

Medical Image Segmentation Using Descriptive Image Features

Meijuan Yang
meijuan.yang@opt.ac.cn

Yuan Yuan
yuanyuan@opt.ac.cn

Xuelong Li
xuelong_li@opt.ac.cn

Pingkun Yan
pingkun.yan@opt.ac.cn

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

Abstract

Segmentation of medical images is an important component for diagnosis and treatment of diseases using medical imaging technologies. However, automated accurate medical image segmentation is still a challenge due to the difficulties in finding a robust feature descriptor to describe the object boundaries in medical images. In this paper, a new normal vector feature profile (NVFP) is proposed to describe the local image information of a contour point by concatenating a series of local region descriptors along the normal direction at that point. To avoid trapping by false boundaries caused by non-boundary image features, a modified scale invariant feature transform (SIFT) descriptor is developed. The number and locations of sample points for building NVFP are determined for each contour point, which are constrained by the neighboring anatomical structures and the statistical consistency of the training features. NVFP is incorporated into a model based method for image segmentation. The performance of our proposed method was demonstrated by segmenting prostate MR images. The segmentation results indicated that our method can achieve better performance compared with other existing methods.

1 Introduction

Segmentation of medical images plays an important role in diagnosis and therapy of diseases. It can provide quantitative pathological information about diseases and help physicians to make decisions. Due to the inefficiency and the inter- and intra-expert variability residing in manual segmentation, accurate and robust automatic segmentation is highly desired. However, fully automatic segmentation of a structure of interest, is still a very challenging task because of the shape and appearance variations of the target structure in different images.

Deformable models have been proven to be a promising approach to segment anatomic structures in medical images [1, 2, 3]. In deformable model based methods, the definition of boundary feature is critical for the segmentation performance. The statistics of the local grey-level appearance in active shape model (ASM) [4], which is a one-dimensional normalized

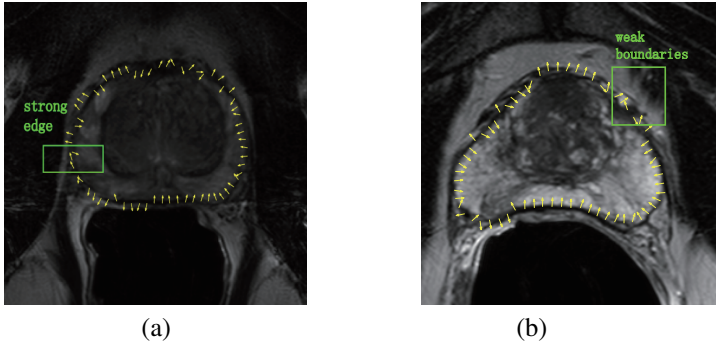


Figure 1: The main orientations of some points are affected by the strong edges in their neighboring region of the true boundary (a) and weak boundaries (b).

first order derivative profile, is widely used. Nevertheless, gradient-based feature is not adequate to characterize the weak and ambiguous boundaries in medical images due to low signal-to-noise ratio, poor image contrast, and the boundary complexity. A number of image appearance models have been proposed to drive the contour towards the boundary of the target. Toth *et al.* [5] employed a multi-feature appearance model incorporating the mean, standard deviation, range, skewness, and kurtosis of intensity values in the vicinity of each contour point to drive the edge detection. Feng *et al.* [6] presented an image appearance model consisting of gradient feature and probability distribution function feature. During the optimization process, one of the two features is selected for each vertex to guide contour deformation. The selection depends on the consistency of the gradient features over the training images. Chen *et al.* [7] proposed to learn a distribution of the intensity histograms in the local neighborhoods of the delineation contours from the training data. The multi-feature [5], the probability distribution function feature [6], and the distribution feature [7] are all region-base features, which are more robust to noise but less accurate than gradient-based features. Therefore, it is desirable to find a robust feature descriptor to describe the object boundaries in medical images.

