SEGMENTATION OF THE PROSTATE IN MR IMAGES BY ATLAS MATCHING

S. Klein[a], U.A. van der Heide[b], B.W. Raaymakers[b], A.N.T.J. Kotte[b], M. Staring[a], and J.P.W. Pluim[a]*

[a]Image Sciences Institute, University Medical Center Utrecht
[b]Department of Radiotherapy, University Medical Center Utrecht
Q0S.459, P.O. Box 85500, 3508 GA Utrecht, The Netherlands

ABSTRACT

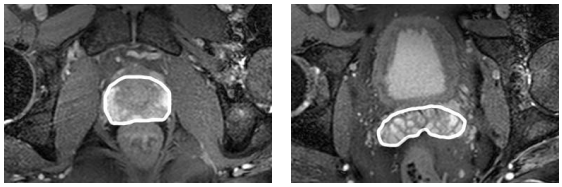
Prostate cancer treatment by radiation therapy requires an accurate localisation of the prostate. For the treatment planning, primarily computed tomography (CT) images are used, but increasingly magnetic resonance (MR) images are added, because of their soft-tissue contrast. In current practice at our hospital, a manual delineation of the prostate is made, based on the CT and MR scans, which is a labour-intensive task. We propose an automatic segmentation method, based on nonrigid registration of a set of prelabelled MR atlas images. The algorithm consists of three stages. Firstly, the target image is nonrigidly registered with each atlas image, using mutual information as the similarity measure. After that, the best registered atlas images are selected by comparing the mutual information values after registration. Finally, the segmentation is obtained by averaging the selected deformed segmentations and thresholding the result. The method is evaluated on 22 images by calculating the overlap of automatic and manual segmentations. This results in a median Dice similarity coefficient of 0.82.

***Index Terms*—** image segmentation, atlas matching, magnetic resonance imaging, prostate

1. INTRODUCTION

Prostate cancer is often treated by radiation therapy. During the radiation, precise targeting of the prostate is important: neighbouring tissue (rectum and bladder) should be spared, while the tumour should receive a maximum dose. Accurate localisation of the prostate is therefore required.

Computed tomography images are commonly used for the treatment planning, because of the electron density information they provide. For delineation of the prostate, MR images can be used in addition to the CT images, because of their good soft-tissue contrast. In current clinical practice at the University Medical Center Utrecht a *manual* delineation of the prostate is made, based on the CT and MR scans, which is a labour-intensive task. Figure 1 shows two example MR slices together with their manual delineations. The images are acquired with a Philips 3T scanner. The MR protocol, a balanced FFE gradient echo sequence, was optimised for visibility of the prostate and rectum.



(a) The base of the prostate (b) The seminal vesicles

Fig. 1. Two example MR slices, zoomed in on the region of interest, with manually delineated prostate (white line).

In this paper, we investigate an automatic segmentation method based on *atlas matching*. The atlas consists of a set of manually labelled images. All atlas images are matched to the patient's image that should be segmented. Then, a selection is made of the atlas images that match well to the patient's image. Finally, the propagated manual segmentations of the selected atlas images are combined into a single segmentation of the patient's image. Multiple atlas images are used, instead of a single image, to account for the large anatomical variability between subjects and for the differences in bladder and rectum filling. Note that the seminal vesicles are considered part of the prostate.

The segmentation method described above is similar to the work of Rohlfing et al. [1]. The difference lies in the atlas selection stage that is added here, which has a small, but significant, positive effect on the segmentation quality. The nonrigid registration method consists of three steps. Firstly, an affine registration is performed, secondly, a nonrigid B-spline based method, and, finally, a nonrigid registration of a region of interest. We show that the last step significantly improves the results for our application.

2. METHOD

The method consists of three stages: registration of all atlas images to the patient image, selection of atlas images that match well to the patient image, and combination of the de-

* Corresponding author: stefan@isi.uu.nl.

formed atlas label images. A set of N accurately labelled images, which serve as an atlas, are assumed to be available. The ith image in this atlas set is referred to as $A_i(\boldsymbol{x})$. The corresponding label image is called $L_i(\boldsymbol{x})$, a binary image, where 'ones' represent prostate tissue and 'zeros' everything else. The patient's image to be segmented is denoted by $P(\boldsymbol{x})$. The goal of the automatic segmentation method is to produce a label image $\hat{L}_P(\boldsymbol{x})$ that accurately defines the prostate of the patient. Ideally, this label image should be equal to a manual segmentation $L_P(\boldsymbol{x})$ created by a radiation oncologist.

