

# PROSTATE DETECTION FROM ABDOMINAL ULTRASOUND IMAGES: A PART BASED APPROACH

Nur Banu Albayrak<sup>1</sup>

Ayşe Betül Oktay<sup>2</sup>

Yusuf Sinan Akgül<sup>1</sup>

<sup>1</sup>GTU Vision Lab, <http://vision.gyte.edu.tr>  
Department of Computer Engineering  
Gebze Technical University  
{albayrak,akgul}@bilmuh.gyte.edu.tr

<sup>2</sup>Department of Computer Engineering  
Istanbul Medeniyet University  
abetul.oktay@medeniyet.edu.tr

## ABSTRACT

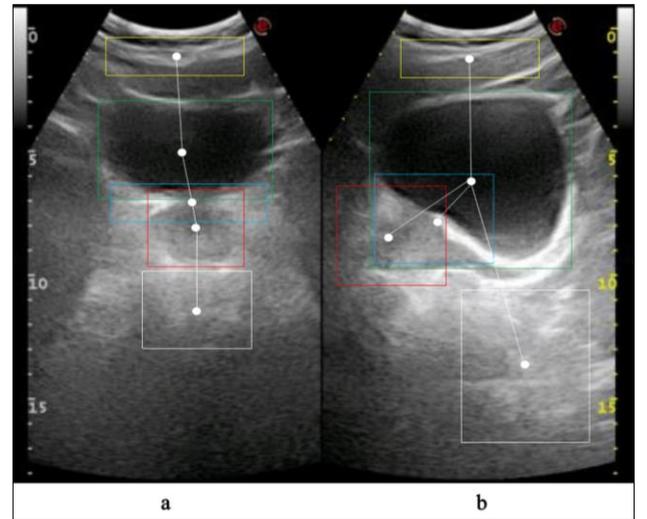
Prostate cancer is one of the most frequent cancers among men. Abdominal Ultrasound scans are very practical alternatives to more precise but inconvenient Transrectal Ultrasound scans for the diagnosis and treatment of prostate cancer. However, detection of the prostate region alone is very difficult for the Abdominal Ultrasound images. This paper presents a new prostate detection method that models the abdominal images as the classes of neighboring anatomical regions of the prostate. The proposed method has two levels: Pixel level detection assigns class scores to each pixel in the image. Model level detection uses these scores to determine the final positions of the anatomical regions in the image. This new approach is very effective for the specific problems of the Abdominal Ultrasound scans. Extensive experiments performed on real patient data with and without pathologies produce very promising results.

**Index Terms**— Prostate Detection, Abdominal Ultrasound, HOG, SVM, Graphical Model

## 1. INTRODUCTION

Prostate cancer is one of the most frequent of cancers in USA. 233,000 new cases and 29,480 deaths of prostate cancer are estimated for 2014 [1]. Ultrasound, Magnetic Resonance Imaging (MRI), and Computed Tomography (CT) are the most commonly used imaging modalities for the diagnosis and treatment of the prostate cancer [4]. Among these modalities, Transrectal Ultrasound (TRUS) has been widely used due to its inexpensiveness, portability, and real time results.