Recently, Shi *et al.* [8] used the combined SIFT [9] features of the points selected evenly from both sides of the deformable contour to segment the lung fields from X-ray images. Their algorithm performs well in segmenting targets with relatively simple structures. However, the performance degrades when the image texture becomes complex. One of the reasons is that the main orientations of the SIFT features are prone to be affected by the strong edges in the neighboring regions of the true boundaries as shown in Fig. 1(a). The misalignment of the main orientations across different training images may degrade the learning ability of the feature training method. Another reason is that some segments of the boundaries are very weak (see Fig. 1(b)). The weakness makes it difficult to compute the exact main orientations in SIFT features.

In addition, the surrounding anatomic structures of different contour parts can vary significantly. If the same feature setting is employed for each contour point, this inhomogeneity will reduce the descriptive ability of the extracted features when being trained together. In ASM based image segmentation, the boundary point is located by searching along the normal direction to find a point, whose feature is most similar to the training features of corresponding contour points. If the huge variation is not eliminated from the training features, it will

be hard to distinguish the desired point from other candidate points. Therefore, eliminating the feature variation according to the anatomical structures surrounding each contour point is a very meaningful task.

In this paper, a new method of defining distinctive image features is proposed for medical image segmentation. Instead of using the general gradient-based and region-based features, we propose to employ a new SIFT based feature to describe the local image appearance of the points along the normal vectors, due to its distinctiveness on appearance description. However, strong edges surrounding the true boundaries and the existence of weak boundaries affect the main orientations when the image texture becomes complex. The misalignment of the main orientations across different training images may degrade the learning ability of the feature training method. To overcome this problem, we propose to modify the main orientations of the SIFT features and combine them to form distinctive image features for medical image segmentation. The obtained feature descriptor is coined as normal vector feature profile (NVFP). In addition, the surrounding anatomic structures of different contour parts can vary significantly. To eliminate the feature variation, proper number and locations of sample points for building NVFP should be determined for each contour point. The determination for each contour point is based on the anatomical constraints and statistical consistency of the training features. After NVFP is trained, it is incorporated into active shape model to improve the accuracy and robustness of segmentation.

The rest of this paper is organized as follows. A detailed description about NVFP is given in Section 2. The image segmentation model using NVFP is explained in Section 3. Experimental results are shown to demonstrate the performance of the proposed method on prostate MR images in Section 4. The conclusion is presented in Section 5.

2 Extracting Descriptive Image Features

In this section, we describe how the new feature of NVFP is computed. As we mentioned earlier, the modified SIFT feature is used in NVFP to obtain the saliency of the image information. Thus, in this section, we start with the modified SIFT feature.

2.1 The Modified SIFT Feature

In medical image segmentation, strong edges surrounding the true boundaries and the existence of weak boundaries affect the main orientations. Consequently, the modified SIFT feature is used to solve the two problems.

First, a difference-of-Gaussian pyramid D is computed to get the gradient magnitude and orientation for each point of the image. $D(x, y, \sigma)$ is computed by

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma), \end{aligned} \quad (1)$$

where $D(x, y, \sigma)$ is a variable-scale Gaussian function. $I(x, y)$ is an input image. $L(x, y, \sigma)$ is the scale space of an image, which is produced by the convolution of $G(x, y, \sigma)$ with $I(x, y)$.

Second, the key point of SIFT is located. In our work, we aim to segment the anatomical structures in medical images. Thus, the contour points of the structures of interest are selected to be the key points.

Third, the main orientation of SIFT is modified. In the original SIFT, the main orientation is the peak of the orientation histogram, which is obtained from the gradient orientations in

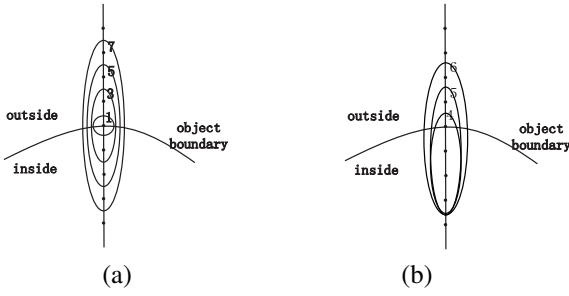


Figure 2: Two patterns of NVFPs are given: (a) appearances of points have low similarity both inside and outside the object, and (b) the appearance of the pixels inside the object is similar. Each ellipse stands for one case in our new profile. The points within the ellipse represent the locations of the sample points and the number marked on the ellipse represents the number of the sample points.

the neighborhood of the key point. However, it is prone to be disturbed by the strong edges in neighboring region of the true boundaries. In order to avoid ambiguities and improve the coherence in different images, we propose to use the normal directions of the contour at the key points for the main orientations, which achieves invariance to image rotation.