2.1. Registration

In the registration stage, each atlas image A_i is matched to the patient image P. A coordinate transformation $T_i(\boldsymbol{x})$ is estimated that maximises the similarity of P and the deformed atlas $A_i \circ T_i$. We use *mutual information* as the measure of similarity [2].

The registration is performed in three steps. Firstly, global pose differences are compensated for by an affine registration. After that a nonrigid registration is performed, using a coordinate transformation that is parameterised by cubic B-splines [3]. Finally, a second nonrigid registration is performed, where the mutual information is evaluated only on a region of interest around the prostate. This forces the registration algorithm to 'concentrate' on the most important region. The region of interest is defined on the atlas image by a 20 times repeated dilation of the atlas label image L_i using a 3×3×3 kernel approximating a sphere. The experiments described in this paper are performed both with and without the last registration step, in order to evaluate its effect on the segmentation results.

In each step a four-level multiresolution scheme is employed. A Gaussian image pyramid is used for the image data. The nonrigid registrations are performed using a B-spline control point spacing of 64, 32, 16, and 8 mm, for each resolution respectively. The transformation that maximises the similarity measure is estimated by an iterative stochastic gradient descent routine.

2.2. Atlas selection

A selection is made of the registered atlas images that are most similar to the patient's image. Similarity is again measured by the mutual information. The measure is computed on the region of interest that was used in the last step of the registration. An atlas A_i is selected if it satisfies the following criterion:

$$\frac{MI(P, A_i \circ T_i)}{\max_j MI(P, A_j \circ T_j)} > \alpha, \qquad (1)$$

where MI denotes the mutual information and $0 \leq \alpha < 1$ is a tunable parameter. A value of 0 means that all atlas scans are included in the selection. A value approaching 1 implies that only the atlas scan with the highest similarity measure is used. In [1] only these extremum settings are investigated.

The set of atlas image indices selected in this stage is called $\mathcal{A}_P$. The subscript indicates that this set can be different for each patient image.

2.3. Label image propagation

The deformed label images of the atlas set $\mathcal{A}_P$ are combined into a single segmentation $\hat{L}_P(\boldsymbol{x})$ of the patient image. This is done by averaging the deformed label images and thresholding the result at a value of 0.5. This is equivalent to a majority voting rule.

3. EXPERIMENTS

The proposed automatic segmentation method is evaluated using a set of 38 MR scans. The scans originate from eight volunteers and were made in the context of another study. Seven volunteers were scanned five times, one volunteer was scanned three times. The time between two scans was at least one day, and the volunteers were asked to try to vary the content of their rectum and bladder, to get as much variety between the scans as possible. The scans have a dimension of 512×512×90 voxels of size 0.49×0.49×1.0 mm. Manual segmentations, made by an experienced observer and approved by a radiation oncologist, are available for each scan. The seminal vesicles were marked separately, which allows us to evaluate the segmentation method on the whole prostate or on the prostate excluding the seminal vesicles. Unless mentioned otherwise, results are presented for segmentation of the prostate including the seminal vesicles.

To reduce computation time and to discard areas of severe intensity inhomogeneities caused by magnetic field inhomogeneities, the scans are manually cropped to a rectangular region roughly encompassing the prostate, bladder, and rectum. This is the only step in the algorithm that requires user intervention. The image intensities are brought into the same range by intensity rescaling and histogram equalisation. Histogram equalisation ensures that not all image intensities fall within only a few histogram bins during computation of the mutual information.

Out of the 38 available scans, 16 scans are manually selected to form the atlas. From each volunteer two scans are included in the atlas. The 22 remaining scans serve as a test set, on which we evaluate the segmentation algorithm. Atlas images that originate from the same person as the test image are excluded from the atlas set, which results in an atlas of 14 images that is used for each test scan.

Evaluation is performed by comparing the automatically generated segmentations with the manual segmentations. A well-known measure of segmentation overlap is the Dice similarity coefficient [4]:

$$DSC(L_P, \hat{L}_P) = \frac{2|L_P \cap \hat{L}_P|}{|L_P| + |\hat{L}_P|}, \qquad (2)$$

where $|X|$ denotes the number of labelled voxels in X. To give an indication of the sensitivity of the DSC, we have dilated the manual segmentations of the test scans using a spherical 3D kernel with a radius of 1, 2, and 3 pixels, and we computed the DSC values of the manual segmentation and their dilated versions. This yielded median DSC values of 0.91, 0.83, and 0.76, respectively. In [5], the intraobserver reproducibility of manually delineating the prostate's peripheral zone on T2-weighted MR images is investigated. A mean similarity coefficient of 0.883 is reported for 1.5T preoperative scans, and 0.838 for 0.5T scans. Interobserver reproducibility for delineations of the prostate in CT images is measured in [6], where a median DSC of 0.82 is reported.