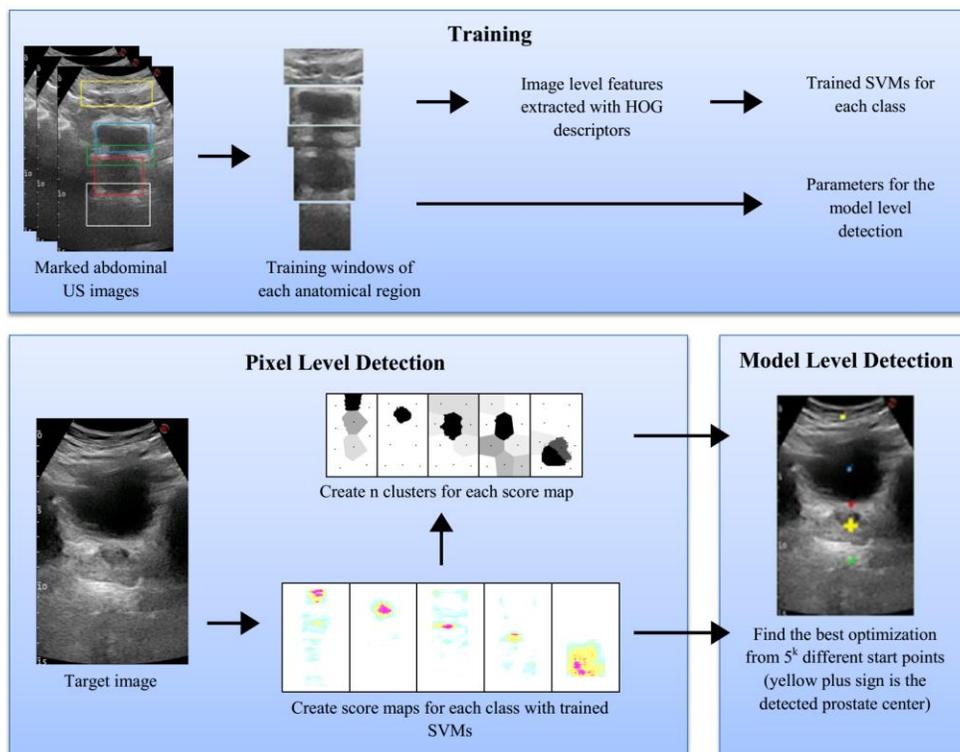
Although TRUS images have high Signal to Noise Ratio (SNR) and they show the prostate as the largest structure in the images, TRUS method is difficult to use regularly during the successive radiotherapy sequences [5] due to patient discomfort [12]. Abdominal Ultrasound (AUS) imaging is an alternative technique that is frequently employed



**Figure 1:** A sample image of the two different views of an AUS scan. Colored boxes show the anatomical regions. Red box includes the prostate region.

where TRUS is not applicable. Despite their practicality, AUS images include unrelated structures and the prostate region of AUS images is smaller than that of the TRUS images. In addition, SNR of AUS images is lower and prostate shows more positional variation. Figure 1 shows transverse and sagittal views of an AUS scan of prostate.

Although there are many studies for the analysis of the prostate in TRUS images [4], [7], [10] there are only a few reported AUS analysis work in the literature. [8] uses an image filtering and a prior shape based approach to find the prostate borders in AUS images. However, this method requires an expert initialization near the target prostate region and it was validated only on about 10 images. We argue that for a fully automatic prostate segmentation system, it is crucial to automatically localize the prostate before segmentation.



**Figure 2:** The framework of the proposed system.

In this paper, we propose a novel prostate detection method for the AUS images. We claim that detection of the prostate region should be done with the help of neighboring anatomical regions so that the unrelated structures in the AUS images (such as the bladder and the abdominal tissues) can act as helping parts of the detection system. Therefore, we propose to simultaneously detect the prostate region and the neighboring regions together in the AUS images. A two-level hierarchical approach is employed for this purpose. At the lower pixel level, the appearance information of the regions is extracted by a pixel classifier. At the higher model level, context information is obtained through the relationships between neighboring regions, which handle the problem of structural and geometrical variations between the patients. The final localization of the prostate is performed by optimizing the configuration of the regions with an enhanced initialization free gradient descent approach.

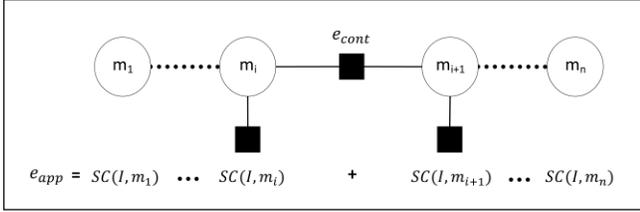
To the best of our knowledge, the proposed method is the first prostate detection system for the AUS images. Our system is initialization free and the novel application of the two level detection system addresses the unrelated structure and low SNR problems very efficiently. Furthermore, our work presents the most extensive experiment set on the AUS images.

## 2. PROSTATE DETECTION FRAMEWORK

The hierarchical framework aims to first identify the candidate regions for prostate and neighboring parts at the lower pixel level and then detect the final prostate location with contextual information at the higher level (Figure 2).