Lastly, the key point descriptor as in the original SIFT is extracted.

2.2 Normal Vector Feature Profile

In deformable models, the point distribution model (PDM) is employed to represent the contour. For each contour point, its distinctiveness is represented by NVFP, which is built by the modified SIFT feature. Let x_i be the coordinates of the i th contour point and n_i is the corresponding normal direction. The next three steps illustrate how NVFP is obtained.

1) For a contour point x_i , sample k points using a fixed spacing δ evenly from both sides along the normal direction. The location of j th pixel is

$$x_i^j = x_i + j \cdot \delta \cdot n_i. \quad (2)$$

The values of k and δ will be discussed in the experiments.

2) Let f_j denote the modified SIFT feature of x_i^j , then $g(x_i) = (f_1, f_2, \dots, f_k)$, which forms a feature vector. Since the SIFT feature consists of 128 elements, the length of a NVFP feature vector is $128 \times k$.

3) Eliminate dissimilar features of the inhomogeneous surroundings from NVFP.

According to the constraints of the surrounding anatomic structures, we design two patterns of NVFPs. In each pattern, there are several cases for x_i . In Fig. 2, each ellipse stands for one case in our new profile. The points within the ellipse represent the locations of the sample points and the number marked on the ellipse represents the number of the sample points. First, appearances of some points have low similarity both inside and outside the object in different images. For this pattern, points of equal number are sampled for the interior and exterior when NVFP is trained. In most cases, the appearance of the pixels inside the object is similar. However, the pixels outside of the object boundary can have quite different appearances. Therefore, the second pattern is to eliminate the inconsistent features caused by the different exterior structures. In general, there is no priori structure constraint to acquire

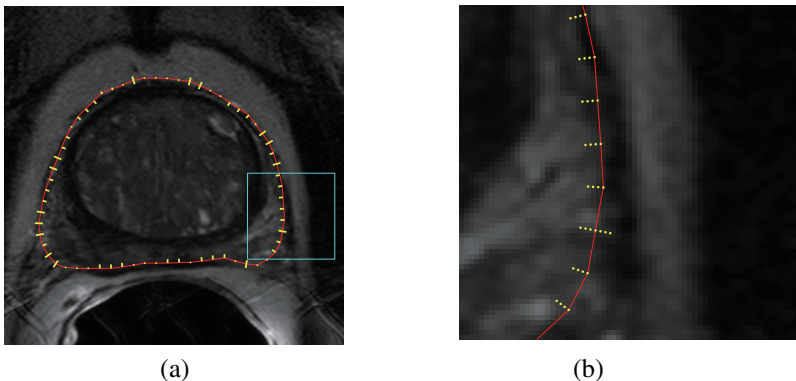


Figure 3: The number and locations of the sample points for the contour points are optimized when the NVFPs are built.

the pattern for each contour point as well as the number and locations of the sample points. Therefore, both of the two patterns are taken into account for each contour point.

Each NVFP is specified by the number and locations of sample points, which is described in detail in Fig. 2. Then, we compute the consistency for each NVFP:

$$C_i = \frac{L}{\sum_{l=1}^L \sigma(l)}, \quad (3)$$

where $L \leq k$ and $\sigma(l)$ is the standard deviation of the feature of l th sample point in the profile. Obviously, C_i measures the similarity of the features of the corresponding contour point in different images. In consequence, the consistency C_i is considered as the criterion to choose the number and locations of sample points when NVFP is built for each contour point.

As illustrated in Fig. 3, the number and locations of sample points for each contour point can be different due to the differences of their surrounding structures. Taking the prostate region at the bottom right as an example, features of the points outside of prostate boundary are not included in the training features, since the appearances of these points are strongly influenced by the varying edges close to the true boundary.

