PROSTATE SEGMENTATION BASED ON VARIANT SCALE PATCH AND LOCAL INDEPENDENT PROJECTION

Yao Wu, Guoqing Liu, Meiyan Huang, Jun Jiang, Wei Yang, Wufan Chen, Qianjin Feng*

School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

ABSTRACT

Accurate segmentation of the prostate in computed tomography (CT) images is very important in image-guided radiotherapy. In the current study, an automatic framework is proposed for prostate segmentation in CT images: first, we propose a novel image feature extraction method, namely, variant scale patch, which can provide rich image information in a low dimensional feature space; second, we take the general idea of sparse representation and design a new segmentation criterion called local independent projection (LIP); third, we use an online update strategy to construct a dictionary to utilize the latest image information. Furthermore, in the proposed LIP, we emphasize locality rather than sparsity, and use local anchor embedding to solve the dictionary coefficients. The proposed method is evaluated based on 201 3D images of 12 patients. Results show that the proposed method is robust in segmenting prostates in CT images.

Index Terms—Prostate segmentation, variant scale patch, local anchor embedding, local independent projection

1. INTRODUCTION

The precise segmentation of the prostate in CT images is important for prostate cancer radiotherapy. In traditional treatments, the prostate is manually segmented by radiologists; however, manual segmentation is time consuming and results obtained by different radiologists often differ. Therefore, effective prostate segmentation methods are valuable in clinical applications.

Recently, many methods have been proposed on this topic. Deformable model based method is commonly used to perform prostate segmentation in CT images [1, 2]. However, the performance of these methods highly depends on an accurate initialization of the deformable model, which is difficult to obtain because of variations in prostate appearance. Registration based method is another common category [3, 4]. However, accurate registration results are difficult to obtain because of the uncertain presence of bladder water and rectum gas. More recently, classification based method has been introduced to prostate segmentation [5, 6]. The current study also belongs to this category. The main contributions in this study are described as follows.

First, we propose a novel image feature extraction algorithm called variant scale patch (VSP). The traditional feature extraction strategy obtains the optimal characteristics by evaluating several candidate features using automatic methods [7, 8] or manual methods, which depend on the performance of the feature selection method or subjectivity of users. Studies [9, 10] have shown that the raw image patch can exhibit good performance in many cases. Since the context information is helpful for prostate segmentation [5], if we want to utilize the context sufficiently and enlarge the patch size, the high dimensional feature will remarkably increase the computational burden. To overcome the limitation of the patch-based feature used in [9, 10], we apply the proposed VSP, which can provide rich image information in a low dimensional feature space.

Second, we design a new classification approach, namely, local independent projection (LIP). In sparse representation-based classification (SRC) [11], a dictionary containing training samples from all classes is used to represent the testing sample, and the dictionary coefficient is solved by using the sparse coding (SC) with $L_1$-norm. Similar to SRC, LIP is based on the concept that a sample can be sparsely represented by a dictionary. In contrast to SRC, we assume that training samples from different classes lie on different submanifolds. The testing sample is independently projected on each submanifold by using a locally linear representation [12].

Third, we use a new dictionary construction approach. A series of daily scans should be obtained during radiotherapy because the appearance of the prostate varies greatly on different treatment days. To utilize the latest image information, we propose an online strategy to update the dictionary.

2. METHOD

2.1. VSP-based feature extraction

The proposed VSP is established by investigating the features of non-uniform visual acuity on human retina. The distribution of cones in the fovea is denser than that in the peripheral area. Thus, the eye captures the scene in variant scales. In the central region, a high number of receptors are used to generate high-resolution images. However, coarse context information is sampled by fewer receptors in the
peripheral region. The proposed VSP is based on the sampling manner of the eye. In this scheme, the sampling density is designed to decrease as the location becomes farther from the central point, which ensures that the extracted features contain detailed information of the central region and coarse context of the peripheral region. At the same time, the dimensionality of the features is maintained at a relatively low level.

To compute the feature at a point \( c \), an image patch \( P_c \) centered at \( c \) is considered. Parameter \( \sigma \) is a variable over \( P_c \), and \( \sigma_p \) (the sampling scale at point \( p \)) increases as \( p \) is located farther from \( c \). Thus, we define \( \sigma_p \) as follows:

\[
\sigma_p = g(D(c,p))
\]

where \( D(c,p) \) is the Euclidean distance between points \( c \) and \( p \). \( g \) is an increasing function. In the current study, \( g \) is defined as an exponential function:

\[
g(D) = \lambda e^{D/\alpha}, \quad \alpha > 0
\]

where \( \alpha \) is used to control the increase in speed of the scale along \( D \). \( \lambda \) determines the sampling interval in point \( c \). Using this scale, we can transform each point in \( P_c \) into a new coordinate space. Considering point \( s' \) located at \( x_s' \) (\( x_s' \) is the relative coordinate to the central point \( c' \)) in VSP, we can calculate the corresponding location \( x_s \) on the original image patch according to the following equation:

\[
x_s^i = x_s'^i \left( \int_{s'}^{s'} \sigma_s ds / D(s',c') \right)
\]

where \( x_s^i \) and \( x_s'^i \) represent the \( i \)th coordinate element of \( x_s \) and \( x_s' \), respectively. \( \int_{s'}^{s'} \sigma_s ds \) denotes the line integral of scale \( \sigma \) from point \( s' \) to point \( s' \) in VSP. According to Eqs. (1) and (2), Eq. (3) can be rewritten as follows:

\[
x_s^i = x_s'^i \left( \int_{c'}^{s'} \lambda e^{D(s',c')/\alpha} ds / D(s',c') \right)
\]

\[
= a \lambda x_s'^i \left( e^{D(s',c')} - 1 \right) / D(s',c')
\]