We use a graphical model to represent the anatomical regions of the AUS images. The model represents the anatomical regions with rectangular parts  $m_i = \{x_i, y_i, h_i, w_i\}$  where  $i = 1, \dots, n$ .  $x_i$  and  $y_i$  are the positions of the parts on an image and  $h_i$  and  $w_i$  are the sizes of the parts. One of these parts corresponds to the prostate region and the other parts correspond to meaningful neighboring anatomical regions, such as bladder and tissue layers between the bladder and the prostate. (Figure 2- Training).

In the pixel level detection, image features of each anatomical region are extracted by a HOG [2] based SVM classifier [13]. HOG descriptors are histograms of gradient orientations which are quantized into a number of bins. In each bin, the number of edges within an angular range is collected. Thus the local gradient information is obtained. Due to the computational cost of HOG descriptor extraction, we use the integral image technique [11] which speeds up the descriptor extraction process.



**Figure 3:** Graphical representation of the model.

For each anatomical region, an SVM classifier is trained with the extracted HOG descriptors. In the testing phase, instead of making a binary classification with SVM, we use the class scores produced by the function  $SC(I, m_i)$ . SC extracts the HOG vector from the region of  $m_i$  on image  $I$  and feeds this data to the trained SVM of the corresponding anatomical region. Finally, SC returns the distance values to the hyper-plane of the HOG based linear SVM. We produce  $n$  different score maps for a given image  $I$  as shown in Figure 2 by running the function SC for all possible  $m_i$  positions. For the overlapping part positions, we take the maximum score.

In the model level detection, the contextual information is incorporated into the system by modeling the geometric relationships between model parts.

Figure 3 shows the graphical representation of the proposed model. Circles represent  $m_i$ . Filled rectangles with a single connection correspond to  $SC(I, m_i)$  while rectangles with double connections show the context information between the neighboring parts.

We employ this model to form an optimization framework for the prostate detection problem. Let  $M = \{m_1, m_2, \dots, m_n\}$  be a configuration of parts. Given an image  $I$ , we seek the maximum a posteriori estimate of the configuration  $M$

$$\hat{M} = \underset{M}{\operatorname{argmax}} P(M|I, \Theta), \quad (1)$$

where  $I$  is the AUS image and  $\Theta$  defines the parameters learned from the training set.  $\hat{M}$  is the optimal configuration. The function  $P(M|I, \Theta)$  consists of appearance level and context level terms and it is modeled by a Gibbs distribution

$$P(M|I, \Theta) = \frac{1}{Z} \exp[-(e_{app} + e_{cont})], \quad (2)$$

The appearance term  $e_{app}$  includes the total appearance information produced at the pixel-level detection stage for all parts.

$$e_{app} = \sum_i SC(I, m_i).$$

The context level term  $e_{cont}$  includes normally distributed distance, orientation, and size terms between the parts.

$$e_{cont} = \alpha f_{dist} + \beta f_{orient} + \delta f_{size},$$

where  $\alpha, \beta, \delta$  are the weights. The distance relation between two parts represents the similarity between the current distance of two parts and the training set.

$$f_{dist} = \sum_{i=2}^n \frac{(|m_i^d - m_{i-1}^d| - \mu_i^d)^2}{(\sigma_i^d)^2},$$

where  $m_i^d = (x_i, y_i)$ ,  $\mu_i^d$  is the average distance between  $m_i^d$  and  $m_{i-1}^d$  in the training set and  $\sigma_i^d$  is the corresponding standard deviation. The orientation and the size terms are also estimated similarly.