The results of detecting the corresponding points of the prostate boundary from MR images are shown in Fig. 4. In this detection process, we perform a Simulated Search [10]. Fig. 4(a) shows the prostate boundary denoted by the red crosses. The corresponding boundary of another prostate image (Fig. 4(b) and Fig. 4(c)) is detected by the similarity of the features between the candidate contour points (Fig. 4(b) and Fig. 4(c)) and the true contour points in Fig. 4(a). The detection results are represented by yellow dots and the contour points of manual segmentation are represented by the red crosses in Fig. 4(b) and Fig. 4(c). It can be seen that NVFP using the modified SIFT features in Fig. 4(b) is more distinctive than NVFP using the original SIFT features shown in Fig. 4(c). In the detection, we aim to validate the distinctiveness of the modified SIFT feature, so the number and locations of sample points for NVFP are the same for all contour points for convenience.

Two prostate segmentation results from the same image are given in Fig. 5. Proper number and locations of sample points for building NVFP for each contour point are used in Fig. 5(a) and the same are used in Fig. 5(b). It can be seen that our method can solve the inhomogeneous problem caused by dissimilar features of gas and bony structures in prostate image.

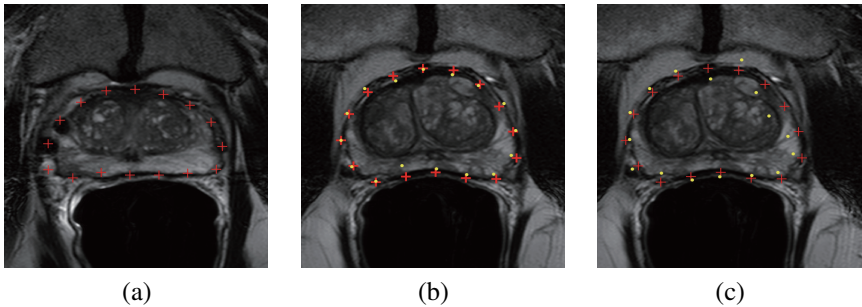


Figure 4: Distinctiveness of NVFP using the modified SIFT feature for correspondence detection in two different prostate images is shown. Manual segmentations are denoted by the red crosses. By comparing the NVFPs of the contour points in (a) and the candidate contour points in (b) and (c), the results of matching two prostate boundary are shown in (b) and (c) by yellow dots. Original SIFT features is used in NVFP in (c) while the modified SIFT is used in NVFP in (b). In the detection, we aim to validate the distinctiveness of the modified SIFT feature, so the number and locations of sample points for NVFP are the same for all contour points.

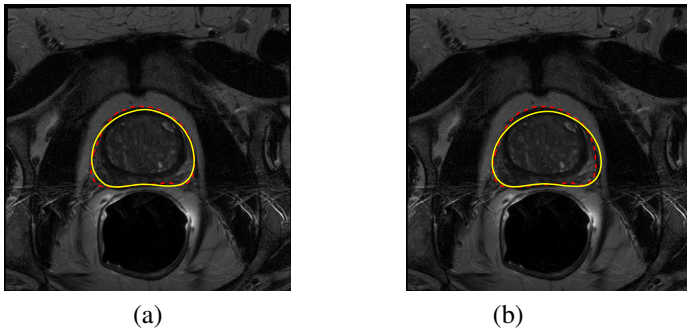


Figure 5: Two prostate segmentation results from the same image are given. The number and locations of the sample points for the profile are optimized for each contour point in (a), while they are the same for all contour points in (b).

