

Automatic Prostate Segmentation in MR Images with a Probabilistic Active Shape Model

Matthias Kirschner¹, Florian Jung¹, and Stefan Wesarg²

¹ Technische Universität Darmstadt,

Graphisch-Interaktive Systeme, Darmstadt, Germany

² Fraunhofer Institute for Computer Graphics Research IGD,
Cognitive Computing & Medical Imaging, Darmstadt, Germany

Abstract. Segmentation of the prostate gland in Magnetic Resonance (MR) images is an important task for image-guided prostate cancer therapy. The low contrast of the prostate to surrounding tissue in MR images makes automatic segmentation very challenging. In this paper, we propose an automatic approach for robust and accurate prostate segmentation in T2-weighted MR scans. We first employ a boosted prostate detector to locate the prostate in the images, and then use a Probabilistic Active Shape Model for the delineation of its contour. Our approach has been quantitatively evaluated on 50 MR images, on which we achieve a median dice coefficient of 0.85 (IQR: 0.09).

1 Background

Magnetic Resonance Imaging (MRI) plays an increasingly important role for treatment of prostate cancer. MRI is used to support biopsies, radiation therapy, or planning of surgeries. Automatic segmentation of the prostate in MR images can greatly enhance the clinical workflow. Although the contrast between soft tissue organs in MR images is better than in Computed Tomography scans, accurate delineation of the prostate's border remains difficult even for the human observer. An additional challenge for automatic segmentation is that MR image acquisition does not provide standardized image intensities. Moreover, MR images may be corrupted by artifacts such as large intensity variations within a single image.

State-of-the art algorithms for prostate segmentation in MR images rely heavily on prior knowledge. Klein et al. [1] use multiple atlas registration and label fusion to segment a scan. Makni et al. [2] first segment the image with an Active Shape Model (ASM) and then refine the resulting segmentation with a Markov Random Field. Toth et al. [3] use a level set-based Active Appearance Model for segmenting the prostate. Expressive texture features are determined during the training phase by filtering the images with various kernels and selecting the most appropriate features.

In this paper, we propose an automatic segmentation algorithm for prostate segmentation in T2-weighted MR images. We extend the Viola-Jones object detection algorithm [4] to 3D and use it to detect the prostate in MRI scans. The

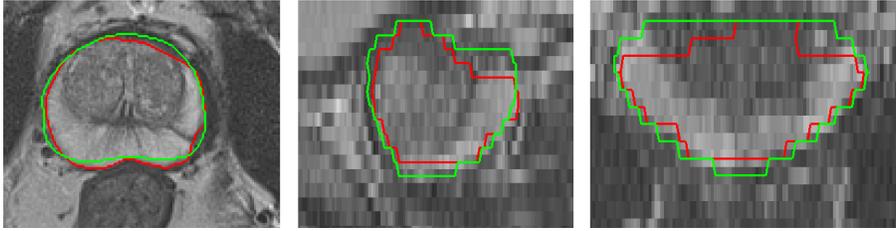


Fig. 1. Qualitative segmentation result of automatic prostate segmentation (from left to right: axial, sagittal and coronal view; Dice coefficient: 0.86). The red contour shows the expert segmentation, the green contour shows our automatic segmentation.

prostate’s contour is then delineated with a Probabilistic Active Shape Model (PASM). Our algorithm is characterized by very efficient object detection and segmentation which is considerably faster than the related approaches presented above. We trained and quantitatively evaluated the approach on 50 studies of the [Promise12](http://promise12.grand-challenge.org)³ segmentation challenge.

2 Methodology

In this section, we present our approach for automatic prostate segmentation, which is comprised of three steps. We first preprocess the image in order to obtain a normalized intensity range for all images (Sec. 2.1). Then, we slide a detector over the image which detects the prostate’s bounding box based on 3D Haar-like features (Sec. 2.2). The bounding box is used to initialize the final step of our algorithm, in which we segment the prostate with a PASM (Sec. 2.3).

2.1 Inhomogeneity correction and intensity normalization

Image intensities in MRI data are not standardized, which means that the intensity of a certain tissue may vary significantly from image to image. Moreover, one can usually observe spatially smoothly varying intensity inhomogeneities within a single MRI scan, the so-called *bias field*. As both non-standardization and the bias field impede accurate and robust segmentation, image preprocessing and intensity normalization are crucial.