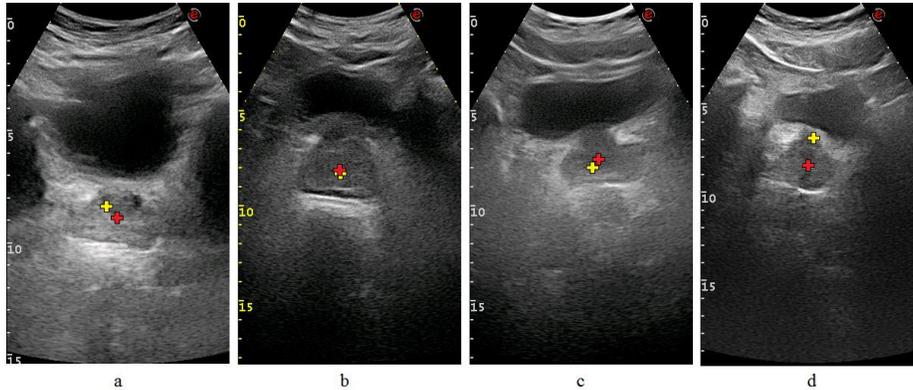
Levenberg-Marquardt Algorithm (LMA) [6] is used to optimize Equation 1 by searching an optimal set of locations  $\hat{M}$  for the parts. LMA is a gradient descent based optimization method which outperforms classical gradient descent approach with complex models. Since this algorithm is dependent on the initialization it does not guarantee globally optimal results. In order to address this problem, we use systematic multiple initializations to minimize chances of getting stuck at the local minima. We observe that a convex function can be optimally solved by LMA and Equation 2 is a convex function if the appearance term is ignored. In order to lower the influence of the appearance term in Equation 2, we cluster the score maps from the pixel level detection phase into  $k$  clusters and use the cluster centers as the positions during the multiple initializations. The K-means algorithm runs on the features of the class scores and the pixel positions. There are  $n^k$  different starts for LMA which seems very high. However, LMA is a very efficient optimization method which makes the overall process fast.

### 3. EXPERIMENTAL RESULTS

The proposed system is tested on AUS images of the transverse view. The dataset consists of 140 AUS images of different subjects. Image sizes vary between 295x470 and 380x610 pixels. The images in the dataset are gathered from subjects that have either normal prostate or abnormal prostate having various diseases. Bounding boxes of each anatomical part are manually marked in all images by an expert for the training and verification processes. We applied Ten Fold Cross Validation (TFCV) in order to evaluate our system.

#### 3.1. Pixel-Level Detection Results

We manually defined  $n=5$  different anatomical regions which are useful for the HOG based SVM. Since our system is validated with TFCV and the total number of images in our dataset is 140, the pixel level detection is performed for each of the images in dataset with 126 training images. Each part is trained and tested individually, so there are 5 different SVM in each TFCV. Therefore, each pixel is given scores with 5 SVMs and totally 5 different score maps are created for each image.



**Figure 4:** Sample detection results for the prostate centers. (a-c) are results with low error, (d) is a detection with high error. Red plus signs are the ground truth centers and yellow plus signs are detected centers.

If we had used the prostate score map as the only score map and choose the best position on this image, we obtain a classical machine learning based detector such as [2]. In this case, average of the distances between the detected centers and the marked prostate centers is 2.1 cm and only 55.7% of the detected results are within the manually marked bounding boxes. This high average distance result is expected because prostate score values represent a model with only one node and it ignores the other parts of the model. Next section presents the results of the proposed hierarchical model with all of the parts with contextual information.

### 3.2. Model-Level Detection Results

Model level detection is built over the results of the pixel level detection. The score map of each part is clustered into  $k=10$  clusters for all 140 images. The centers of these clusters are transferred to the model level as the initial points of the multiple LM optimization method. The minimum of the optimization results of different initializations are selected as the resulting positions for a target image.

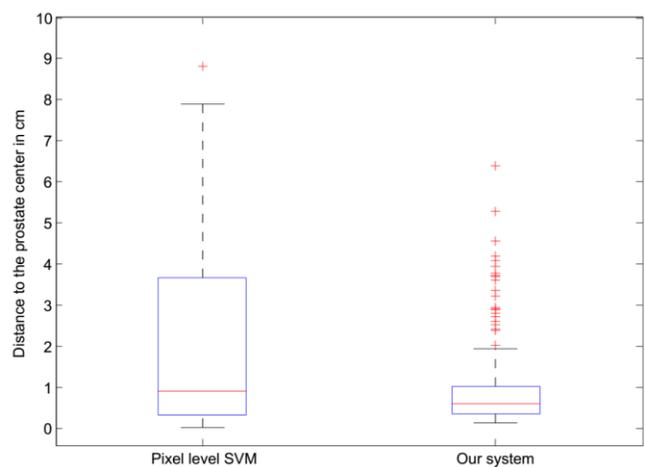
The average distance between the detected prostate center and the hand marked ground truth for 140 images is 1.08 cm and 83.5% of detected centers are inside of the bounding box of the ground truth.