3 Image Segmentation Model

In this section, a new energy function is defined.

$$E = E_{shape} + E_{NVFP}, \quad (4)$$

where E_{shape} represents the energy of constrained shape from training images. E_{NVFP} denotes the variation of NVFP appearance. The point distribution model is constructed to represent the shape \mathbf{x} as follows.

$$\mathbf{x} = (x_1, x_2, \dots, x_M), \quad (5)$$

where x_i is a 1-D vector that represents the coordinates of the i th contour point and M is the number of contour points. E_{shape} is determined by the shape model which derives from the

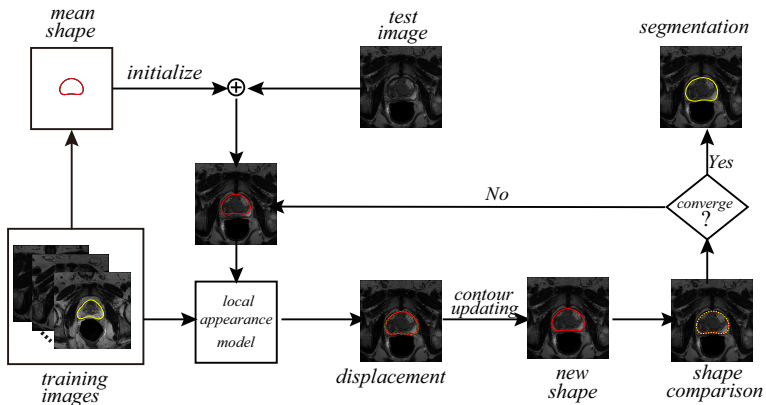


Figure 6: The flow chart of the proposed segmentation model.

principal component analysis (PCA) of training shapes.

The training of local appearance model is described in Section 2. A proper way to minimize E_{NVFP} can improve the accuracy of segmentation. The Mahalanobis distance assumes a normal distribution of training NVFPs. However, this assumption is not always hold. In addition, the dimensionality of NVFP may be larger than the number of the training data. In this case, Mahalanobis distance is an approximated measure. Therefore, Euclidean distance is employed in our work. Minimizing E_{NVFP} based on the Euclidean distance can obtain the contour point. The target point is located according to the objective function

$$x_i^{target} = \arg \min_{\{x_i^j | j = -T, \dots, +T\}} (g(x_i^j) - g_s(x_i)), \quad (6)$$

where $g(x_i^j)$ presents the NVFP of the j th point along the normal direction for x_i , $g_s(x_i)$ is the matching feature and T denotes the maximum displacement. The target point is to guide the contour deformation.

The proposed segmentation method proceeds as shown in Fig. 6.

- 1) Train the shape model and local appearance model.
- 2) Initialize the shape in test image with mean shape.
- 3) For $2T+1$ candidate points along the normal profile, compute $g_s(x_i)$ for j th point and match it with the training NVFPs to obtain the displacement, then move the contour point to x_i^{target} .
- 4) Update the current contour and compare it with previous contour. It converges when the current contour is close enough to the previous contour. If it doesn't converges, go to 3).

4 Experiments

In our experiments, 40 MR images of the prostate were used to validate the performance of the proposed method. The images have the size of 512×512 with the resolution of $0.3mm \times 0.3mm$. The manual segmentation done by experts was considered as the ground truth. Mean absolute distance (MAD) [5] is the quantitative performance measure to evaluate the proposed method, which is the average distance between the automatic segmentation and

the ground truth.

$$MAD = \frac{1}{M} \sum_{m=1}^M (\|c_m - d_m\|, c_m \in X_a, d_m \in X_g), \quad (7)$$

where X_a denotes the automatic segmentation and X_g denotes the manual segmentation.

There are two sets of experiments in this paper. The former is to select the parameters used in our method. The latter is to compare the proposed method with the classic ASM [4].

The parameters in this paper are as follows. Among these parameters, M and T are set manually, while k and δ are tested in different values and set according to the experimental results. The performance of the proposed feature NVFP varies due to different values of k and δ .

Table 1: Parameters used in our experiments

Parameters	Value and description
N	39 (number of training samples)
M	64 (number of points in each contour)
k	1,3,5,7,9,11 (number of sample points to build the profile for each contour point)
δ	1,2,...,16 (spacing of the adjacent two points of the profile)
T	6 (maximum of the displacement along the normal direction)

Table 2: MAD of different parameters (unit in mm)

k	$\delta=1$	$\delta=2$	$\delta=3$	$\delta=6$	$\delta=16$
11	1.13	1.26	1.43	2.75	4.56
9	1.12	1.27	1.27	1.39	7.51
7	1.12	1.21	1.71	1.88	1.81
5	1.16	1.17	1.21	1.50	2.15
3	1.17	1.16	1.16	1.21	2.36
1	1.19				