We first estimate and remove the bias field. This is done using *coherent local intensity clustering* (CLIC) [5], which computes the bias field by an energy minimization based on local Fuzzy-C-Means clustering. As parameters for CLIC, we chose $\sigma = 4$ for the scale of the truncated Gaussian which defines the size of the local kernel, and use $q = 2$ as fuzzifier for the Fuzzy-C-Means. These values have been recommended by Li et al. [5]. Moreover, we clustered into $K = 4$ intensity classes, as this choice gave visually the best results.

³ <http://promise12.grand-challenge.org>

After bias field computation, we normalize and rescale the image intensities. Image normalization is often done by shifting the mean intensity to zero and scaling the intensities such that the standard deviation equals one. For more robust normalization, we use the median image intensity I_{med} and the median absolute deviation MAD_I as an estimate for mean and standard deviation, respectively. The MAD_I is defined by

$$\text{MAD}_I = \text{median}_v |I(v) - I_{\text{med}}|, \quad (1)$$

where $I(v)$ denotes the image intensity at position v . For normal distributed data, $1.4826 \cdot \text{MAD}_I$ is approximately equal to the standard deviation.

We scale the image intensities such that they are in the interval $[0, 1000]$ using the equation

$$I_{\text{norm}}(v) = \frac{(I(v) - I_{\text{med}}) \cdot 167}{1.4826 \cdot \text{MAD}_I} + 500. \quad (2)$$

Rescaled intensity values $I_{\text{norm}}(v)$ outside the bounds of the interval $[0, 1000]$ are clamped accordingly.

2.2 Prostate detection

We identify a region of interest (ROI) in the image that contains the prostate by adapting the face detection algorithm of Viola and Jones [4]. We slide a detection window of fixed size through the image, and compute for each subimage defined by the current detector position whether it contains the prostate or not. As a classifier, we use boosted decision tree stumps that classify on the basis of 3D Haar-like features. The most discriminative features are selected during the training phase by adapting Adaboost as described by Viola and Jones [4].

The tissue within the prostate’s bounding box itself is relatively homogeneous and does not contain features that are expressive enough for reliable classification. In order to improve the detection accuracy, we train our classifier such that it detects a slightly enlarged bounding box which also contains surrounding tissue. By enlarging the bounding boxes of the training data by a factor of two in x and y direction, meaningful features can be learned which allow for robust prostate detection. After the detection phase, the detected ROI is shrunk to obtain the prostate’s bounding box.

2.3 Probabilistic Active Shape Model

Having determined its bounding box, we segment the prostate with the PASM [6], which is a flexible variant of the ASM [7] that allows for a more accurate delineation. The PASM employs a Statistical Shape Model (SSM) in order to avoid that the segmentation leaks into neighbouring structures. The SSM is learned from a set of S training shapes. Each shape is represented by N 3D landmarks, which have been concatenated to a $3N$ -dimensional vector \mathbf{x}_i , $1 \leq i \leq S$. We

assume that the shapes are in correspondence, which means that landmark k describes the same anatomical feature on all shapes. Correspondence has been established with a nonrigid mesh registration algorithm, whose description is beyond the scope of this paper.

The SSM is constructed by a Principal Component Analysis of the training shapes. We compute the mean shape $\bar{\mathbf{x}} = \frac{1}{S} \sum_{i=1}^S \mathbf{x}_i$ and the $S - 1$ eigenvectors $\mathbf{p}_1, \dots, \mathbf{p}_{S-1}$ with non-zero eigenvalues $\lambda_1 \geq \dots \geq \lambda_{S-1} > 0$ of the covariance matrix $\mathbf{C} = \frac{1}{S-1} \sum_{i=1}^S (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$. We select the dimension t of the SSM's eigenspace such that it represents 98% of the observed variance, that is $t = \operatorname{argmin}_{t'} \left\{ (\sum_{i=1}^{S-1} \lambda_i)^{-1} (\sum_{i=1}^{t'} \lambda_i) \geq 0.98 \right\}$.

With the SSM, each shape \mathbf{x} can be represented by $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} + \mathbf{r}$, where $\mathbf{P} = (\mathbf{p}_1, \dots, \mathbf{p}_t)$ is a matrix containing the first t eigenvectors, $\mathbf{b} = \mathbf{P}^T(\mathbf{x} - \bar{\mathbf{x}})$ and \mathbf{r} is a residual vector.