To implement VSP conveniently, we define a sampling template \( S' = \{ s_i \}_{i=1}^{N} \) consisting of \( N \) points on regular grids in VSP [Fig. 1(a)]. According to Eq. (4), the template \( S' \) can be transformed into a warped template \( S = \{ s_i \}_{i=1}^{N} \) on the original image space [Fig. 1(b)]. The warped template is further truncated using a circular region around the central point \( c \) because the farthest sampling points are mostly irrelevant to the central point. The sampling points in this circular region remain and compose the final template \( S_{final} = \{ s_i \}_{i=1}^{N} \) [Fig. 1(c)]. For a sampling point \( s_i \) in \( S_{final} \), the covered region on the original patch can be approximated by a circle neighborhood with a diameter of scale \( \sigma_{s_i} \) [Fig. 1(d)]. The image contents in this circle neighborhood are extracted as the subfeature noted as \( f_{s_i} \) at point \( s_i \), and the subfeatures of all \( s_i \) in \( S_{final} \) compose the feature of central point \( c \) as \( f = \{ f_{s_1}, f_{s_2}, ..., f_{s_M} \} \) [Fig. 1(d)].

Fig. 2 shows the variation of VSP under different \( a \) (in Eq. (2)) in 2D images. In this figure, the size of the sampling template is \( 5 \times 5 \), and 21 sampling points remain in the truncated template after removing four farthest sampling points. As parameter \( a \) decreases, the covered region of VSP enlarges accordingly. If parameter \( a \) is set to be 0.5, the area covered by VSP is larger than \( 100 \times 100 \) [Fig. 2(a)]. That is to say, if we take the mean intensity as the subfeature in VSP, the final feature of the central point can be represented by a vector with only 21 dimensions.

![Fig. 1. Demonstration of VSP construction.](image1.png)

Numerous approaches, such as gray level histogram, mean and variance of image intensities, can be exploited in each circle to extract the subfeature in VSP, where users can choose applicable methods based on their special tasks.

![Fig. 2. The variation of circles using different a in VSP.](image2.png)

2.2. Classification based on LIP

2.2.1. Basic idea of LIP

In the current study, a novel classification method called LIP-based classification is proposed. Here, we firstly present an assumption.

**Assumption 1:** Samples from different classes are located on different non-linear submanifolds, and a sample can be approximately represented as a linear combination of several nearest neighbors from the same submanifold.

For \( N \)-class classification problem, this assumption indicates that the samples lie on a manifold \( \{ M_i \}_{i=1}^{N} \), which
consists of \( N \) submanifolds, where \( M^i \) represents the submanifold associated with the \( i^{th} \) class. For a dictionary \( D = \{D^1, \ldots, D^N\} \), which consists of \( N \) subdictionaries, where \( D^i = \{d^i_1, d^i_2, \ldots, d^i_t\} \) consists of \( t \) samples from the \( i^{th} \) submanifold. A testing sample \( f \) can be projected onto each submanifold using the following linear representation:

\[
f = D^i a + e^i = \sum_{j=1}^{t} a_j d^i_j + e^i
\]

where \( a = (a_1, a_2, \ldots, a_t)^T \) denotes the coefficient vector of \( D^i \); \( e^i \) represents the reconstruction error for \( f \) associated with \( M^i \). The classification rule is expressed as follows:

\[
l^* = \min_i \|e^i\|_2
\]

According to Eq. (6), \( f \) is labeled to the class \( l^* \) with the minimum reconstruction error.

### 2.2.2. Implementation of LIP-based classification

There are three steps in LIP-based classification: dictionary construction, coding, and classification.

In the dictionary construction step, we use the \( k \)-means method to cluster the training samples and select the cluster centers to form a compact dictionary.

In the coding step, we emphasize the locality property of the representation and use local anchor embedding (LAE), which focuses on locality rather than sparsity by limiting linear coding with a local neighborhood. The cost function of LAE can be formulated as follows:

\[
a = \arg\min_a \left\| f - \sum_{j=1}^{t} a_j d^i_j \right\|_2^2
\]

s.t. \( \forall d^i_j \in N_j(k), a_j = 0 \).

In this section, we evaluate the effectiveness of the proposed method. To compare with relevant prostate segmentation methods, we use five popular evaluation measures: dice similarity coefficient, average surface distance (ASD), centroid distance (CD), probability of detection (PD), and probability of false alarm (FA).

