmentation with $(k, \delta) = (7, 1)$ is near to that with $(k, \delta) = (9, 1)$. In order to save the computation time on the premise of not decreasing the accuracy, the values of k and δ are set to 7 and 1, respectively. Some segmentation results are shown in Fig. 10.

C. Validation of the Proposed Method

Considering the common case that the number of available training images is small in practice, the method of leave-one-out was used to present the statistical segmentation results. To segment one image, the other 51 prostate images were employed to train the shape model and the appearance model. The experiments are conducted with the classic ASM, ASM with multifeature (ASM with MF), active appearance model (AAM), the truncated NVFP, and truncated NVFP with local PCA. Then, the segmentation results are visualized and quantified to evaluate the efficiency of the proposed feature. For convenience, ASM NVFPt and ASMs NVFPt are the short names of the last two methods, respectively. The adopted methods were initialized by using the mean shape of the training shapes. Since the ASMs-based method are prone to be converged to the local minima, mean shape is translated so that the initial shape will not be too far away to the true boundary. The five methods share the same initial shapes for fair comparison.

1) Distinctiveness of the Proposed Feature: In Fig. 13, some segmentation results are visualized. In the left two columns, the segmentation of all the five methods are close to the ground truth. However, for some segments of the shape, the proposed method with NVFPt are more prone to differentiate the target points from other candidates located on the direction being perpendicular to the boundary. The reason is that the distinctive local image appearance descriptors steer the shape deformation, which is the similarity transformation of the mean shape, to the true boundary. In the right two columns, it can be seen that the proposed image appearance model captures the similarity transformation parameters leading

to better segmentation results compared to the other methods. Consequently, the mean quantitative segmentation result of the ASM with the proposed local image appearance descriptor is 1.35 mm, which is more accurate than 1.68 and 2.00 mm of AAM and ASM, respectively. The 2-D ASM incorporating multifeature in [7] is denoted by ASM with MF. The DSC of ASM with MF and ASM is 92.5 and 91.5, respectively. The segmentation results indicate that multifeature is more distinctive and robust than the 1-D gradient profile ($p < 0.01$).

2) Redundancy of the Image Appearance: In improving the distinctiveness of the image appearance descriptor, a general way is to add several features of different categories (i.e., combining the region-based and boundary-based features). Sequently, a large feature vector can be achieved. Another way is to expand the region in which the image appearance descriptor is computed (i.e., expanding the 1-D profile to 2-D profile). The two techniques can increase the distinctiveness of the image appearance descriptor to some extent. However, the redundant information is introduced inevitably. How to remove the redundant information is worth researching. k is the number of sample points to build the initial profile NVFP and the large value indicates that more information is included in the initial NVFP. However, the accuracy of the segmentation results are not always increasing as k increases.

3) Coarse-to-Fine Segmentation: In Table 1, it can be seen that there is a little improvement in the accuracy, comparing the ASM NVFPt and ASMs NVFPt. A paired t-test on the ASD values is performed showing that the ASDs for the ASM NVFPt are significantly ($p < 0.01$) greater than the ASDs for ASMs NVFPt.

4) Outlier: From Fig. 11(a) and (b), an outlier exists in the statistical analysis of the segmentation results. The outlier image is shown in Fig. 12. It can be seen that the boundary shows low contrast compared with the neighboring regions. In Fig. 12(b), the ground truth of the outlier and the mean shape derived from training samples are both visualized in the outlier image. The ground truth shows large variation compared with the mean shape. The general shape of the prostate is like the chestnut. Even if the proposed feature can detect the boundary points, the mean shape is hard to be transformed to a shape close to the ground truth.

V. Conclusion

In this paper, we propose a distinctive image appearance descriptor to describe the local regions surrounding landmarks and the proposed appearance descriptor is incorporated in the ASM to segment the prostate from T2w MR images. The SVA-SIFT feature is employed to build the image appearance description of NVFP. In order to make the image appearance descriptor more distinctive, the SIFT feature and NVFP are optimized. During the optimization, a measurement is introduced to assess the distinctiveness of the features. The measurement integrates two principles of learning ability of the objects of the same class and the distinctive ability of the objects of different objects into the distinctive function. Compared to the traditional assessing measurement, the proposed measurement can be implemented for each landmark directly in the training stage, not after the complete segmentation. The NVFPt are designed according to anatomical constraints for each

landmark, removing the redundancy information at the same time. The coarse-to-fine strategy is employed with local PCA of the segments of the global shapes, improving the segmentation accuracy. The experimental results show the distinctiveness of the proposed image appearance descriptor and the efficiency of the proposed measurement. In our future work, the proposed method will be extended to 3-D segmentation of the medical images.