The PASM is initialized by placing the mean shape $\bar{\mathbf{x}}$ into onto the prostate in the image. A scale factor and a translation vector that map $\bar{\mathbf{x}}$ from the SSM's coordinate system to the image coordinate system can be easily computed from the bounding box of $\bar{\mathbf{x}}$ and the detected prostate bounding box. After initialization, the shape model is iteratively adapted to the image. In each iteration, the shape is deformed such that it matches image features detected by an appearance model (Sec. 2.4). The deformed shape $\hat{\mathbf{x}}$ is then constrained with the SSM by minimizing an energy (Sec. 2.5). The algorithm terminates after a fixed number of 30 iterations.

2.4 Appearance model

Our appearance model is based on oriented 1D intensity profiles which consist of seven sampled intensity values. For each profile \mathbf{f} , we can define a set of 1D Haar-like features: Each feature $H = (j, l_1, l_2, l_3)$ can be uniquely defined by an offset j and three subprofile lengths $l_1 \geq 1$ and $l_2, l_3 \geq 0$. The feature value for H is defined by

$$h(\mathbf{f}, H) = \sum_{i=j}^{\hat{l}_1-1} f_i - \sum_{i=\hat{l}_1}^{\hat{l}_2-1} f_i + \sum_{i=\hat{l}_2}^{\hat{l}_3-1} f_i \quad (3)$$

with $\hat{l}_k = j + \sum_{i=1}^k l_k$.

In the training phase, we sample for each landmark profiles on the boundary which we label with 1 as well as slightly displaced profiles which we label with 0. Then, we train a boosted classifier that discriminates between boundary and non-boundary profiles based on 1D Haar features. Again, we use decision tree stumps as base classifiers and select the most discriminative features with Adaboost.

During segmentation, we sample several intensity profiles in the vicinity of a landmark and use the trained classifier to compute the probability of a profile that it is a boundary profile. The deformed shape $\hat{\mathbf{x}}$ is determined by optimal surface detection [8], in which we select a consistent set of image features across the whole shape by solving a max-flow problem.

2.5 Constraining shapes with energy minimization

The deformed shape $\hat{\mathbf{x}}$ is constrained by minimizing the energy

$$E(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{w}) = \alpha \cdot (E_{\text{image}}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{w}) + E_{\text{local}}(\mathbf{x})) + E_{\text{shape}}(\mathbf{x}) \quad (4)$$

where α is a balancing parameter ($\alpha = 0.5$ throughout this work), and $\mathbf{w} \in \mathbb{R}$ contains a weight for each landmark that assesses the confidence of its appearance model. The energy minimization is performed in the coordinate system of the SSM, which means that scale, translation and rotation must not be affected by the optimization. The transformation from image to model coordinate system is done as in the standard ASM [7]. For minimization, we use the limited-memory BFGS algorithm.

The global shape energy $E_{\text{shape}}(\mathbf{x})$ ensures that the shape is similar to the training shapes. It is defined by

$$E_{\text{shape}}(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^t \frac{b_i}{\lambda_i} + \frac{S-t-1}{2 \cdot \sum_{i=t+1}^{S-1} \lambda_i} \|\mathbf{r}\|^2 \quad (5)$$

and approximates the negative log-likelihood of a shape, under the assumption of a Gaussian shape distribution.

It is important to note that the global shape energy does not restrict shapes to the subspace spanned by the t principal eigenvectors, as the magnitude of the residual vector \mathbf{r} is allowed to be larger than zero. This allows a more flexible adaptation of the SSM to unseen shapes than in the standard ASM. A local shape energy regularizes this additional deformation such that the organ’s contour remains smooth. It is defined by

$$E_{\text{local}}(\mathbf{x}) = \sum_{i=1}^N \sum_{j \in \mathcal{N}(i)} \|\mathbf{x}^{(i)} - \boldsymbol{\mu}^{(i)} - \mathbf{x}^{(j)}\|^2, \quad (6)$$

where $\mathbf{x}^{(i)} \in \mathbb{R}^3$ denotes the coordinates of landmark i , $\mathcal{N}(i)$ denotes the set of neighbors of i , and $\boldsymbol{\mu}^{(i)} = \bar{\mathbf{x}}^{(i)} - \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \bar{\mathbf{x}}^{(j)}$ is the average relative position of a landmark to its neighbours. The integration of $\boldsymbol{\mu}^{(i)}$ in Eq. 6 guarantees that characteristic curvature features of the shapes are preserved.