4.1 Selecting The Parameters

The two parameters of k and δ determine the performance of the proposed feature of NVFP, so they were selected through the segmentation results. In this set of the experiments, points of equal number were sampled on each side of the contour along the normal direction. By comparing the performance of using different numbers of feature points, we aim to obtain the range of the numbers for feature extraction and the proper spacing between the adjacent points along the normal direction. Within the range, least redundant feature information is included in the appearance profile. From the quantitative segmentation results shown in Table 2 and some qualitative results shown in Fig. 7, we can conclude that the NVFP is most distinctive when k and δ are set to 9 and 1, respectively. However, the value of k used in the following experiments is 7 through the trade-off between the computation time and the accuracy.

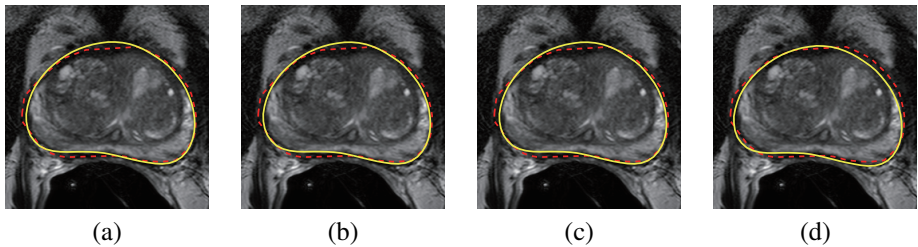


Figure 7: The prostate segmentation results from MR images. The yellow solid line shows the automatic segmentation results from different parameters and the dash red line is the manual segmentation. The parameters of k and δ are 7 and 1 in (a), 1 and 1 in (b), 9 and 3 in (c), 9 and 6 in (d) respectively.

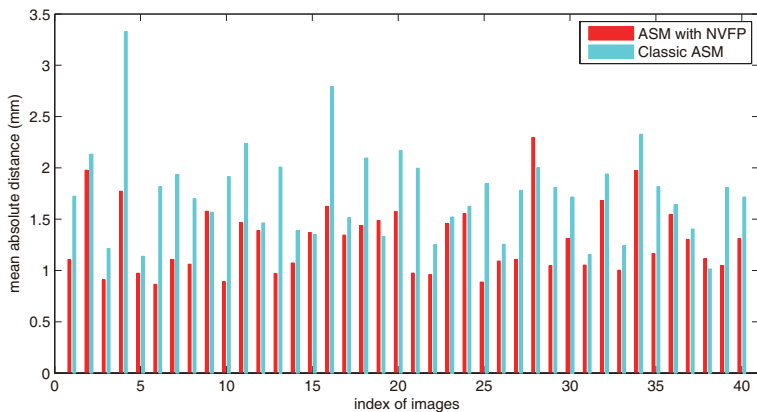


Figure 8: The mean absolute distance is shown for each image. The left bar (in red) is the results of ASM with NVFP while the right bar (in cyan) bar is the results of Classic ASM.

Table 3: MAD of two methods (unit in mm)

Algorithm	Mean \pm std	Min	Max	Med
ASM with NVFP	1.30 \pm 0.41	0.87	2.30	1.23
Classic ASM	1.74 \pm 0.45	1.01	3.30	1.72

4.2 Cross-validation of The Proposed Method

Because of the limited number of training samples, the method of leave-one-out was employed for the validation of the proposed method. Some qualitative results of the segmentation are given in Fig. 8 and Fig. 9 and quantitative results in Table 3. The average MAD is $1.30mm$ in ASM with NVFP and $1.74mm$ in classic ASM. The profile using proper number and locations of sample points makes it easier to distinguish the points located along the normal direction. Also, the qualitative segmentation results show that ASM with NVFP is more likely to latch on the weak and ambiguous boundaries and more robust to strong edge in the neighboring region of the true boundary. It can be seen that feature of NVFP improves the segmentation performance greatly.