Acknowledgments

This work was supported in part by the National Basic Research Program of China (973 Program) under Grant 2012CB719905, the National Natural Science Foundation of China under Grant 61105012 and Grant 61172142, and the Open Project Program of the State Key Lab of CAD&CG under Grant A1116, Zhejiang University.

References

- [1]. Cootes TF, Taylor CJ, Cooper DH, and Graham J, "Active shape models—their training and application," *Comput. Vis. Imag. Understand.*, vol. 61, no. 1, pp. 38–59, Jan. 1995.
- [2]. Shen D, Zhan Y, and Davatzikos C, "Segmentation of prostate boundaries from ultrasound images using statistical shape model," *IEEE Trans. Med. Imag.*, vol. 22, no. 4, pp. 539–551, Apr. 2003.
- [3]. Zhang W, Yan P, and Li X, "Estimating patient-specific shape prior for medical image segmentation," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2011, pp. 1451–1454.
- [4]. van Ginneken B, Frangi AF, Staal J, ter Haar Romeny BM, and Viergever MA, "Active shape model segmentation with optimal features," *IEEE Trans. Med. Imag.*, vol. 21, no. 8, pp. 924–933, Aug. 2002.
- [5]. Yan P, Xu S, Turkbey B, and Kruecker J, "Discrete deformable model guided by partial active shape model for TRUS image segmentation," *IEEE Trans. Bio. Eng.*, vol. 57, no. 5, pp. 1158–1166, 5 2010.
- [6]. Cootes TF, Edwards GJ, and Taylor CJ, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.
- [7]. Toth R, Tiwari P, Rosen M, Reed G, Kurhanewicz J, Kalyanpur A, Pungavkar S, and Madabhushi A, "A magnetic resonance spectroscopy driven initialization scheme for active shape model based prostate segmentation," *Med. Imag. Anal.*, vol. 15, no. 2, pp. 214–225, 2011.
- [8]. Yang M, Yuan Y, Li X, and Yan P, "Medical image segmentation using descriptive image features," in *Proc. Brit. Mach. Vision Conf.*, 2011, pp. 94.1–94.11.
- [9]. Stough J, Broadhurst RE, Pizer SM, and Chaney EL, "Regional appearance in deformable model segmentation," in *Proc. 20th Int. Conf. Inf. Process. Med. Imaging*, 2007, pp. 532–543.
- [10]. Kirschner M, Becker M, and Wesarg S, "3d active shape model segmentation with nonlinear shape priors," in *Proc. Int. Conf. Med. Med. Image Comput. Comput.-Assisted Intervention Conf.*, 2011, vol. 2, pp. 492–499.
- [11]. Zheng Z, Jiong J, Chunjiang D, Liu X, and Yang J, "Facial feature localization based on an improved active shape model," *Inf. Sci.*, vol. 178, no. 9, pp. 2215–2223, 2008.
- [12]. Feng Q, Foskey M, Chen W, and Shen D, "Segmenting CT prostate images using population and patient-specific statistics for radiotherapy," *Med. Phys.*, vol. 37, no. 8, pp. 4121–4132, 2010. [PubMed: 20879572]
- [13]. Georgescu B, Zhou XS, Comaniciu D, and Gupta A, "Database-guided segmentation of anatomical structures with complex appearance," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Image Understanding*, 2005, vol. 2, pp. 429–436.
- [14]. Mikolajczyk K and Schmid C, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005. [PubMed: 16237996]
- [15]. Zhan Y, Dewan M, and Zhou XS, "Cross modality deformable segmentation using hierarchical clustering and learning," in *Proc. Int. Conf. Med. Med. Image Comput. Comput.-Assisted Intervention Conf.*, 2009, vol. 1, pp. 1033–1041.
- [16]. Yang M, Zhang L, Feng X, and Zhang D, "Fisher discrimination dictionary learning for sparse representation," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 543–550.