### 3. EXPERIMENTAL RESULTS

In this section, a series of experiments are presented to evaluate the performance of the proposed method. To compare with relevant prostate segmentation methods, we use five popular evaluation measures: dice similarity coefficient, average surface distance (ASD), centroid distance (CD), probability of detection (PD), and probability of false alarm (FA).

#### 3.1. Subjects

The dataset used in the current experiments consists of 12 patients, each with more than 10 daily CT scans and with a total of 201 3D CT images. The voxel resolution of each image is 1 mm \( \times \) 1 mm \( \times \) 3 mm. Each scan has an XY size of 512 \( \times \) 512 with more than 44 slices. Manual segmentations of the prostates in all images are provided to evaluate segmentation results.

#### 3.2. Effectiveness of VSP-based feature extraction

In this section, we evaluate the effectiveness of the proposed VSP. Here, we use the mean intensity in each circle to denote the subfeature in VSP. Compared with the patch-based feature used in [9, 10], which is obtained by vectorizing the intensities of the image patch centered at the central point, the mean Dice ratio of VSP-based feature is enhanced from 78.7% to 91.6%, with an increase of 12.9% (t test; \( p < 0.001 \)). Fig. 3 shows some segmentation results using VSP- and patch-based features, respectively.

![Fig. 3. Results of the seventh daily scan of the ninth patient using VSP (first row) and patch-based features (second row). Red curves show the ground truth and yellow curves show the automatic segmentation results obtained by using different features. The corresponding Dice ratios are 91.57% and 79.23%, respectively.](image)

#### 3.3. Effectiveness of LIP-based classification

To evaluate the effectiveness of LIP, we compare it with SRC [11]. Here, we use the same dictionary in LIP and SRC with different coding methods and classification criterions. After investigating the results of the 12 patients [Fig. 4], we find that the mean Dice ratio of LIP is 10.8% higher than that of SRC (t test; \( p < 0.001 \)), which shows that the proposed LIP is effective in classification.
3.4. Comparison with relevant methods

To evaluate the performance of the proposed method, we compare it with six relevant methods. Methods proposed by Chen et al. [1] and Feng et al. [2] are based on deformable models. Methods proposed by Davis et al. [3] and Liao et al. [4] are based on the registration technique. The other two methods [5, 6] belong to the classification category. Table 1 shows the results of these methods under five evaluation measures. The best results are bolded in this table.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Mean Dice</th>
<th>Mean ASD</th>
<th>Mean CD (x/y/z)</th>
<th>Median PD</th>
<th>Median FA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen et al. [1]</td>
<td>N/A</td>
<td>1.10</td>
<td>N/A</td>
<td>0.84</td>
<td>0.13</td>
</tr>
<tr>
<td>Feng et al. [2]</td>
<td>0.893</td>
<td>2.08</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Davis et al. [3]</td>
<td>0.820</td>
<td>-0.26/0.35/0.22</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Liao et al. [4]</td>
<td>0.899</td>
<td>1.08</td>
<td>0.21/0.12/0.29</td>
<td>0.89</td>
<td>0.11</td>
</tr>
<tr>
<td>Li et al. [5]</td>
<td>0.908</td>
<td>1.40</td>
<td>0.18/-0.02/0.57</td>
<td>0.90</td>
<td>0.10</td>
</tr>
<tr>
<td>Liao et al. [6]</td>
<td>0.909</td>
<td>0.97</td>
<td>0.17/0.09/0.18</td>
<td>0.90</td>
<td>0.08</td>
</tr>
<tr>
<td>Our method</td>
<td><strong>0.916</strong></td>
<td>1.20</td>
<td><strong>0.09/0.09/0.20</strong></td>
<td><strong>0.93</strong></td>
<td><strong>0.09</strong></td>
</tr>
</tbody>
</table>

It can be observed that the proposed method yields the highest mean Dice and median PD among these methods. Although the mean CDs along y and z directions of the proposed method are slightly larger than those in Li et al. [5] and Liao et al. [6], respectively, the mean CD along the x direction is better than that of other methods. The mean ASD and the median FA of the proposed method are also acceptable. Table 1 demonstrates that the proposed method is effective and robust in segmenting prostate in CT images.

4. CONCLUSION

In the current study, a novel automatic method is proposed for prostate segmentation in CT images. A new VSP-based feature extraction method is designed to address the limitation of the patch-based feature extraction. The proposed VSP can use a lower dimensional feature space to represent more image information, thereby reducing computational burden and achieving more accurate results. Furthermore, a new classification method called LIP is proposed, where the testing sample is independently projected on each submanifold by using a locally linear representation. In addition, we use an online update strategy to obtain an informative dictionary. Experimental results demonstrate that the proposed method is effective for prostate segmentation in CT images.

ACKNOWLEDGEMENT

This research is supported by the grants from the National Basic Research Program of China (973 Program, No. 2010CB732505), National Natural Science Funds of China (Nos. 81101109 and 31371009), and Program of Pearl River Young Talents of Science and Technology in Guangzhou (No. 2013J2200065).

REFERENCES