The difference in the performances of the pixel level score map with HOG based SVM and the two level detection shows the effectiveness of the proposed approach. Figure 5 shows the box plot of the distances to the marked centers of pixel level SVM and the proposed system. The proposed system outperforms the classical pixel level SVM.

Figure 4 shows sample prostate detection results of 4 different AUS images in the dataset. Note that marking the prostate center is a difficult job and there is a great variation in the prostate positions marked by the experts.

## 4. CONCLUSIONS

We presented a novel method that brings the pixel-level information and the context information together under the two level framework for prostate AUS images. The proposed method for the prostate center detection fits constraints of the problem very well because the prostate center can be localized more effectively with the help of the neighboring structures. Furthermore, the structural differences between the patients are nicely handled by the optimization framework. The two level structure of the detection system helps us separate pixel level and model level knowledge. For the pixel level knowledge an HOG based SVM works efficiently. For the future work, we are planning to define the anatomical regions automatically and use the results from this system to segment the prostate borders.



**Figure 5:** Comparison between the detection results of classical SVM the and proposed method. The central mark at the box is the median, the edges of the box are the 25<sup>th</sup> and 75<sup>th</sup> percentiles, the plus signs are the outliers.

## 5. REFERENCES

- [1] National Cancer Institute Seer Stat Fact Sheets, <http://seer.cancer.gov/statfacts/html/prost.html>, Feb. 2014
- [2] Dalal, N, Triggs, Bill.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 886-893, IEEE (2005)
- [3] Felzenszwalb, P, McAllester, D, Ramanan, D.: A discriminatively trained, multiscale, deformable part model. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, IEEE (2008)
- [4] Ghose, S., Oliver, A., Martí, R., Lladó, X., Vilanova, J. C., Freixenet, J, Meriaudeau, F.: A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images. *Computer methods and programs in biomedicine*, 108(1), 262-287 (2012)
- [5] Betrouni, N., Vermandel, M., Pasquier, D., Maouche, Rousseau, J.: Segmentation of abdominal ultrasound images of the prostate using a priori information and an adapted noise filter. *Computerized Medical Imaging and Graphics*, 29(1), 43-51 (2005)
- [6] Marquardt, D., W.: An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics*. 11.2, 431-441 (1963)
- [7] Yuan; J, Qiu; W., Rajchl, M.; Ukwatta, E., Tai; X, Fenster, A.: Efficient 3D Endfiring TRUS Prostate Segmentation with Globally Optimized Rotational Symmetry. In: IEEE Conference on Computer Vision and Pattern Recognition, pp.2211-2218, IEEE (2013)
- [8] Ghanei, A., Soltanian-Zadeh, H., Ratkewicz, A., Yin, F., F.: A three-dimensional deformable model for segmentation of human prostate from ultrasound images. *Medical Physics*. 28.10, 2147-2153 (2001)
- [9] Toth, R.; Madabhushi, A.: Multifeature Landmark-Free Active Appearance Models: Application to Prostate MRI Segmentation, *IEEE Transactions on Medical Imaging*, pp.1638,1650, IEEE (2012)
- [10] Zhang Y., Matuszewski B. J., Histace A., Precioso F., Kilgallon J., and Christopher Moore.: Boundary delineation in prostate imaging using active contour segmentation method with interactively defined object regions. In: Proceedings of the 2010 international conference on Prostate cancer imaging: computer-aided diagnosis, prognosis, and intervention, Anant Madabhushi, Jason Dowling, Pingkun Yan, Aaron Fenster, and Purang Abolmaesumi (Eds.), Springer-Verlag, pp. 131-142. Berlin, Heidelberg, (2010)
- [11] Viola, P., Jones, M.: Robust real-time object detection. *International Journal of Computer Vision* 4, 34-47 (2001)
- [12] De Sio, M., D'armiento, M., Di Lorenzo, G., Damiano, R., Perdonà, S., De Placido, S., Autorino, R.: The need to reduce patient discomfort during transrectal ultrasonography-guided prostate biopsy: what do we know? *BJU International*. 96(7), 977-983 (2005)
- [13] Oktay, A., B., Akgul, Y., S.: Localization of the Lumbar Discs using Machine Learning and Exact Probabilistic Inference. In: *Medical Image Computing and Computer-Assisted Intervention*. pp. 158-165. Springer, Berlin, Heidelberg, 2011