The DSC does not provide insight in the spatial distribution of the segmentation errors. To visualise the distance between the boundaries of two approximately circular-shaped objects, a spherical coordinate mapping is sometimes used [6, 7]. We use this technique to assess the quality of the corpus segmentation (prostate *without* seminal vesicles). The position on the boundary of the corpus is parameterised by two angles, θ and ϕ, where the centre of mass of the manual segmentation serves as the centre of rotation. The shortest Euclidian distance between the manual and automatic segmentation boundaries can be plotted as a function of θ and ϕ, for each test scan. The results for the 22 test scans are summarised by plotting the first quartile, median, and third quartile as a function of θ and ϕ.

4. RESULTS

The DSC values are computed for all 22 test scans, with different values of α. Figure 2 summarises the results in a box-and-whisker plot. Each box-and-whisker visualises the distribution of DSC values for a specific value of α. Results for α between 0 and 0.4 are not shown since they are identical to $\alpha = 0$. A Wilcoxon signed rank test is performed to assess the influence of α. Significant differences ($p < 0.05$) with respect to $\alpha = 0$ and $\alpha = 0.99$ are indicated above the graph by 'a' and 'b', respectively. Based on Figure 2, we conclude that a value of α in the range 0.6-0.85 gives the best results. For further evaluation, we select $\alpha = 0.7$.

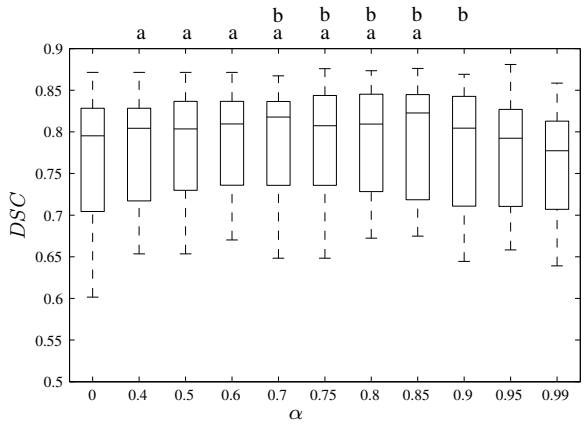


Fig. 2. A box-and-whisker plot of the Dice similarity coefficients for different values of α. The box shows the first quartile, the median and the third quartile of the data. The whiskers display the total range of the data. Significant differences with respect to $\alpha = 0$ and $\alpha = 0.99$ are indicated by 'a' and 'b', respectively.

The experiments are repeated *without* the additional registration on the region of interest, see Sec. 2.1. Figure 3 compares the DSC values for each test scan, for $\alpha = 0.7$. The median DSC value is 0.75 for the experiments without the registration on the region of interest. With the additional registration step, a median of 0.82 is achieved. The Wilcoxon test on the differences results in a $p < 0.0001$.

Figure 4 shows the spatial distribution of the segmentation errors, again for $\alpha = 0.7$. A map of the surface is given in Figure 4(a). First quartile, median, and third quartile of the distance between the manually and automatically determined corpus boundaries are shown in the remaining figures. From the figures it is evident that the largest errors occur at the border between the corpus and the vesicles. In Figure 4(b) it can be seen that in 75% of the test scans the error at the vesicles was larger than 2.5 mm approximately. At the boundary between prostate and rectum smaller errors are made. Figure 4(c) and 4(d) show that in 50% of the test cases the distance remains below 1 mm, and in 75% of the cases below 2.5 mm approximately. The edge between prostate and bladder is determined with similar accuracy. Highest accuracy is found at the left and right side of the base of the prostate (θ around 0°, 180°, and 360°, ϕ from 90° to 135°). The tip of the apex (ϕ close to 180°) is less accurately determined.