Finally, the image energy ensures that the optimized shape \mathbf{x} is close to the detected appearance features. It is defined by

$$E_{\text{image}}(\mathbf{x}; \hat{\mathbf{x}}; \mathbf{w}) = \sum_{i=1}^N \|w_i(\mathbf{x}^{(i)} - \hat{\mathbf{x}}^{(i)}) + \sum_{j \in \mathcal{N}(i)} w_j(\mathbf{x}^{(j)} - \hat{\mathbf{x}}^{(j)})\|^2. \quad (7)$$

In contrast to the image energy previously proposed [6], $E_{\text{image}}(\mathbf{x}; \hat{\mathbf{x}}; \mathbf{w})$ penalizes the net deviation of neighboring landmarks from their corresponding image features. By this, the image energy counteracts a shrinking force caused by the local shape energy. The texture weights w_i are normalized such that they sum up to N . Thus, image energy and local shape energy are automatically balanced.

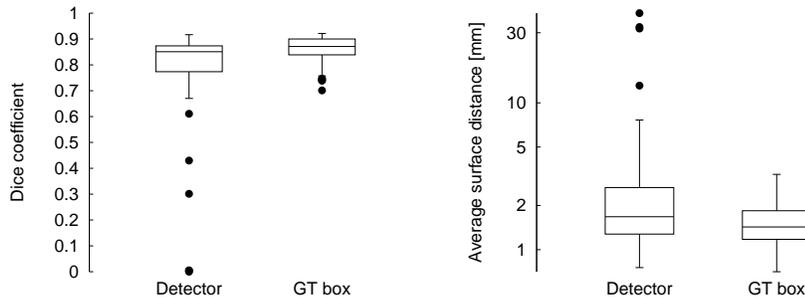


Fig. 2. Boxplot showing quantitative segmentation results on the 50 training scans. *Detector* corresponds to the automatic approach. *GT box* corresponds to initialization of the shape model using the ground truth bounding box.

3 Experimental Design

We trained and tested our algorithm quantitatively on the 50 training cases of the [Promise12](#) challenge. Moreover, we performed qualitative tests on the 30 test cases of the challenge, for which no ground truth was provided. In order to avoid training and testing on the same data, experiments on the training data was done using leave-one-out appearance model training and 5-fold cross validation for prostate detector training. Experiments on the test data exploited the complete training data. We measured the segmentation accuracy compared to the ground truth with Dice Coefficient, Average Symmetric Surface Distance (ASD) and Hausdorff Distance (HD). As the prostate detector does not always estimate the prostate’s bounding box well, we performed additional experiments by initializing the PASM using the ground truth bounding box (*GT Box*) in order to quantify the segmentation accuracy of the PASM under optimal initialization.

4 Results and Discussion

4.1 Quantitative Results

Quantitative results of our experiments are listed in Figure 2 and Table 2. The automatic algorithm achieves in most cases an accurate segmentation of the prostate. For 37 out of 50 scans, the Dice coefficient is larger than 0.8. However, on four scans, the segmentation failed completely ($Dice < 0.5$), which was caused by misdetection of the prostate. Segmentation results with low dice coefficient mainly coincide with slightly displaced bounding box estimates, or misjudgment of the prostate’s size. The quantitative results we obtained when using the ground truth bounding box for initializing the shape model shows that the segmentation itself is robust given appropriate initialization. On the 30 test cases, the prostate was accurately located in every scan.

Table 1. Quantitative results on the training data with different initialization. Shown are median and interquartile range (IQR) as well as mean and standard deviation (SD) of the respective measure from 50 experiments.

	Dice				ASD [mm]				HD [mm]			
	Median	IQR	Mean	SD	Median	IQR	Mean	SD	Median	IQR	Mean	SD
GT box	0.87	0.06	0.86	0.06	1.43	0.64	1.60	0.63	9.14	2.62	9.51	2.73
Detector	0.85	0.09	0.77	0.23	1.68	1.13	4.10	7.81	10.65	4.95	15.00	14.72

Table 2. Overview of implementation details and the efficiency of our approach.

