# An automatic multi-atlas based prostate segmentation using local appearance-specific atlases and patch-based voxel weighting

Qinquan Gao[1],Daniel Rueckert[1], Philip "Eddie" Edwards[2]

[1] Department of Computing
Imperial College London, United Kingdom
q.gao10@imperial.ac.uk
[2] Department of Surgery and Cancer
Imperial College London, United Kingdom
Eddie.Edwards@imperial.ac.uk

**Abstract.** Prostate segmentation facilitates prostate cancer detection, in conjunction with other parameters, can help predict the pathologic stage of disease. Segmentation of anatomy may also help to improve the outcome of robotic-aided laparoscopic prostatectomy (RALP) by augmented reality image guidance. In this paper, we present a fully automated multi-atlas segmentation pipeline for multi-center and multi-vendor MRI prostate segmentation using a multi-atlas approach with local appearance-specific voxel weighting. Segmenting prostates with a large variation of shape and intensity still remains a significant challenge. In this work, the atlases with the most similar global appearance are classified into the same categories. Sum-of-square local intensity difference after affine registration is used for atlas selection and after non-rigid registration, a local patch-based atlas fusion is performed using voxel weighting based on the local patch distance. Such multi-atlas segmentation is a widely used method in brain segmentation. We thoroughly evaluated the method on 50 training images by performing a leave-one-out study. Dice coefficient and overlap rate are used to quantify the difference between the automatic and manual segmentation. Compared to the manual gold standard segmentation, our proposed method produce favorable outcomes in these highly variable data sets, with an average dice coefficient $0.8467 \pm 0.0435$. The result shows that the algorithm presented could be used to delineate prostate from diverse MRI images, and therefore is available for a variety of clinical applications.
Keywords: Multi-atlas segmentation, atlas segmentation, image registration, patch-based segmentation, local weighting

## 1 Introduction

Accurate segmentation and location of the prostate is crucial for prostate cancer detection and staging, surgical planning and image-guided robotic-aided laparoscopic prostatectomy (RALP) with augmented reality (AR) [2]. Currently, the majority of segmentation is done by well-trained radiologists based on the anatomical knowledge with the appearance of the scans to identify the relevant physical structures. This is very time consuming to achieve manually, especially for a large number of segmentation. Therefore, there is a pressing need for fast, automatic segmentation methods for clinical applications. There is much previous work based on the well known statistical shape model and probabilistic atlas priors learned from training data [5].

In a recent study, multi-atlas segmentation provides the best accuracy compared to a number of algorithms for the segmentation of subcortical structures [1]. However, as there are large shape variations and intensity differences on prostate scans with different acquisition protocols, it still remains challenging to provide a robust and fully automated segmentation across various scans. In order to tackle this problem, in multi-atlas based segmentation, the atlas in the database which is the most similar to the query image is used [3]. Several methods have been investigated and compared to improve this selection by Lotjonen [6]. Also, during the multi-atlas fusion process, it is reasonable to expect the atlases whose reference images are more similar to the target image should contribute more [10].

In this paper, aiming for a more automatic and accurate segmentation, we introduce a multi-atlas segmentation using local appearance-specific atlases, which is more robust to inter-subject variation. The atlas database was classified into different categories and the most similar atlases in the region of interest are selected for multi-atlas registration by comparing the sum of square intensity distance after affine registration. The select atlases are non-rigidly aligned to the target image, and a patch-based local voxel weighting strategy is introduced, which was recently proposed for use in patch based brain segmentation [4]. The weighting of mapping agreement from atlas to target is proposed to make the final atlas fusion more robust. The proposed method was evaluated on the 50 training data, which is a representative set of the types of MR images acquired in a clinical setting.

## 2 Method

In multi-atlas based segmentation, the most similar atlases create more accurate transformations to the target image, therefore getting better label estimation. In this paper, aiming for a robust and accurate segmentation of multi-center and multi-vendor MRI prostate scans, appearance-specific atlas selections and a patch-based local weighting strategy for atlas fusion are introduced. An initial denoising and intensity inhomogeneity is performed on all images. Atlases are classified into two categories: normal MRI scans $A_n$ and scans taken with a transrectal coil $A_m$. This is easily achieved by examining the intensity variation around the rectum since the transrectal coil produces significant physical distortion but also has a characteristic bright appearance in the local region near the coil itself. During the segmentation, the sub-atlas database, whose atlas appearance is closest to the new target, is chosen as the initial atlas database.

After that, the top N similar atlases are further chosen for atlas recitation by measuring intensity difference in the region of interest around prostate. After all the selected atlas registered non-rigidly to a target image, the resulting transformation is used to propagate the anatomical structure labels of the atlas into the space of the target image. Finally, based on patch-based voxel weighting, the label that the majority of all warped labels predict for each voxel is used for the final segmentation of the target image. The pipeline of multi-atlas segmentation of the prostate is divided into the following parts: atlas database construction, appearance-specific atlas selection, multi-atlas pairwise registration, and atlas propagation and fusion with local voxel weighting, shown in Figure 4.