- [17]. Lowe DG, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [18]. Scovanner P, Ali S, and Shah M, "A 3-dimensional sift descriptor and its application to action recognition," in *Proc. ACM 15th Int. Conf. Multimedia*, 2007, pp. 357–360.
- [19]. Urschler M, Bauer J, Ditt H, and Bischof H, "Sift and shape context for feature-based nonlinear registration of thoracic CT images," in *Proc. Second ECCV Int. Conf. Comput. Vis. Approaches Med. Image Anal.*, 2006, pp. 73–84.
- [20]. Peters J, Ecabert O, Meyer C, Kneser R, and Weese J, "Optimizing boundary detection via simulated search with applications to multi-modal heart segmentation," *Med. Imag. Anal.*, vol. 14, no. 1, pp. 70–84, 2010.
- [21]. Davis BC, Foskey M, Rosenman JG, Goyal L, Chang S, and Joshi S, "Automatic segmentation of intra-treatment CT images for adaptive radiation therapy of the prostate," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention Conf.*, 2005, pp. 442–450.

Fig. 1. Texture of the prostate image. The major challenges are caused by the existence of weak boundary (a) and spurious boundary (b).

Fig. 2.

Principal orientations of SIFT features. The points inside the rectangle in (a) and (b) are at corresponding locations. However, the computed principal orientations are susceptible to the strong edge surrounding those points. The dissimilarity of the two SIFT features degrades the feature learning ability.

Fig. 3. Distinctive function $S(F)$ for SVA-SIFT feature with (a) different image scales and (b) different variances.

Fig. 4. Correspondence detection using SVA-SIFT feature and SIFT feature in two different prostate images. The manually annotated landmarks are denoted by the red crosses. The matched points are displayed in yellow dots.

Fig. 5. Two patterns of NVFPt. (a) Equal number of inside and outside feature points, and (b) more number of feature points inside than outside.

Fig. 6.
Number and locations of the sample points for each landmark after being optimized for NVFPt.

Fig. 7.
Flowchart of the proposed method.

Fig. 8.

Segmentation results under five image scales (from left to right): optimized scale space (ASD = 1.34 mm), 1024 1024 (ASD = 1.37 mm), 512 × 512 (ASD = 1.49 mm), 256 × 256 (ASD = 1.55 mm), and 128 × 128 (ASD = 1.90 mm). The red dashed contour denotes the yellow solid contour is the ground truth.

Fig. 9. Segmentation performance with different parameter settings.

Fig. 10.

Segmentation results with different parameter settings. The red dashed contour denotes the automatic segmentation result and the yellow solid contour is the ground truth. The values of parameters k and δ are 7 and 1 in (a), 1 and 1 in (b), 9 and 3 in (c), and 9 and 6 in (d).

Fig. 11. Box-and-whisker diagrams showing the statistical segmentation results of five methods measured in ASD and DSC, respectively.

Fig. 12.
Outlier indicated by the statistical analysis of the segmentation results.

Fig. 13.

Segmentation results obtained by the five different methods (from the top row to the bottom row): ASM, ASM with MF, AAM, ASM with NVFPt, and ASMt with NVFPt. The red dashed contour is the result of the automatic segmentation and the yellow solid contour denotes the ground truth.

TABLE 1

Accuracy of Prostate Segmentation for Different Algorithms

Algorithm	Measures	Mean $\pm$ std	Min	Max	Med
ASM	ASD(mm)	2.01 $\pm$ 0.70	0.97	4.12	1.87
	DSC(%)	91.5 $\pm$ 3.3	79.4	96.0	92.4
ASM with MF	ASD(mm)	2.04 $\pm$ 0.65	1.17	4.19	1.97
	DSC(%)	92.5 $\pm$ 2.9	82.0	96.3	93.1
AAM	ASD(mm)	1.68 $\pm$ 0.45	0.84	2.89	1.55
	DSC(%)	93.2 $\pm$ 2.5	85.4	97.9	93.3
ASM with NVFPt	ASD(mm)	1.35 $\pm$ 0.42	0.58	2.74	1.37
	DSC(%)	94.7 $\pm$ 3.2	85.9	99.4	95.5
ASMs with NVFP	ASD(mm)	1.32$\pm$0.42	0.58	2.74	1.33
	DSC(%)	94.7$\pm$2.2	86.1	97.8	95.1