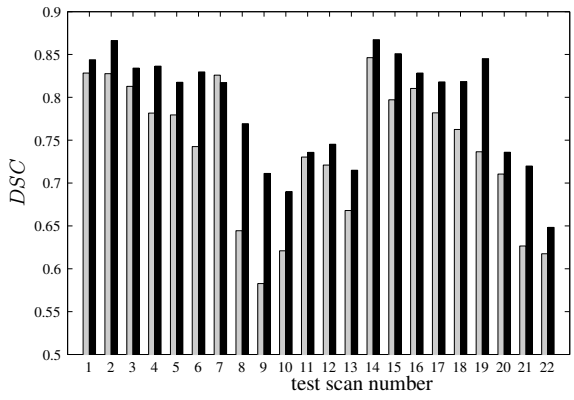
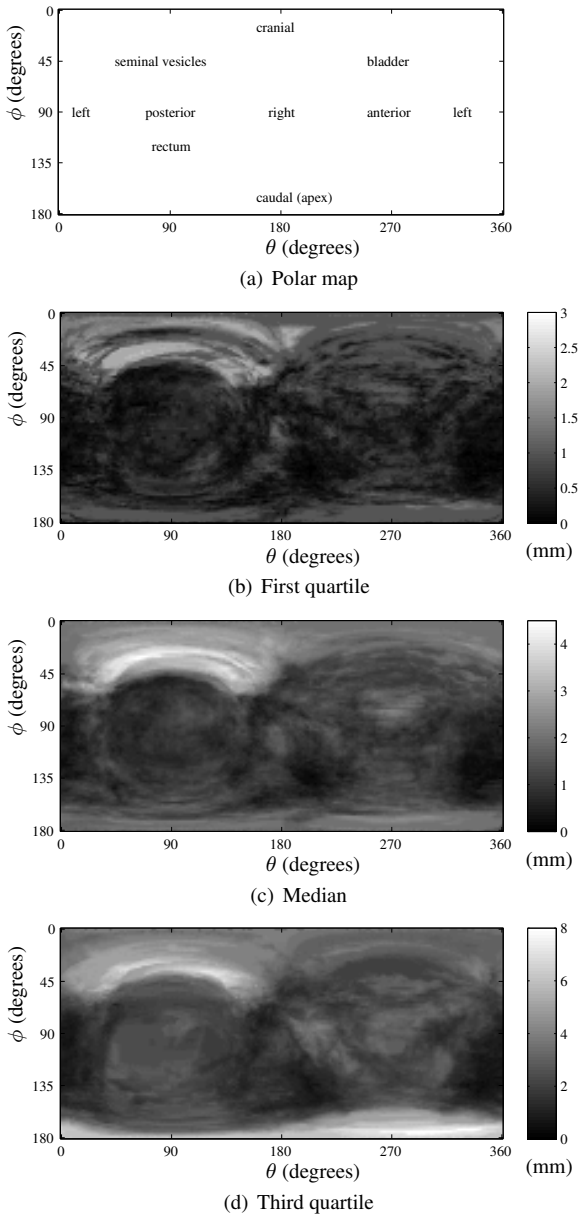


Fig. 3. The DSC values for each test scan. Grey bars correspond to the results obtained *without* the registration on the region of interest; black bars present the results for the complete registration method. In both cases $\alpha = 0.7$ was used.



(a) Polar map

(b) First quartile

(c) Median

(d) Third quartile

Fig. 4. A polar map of the spatial error distribution of the prostate segmentation, *excluding* the seminal vesicles.

5. CONCLUSION

An automatic prostate segmentation method for pelvic MR images has been investigated in this work. The method is based on matching of manually segmented atlas images. To account for the large variability in shape, multiple atlas images are combined.

Evaluation has been performed on volunteer data. The method was tested on 22 test scans, using an atlas consisting of 14 images. The segmentations obtained by the automatic segmentation method are compared to manual segmentation by means of the Dice similarity coefficient and by plots of the spatial error distribution. A median Dice similarity coefficient of 0.82 is achieved. The atlas selection stage and the nonrigid registration step that takes into account the region of interest are both shown to improve the results significantly. In the spatial error distribution plots, it can be seen that on a large part of the prostate surface the segmentation errors remain below 2-3 mm for 75% of the test scans. The most serious segmentation errors are made in the seminal vesicles, which is confirmed by visual inspection. Imaging artefacts caused by air in the rectum also influence the results negatively in some cases.

For good performance of the proposed algorithm, it is important that the atlas contains sufficient variation. Therefore, future work includes optimisation of the atlas composition. Furthermore, the algorithm will be evaluated on patient data.

6. ACKNOWLEDGMENTS

The authors thank Ellen Kerkhof for providing the manual segmentations.

7. REFERENCES

[1] T. Rohlfing et al., "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains," *NeuroImage*, vol. 21, no. 4, pp. 1428–1442, 2004.

[2] P. Thévenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2083–2099, 2000.

[3] D. Rueckert et al., "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[4] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[5] K. H. Zou et al., "Statistical validation of image segmentation quality based on a spatial overlap index," *Academic Radiology*, vol. 11, pp. 178–189, 2004.

[6] M. Foskey et al., "Large deformation three-dimensional image registration in image-guided radiation therapy," *Physics in Medicine and Biology*, vol. 50, pp. 5869–5892, 2005.

[7] C. Rasch et al., "Definition of the prostate in CT and MRI: a multi-observer study," *Int. J. Radiation Oncology Biol. Phys.*, vol. 43, pp. 57–66, 1999.