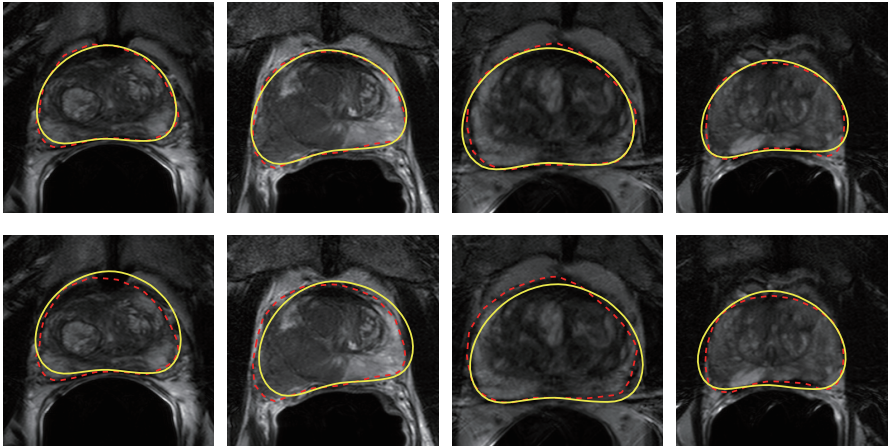


Figure 9: Some qualitative segmentation results of ASM based methods. The ASM based segmentation results are shown as yellow solid line and the manual segmentations are shown as red dash line. The top and the bottom were obtained by using the proposed method and the standard ASM method, respectively.

5 Conclusion

In this paper, a modified SIFT feature is used to form the normal vector feature profile, which is to capture the distinctiveness of a point. Then, a proper number and locations of the sample points for NVFP is determined for each contour point. For the high dimensionality of the new feature and small samples, Euclidean distance is used in our work. The contributions are proven to be effective in segmenting prostate in MR images. In our future work, the proposed method will be extended to 3D models in medical image segmentation.

6 Acknowledgments

The presented research work is supported by the National Basic Research Program of China (973 Program) (Grant No. 2011CB707000) and the National Natural Science Foundation of China (Grant No. 61072093).

References

- [1] T. McInerney, and D. Terzopoulos. Deformable models in medical image analysis. *Medical Image Analysis*, 1(2):91-108, June, 1996.
- [2] P. Yan, S. Xu, B. Turkbey, J. Kruecker, Adaptively Learning Local Shape Statistics for Prostate Segmentation in Ultrasound, *IEEE Trans. Biomedical Engineering*, vol. 58, no. 3, pp. 633-641, 2011.
- [3] W. Zhang, P. Yan, X. Li, Estimating Patient-specific Shape Prior for Medical Image Segmentation, *IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1451-1454, Chicago, USA, 2011.

- [4] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical images. *Image and Vision Computing*, 12(6):355-365, July-August, 1994.
- [5] R. Toth, P. Tiwari, M. Rosen, G. Reed, J. Kurhanewicz, A. Kalyanpur, S. Pungavkar, and A. Madabhushi. A magnetic resonance spectroscopy driven initialization scheme for active shape model based prostate segmentation. *Medical Image Analysis*, 15(2):214-225, April, 2011.
- [6] Q. Feng, M. Foskey, W. Chen, and D. Shen. Segmenting CT prostate images using population and patient-specific statistics for radiotherapy. *Medical Physics*, 37(8): 4121-4132, August, 2010.
- [7] S. Chen, D. M. Lovelock, and R. J. Radke. Segmenting the prostate and rectum in CT imagery using anatomical constraints. *Medical Image Analysis*, 15(1):1-11, February, 2011.
- [8] Y. Shi, F. Qi, Z. Xue, L. Chen, K. Ito, H. Matsuo, and D. Shen. Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics. *IEEE Transactions on Medical Imaging*, 27(4):481-494, April, 2008.
- [9] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal Computer Vision*, 60(2):91-110, November, 2004.
- [10] J. Peters, O. Ecabert, C. Meyer, R. Kneser, and J. Weese. Optimizing boundary detection via Simulated Search with applications to multi-modal heart segmentation. *Medical Image Analysis*, 14(1):70-84, February, 2004.