		Parameter	Value
Algorithm		<i>Language:</i>	C++
		<i>Libraries/Packages:</i>	Insight Toolkit (ITK), Visualization Toolkit (VTK)
		<i>Multi-Threaded:</i>	only used during training (OpenMP)
		<i>User Interaction:</i>	none
Machine		<i>CPU Clock Speed:</i>	2.5 GHz
		<i>Machine CPU Count:</i>	4
		<i>Machine Memory:</i>	8 GB
		<i>Memory Used During Segmentation:</i>	≈ 300 MB
Time		<i>Training Time (Prostate detector):</i>	36 hours (for 50 studies)
		<i>Training Time (Appearance model):</i>	8 minutes (for 50 studies)
		<i>Image Preprocessing Time:</i>	6-8 minutes (per study)
		<i>Prostate Detection Time:</i>	≈ 1 second (per study)
		<i>Shape Model Adaption Time:</i>	≤ 18 seconds (per study)
		<i>Total Segmentation Time:</i>	6-8 minutes (per study)

4.2 Implementation Details & Efficiency

Our approach was implemented in C++. We used the open source library Insight Toolkit (ITK) (www.itk.org) for image processing and its sublibrary VNL for numerical computations such as the BFGS optimization. While we used OpenMP (www.openmp.org) to parallelize the training phase, the segmentation algorithm itself is single-threaded. The training of the prostate detector is relatively complex and requires up to 36 hours on a standard desktop PC. Moreover, we require approximately eight minutes for training the appearance models of all landmarks. The segmentation algorithm requires between 6 to 8 minutes for segmenting a study. Most of the running time is spent for image preprocessing, in particular for bias field correction. After preprocessing, prostate detection and the actual segmentation requires less than 20 seconds. An overview of implementation details and the efficiency of our approach are listed in Table 2.

5 Concluding Remarks

In this paper, we presented an automatic algorithm for segmenting the prostate in T2-weighted MR images. Our evaluation shows that the segmentation algorithm achieves accurate segmentation of the prostate given an appropriate initialization of the shape model. The initialization is provided by a very fast prostate detector, which estimates the prostate's bounding box accurately in most images. However, in some rare cases the detector fails to detect the prostate or misjudges its size. In these cases, accurate segmentation is not possible. Considering the good performance of the detector on the test set, we believe that the robustness of the detector can be considerably improved by increasing the training set. For use in clinical practice, it would also be possible to reinitialize the model manually in case of misdetection.

With our approach, both prostate detection and segmentation can be done in a few seconds on preprocessed image data. The bottleneck of our implementation is currently the bias-field correction. We expect that a multi-threaded implementation of CLIC [5] would at least halve the computation time on today's multi-core CPUs. Alternatively, the image filter can be replaced by faster algorithms with the same purpose.

References

1. Klein, S., van der Heide, U.A., Lips, I.M., van Vulpen, M., Staring, M., Pluim, J.P.W.: Automatic segmentation of the prostate in 3d mr images by atlas matching using localized mutual information. *Medical Physics* **35**(4) (2008) 1407–1417
2. Makni, N., Puech, P., Lopes, R., Dewalle, A., Colot, O., Betrouni, N.: Combining a deformable model and a probabilistic framework for an automatic 3D segmentation of prostate on MRI. *International Journal of Computer Assisted Radiology and Surgery* **4** (2009) 181–188
3. Toth, R., Madabhushi, A.: Multi-feature landmark-free active appearance models: Application to prostate MRI segmentation. *IEEE Transactions on Medical Imaging* (2012) (preprint; IEEE Early Access Articles).
4. Viola, P., Jones, M.: Robust real-time object detection. *International Journal of Computer Vision* **57**(2) (2001) 137–154
5. Li, C., Xu, C., Anderson, A.W., Gore, J.C.: MRI tissue classification and bias field estimation based on coherent local intensity clustering: A unified energy minimization framework. In Prince, J.L., Pham, D.L., Myers, K.J., eds.: IPMI. Volume 5636 of LNCS. (2009) 288–299
6. Kirschner, M., Wesarg, S.: Active shape models unleashed. In: *Proc. SPIE Medical Imaging 2011: Image Processing*. (2011)
7. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - their training and application. *Computer Vision and Image Understanding* **61**(1) (1995) 38–59
8. Heimann, T., Münzing, S., Meinzer, H.P., Wolf, I.: A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. In Karssemeijer, N., Lelieveldt, B., eds.: IPMI. Volume 4584 of LNCS. (2007) 1–12