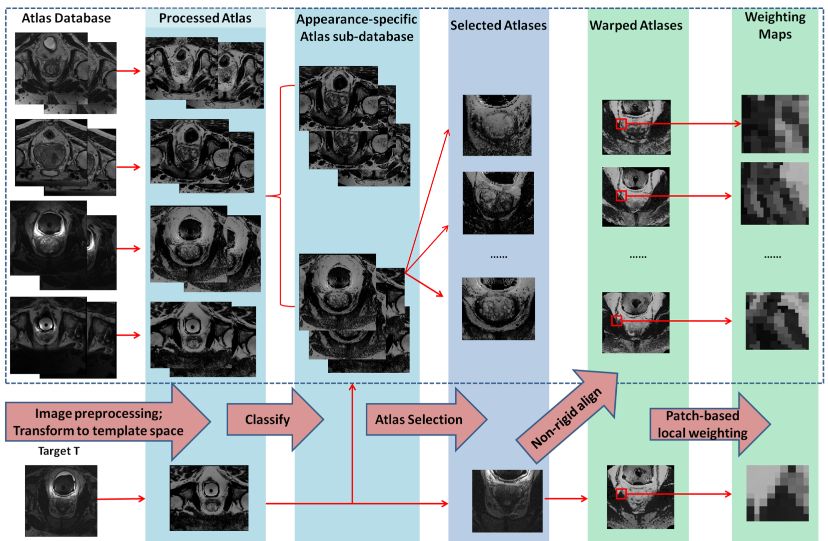


**Fig. 1.** The pipeline of multi-atlas segmentation of the prostate: According to the specific segmentation problem, after image preprocessing, the atlases are first classified into two categories based on the local appearance around rectum: normal MRI scans and scans with transrectal coil. For an unseen new scan segmentation, the sub-atlas database with the closed appearance distribution is automated chosen as segmentation priors. Then, comparing the intensity variation in the region of interest around the prostate with the target, the most similar N atlases are chosen for atlas registration. After aligning all the selected atlases with the target image non-rigidly, a patch-based local weighting are applied to implementing the atlas fusion, getting each voxel's label estimation.

### 2.1 Image preprocessing for atlas database construction

In the training set, 50 transversal T2-weighted MR images of the prostate are provided, which are representative of the types of MR images acquired in a clinical setting. The data is multi-center and multi-vendor and has different acquisition protocols, such as differences in slice thickness and the presence of an endorectal coil in some images, which makes the intensity of images inhomogeneous. During atlas database construction, variability caused by image formation is minimized by performing denoising, an inhomogeneity correction, and an inter-subject intensity normalization. To remove the intensity bias introduced by the Racian nature of noise, a Rician adaption of non-local means [11] was used for denoising. Then, the well-known N3 intensity nonunifornity correction [9] was applied to ensure that each tissue type should have a same intensity. Finally, all the atlases are transformed into the template space by affine registration. The intensity of images were then normalized together in the template space by performing the method proposed by Nyul and Udupa [7]. This makes the contrast and luminance of each tissue type more consistent across the training images in the database. After all these procedures, an atlas database was constructed with the preprocessed MR images and their corresponding manual segmentation, representing as $\sum_N A(I_i, L_i)$.

### 2.2 Local appearance-specific atlas selection

In the atlas database, the endorectal coil influence makes the appearance of the scan significantly difference in both shape and intensity, which easily leads to faulty registrations. In order to tackle this problem, firstly, by comparing the intensity difference around the region of rectum, the atlas database is classified into two sub-database, representing as $A = \{\sum_n A(I_i, L_i), \sum_m A(I_j, L_j)\}$: atlases with normal MRI scans $\sum_n A(I_i, L_i)$ and atlases taken with an endorectal coil $\sum_m A(I_j, L_j)$. The most suitable sub-database will be automatically chosen for a new unseen target segmentation.

Then, in the chosen sub-database, the top N atlases with the most similar appearance around the prostate region are selected for final multi-atlas segmentation. Atlas selection in these two steps are base on the L2 norm: the sum of squared intensity differences $\triangle(A_i, \mathcal{L})$ between atlas $A_i$ and target image $\mathcal{L}$ , defined over a region of interest $R$, measuring local image appearance:

$$\triangle(\mathcal{L}, A_i) = \sum_{j \in N} ||\mathcal{L}(x_j) - A_i(x_j)||^2 \tag{1}$$

### 2.3 Patch-based voxel weighting

During the multi-atlas fusion process, would like the atlases whose reference images are more similar to the target image to contribute more than those reference images are less similar. Also, the accuracy of the transformation from atlas to target is crucial for more accurate label estimations. In this paper, we propose a more robust and improved weighting strategy for atlas label fusion that combines the mapping agreement weighting with the patch-based weighting for each voxel label. This weight is based on the similarity of a patch surrounding voxel

$x_i$ and patches in a local neighborhood of all non-rigid aligned atlas images $A_s$. Therefore, the segmentation problem can be denoted as follows:

$$\Gamma(x_i) = \frac{\sum_{s=1}^{N} w_{T(A_s \to \mathcal{L})} \sum_{j \in V} w(x_i, x_{s,j}) L_{s,j}}{\sum_{s=1}^{N} w_{T(A_s \to \mathcal{L})} \sum_{j \in V} w(x_i, x_{s,j})} \tag{2}$$

where $\Gamma_i$ is the estimated label for voxel i, $L_{s,j}$is the label by expert to voxel $x_j$in atlas $s$ , $V$ is the search window size, $w(x_i, x_{s,j})$ is the weight assigned to label $L_{s,j}$ by patches comparison, from patch surrounding $x_i$ to that surrounding $x_{s,j}$, as following:

$$w(x_i, x_{s,j}) = exp^{\frac{-\triangle(P_i^{\mathcal{L}}, P_j^{A_s})}{h}} \tag{3}$$

where $h$ is a decay parameter, which is set to the minimum patch distance as proposed in [4].

The other weight $w_{T(A_s \to \mathcal{L})}$ is defined by the accuracy of the non-rigid mapping from each atlas to the target image. This is measured by comparing the intensity between the warped atlas $T(A_s \to \mathcal{L})A_s^L$and target image $\mathcal{L}^L$ under the segmented label region using normalized matural information, denoted as:

$$w_{T(A_s \to \mathcal{L})} = NMI(T(A_s \to \mathcal{L})A_s^L, \mathcal{L}^L) \tag{4}$$

It is a crucial term to reduce the negative influence by mis-registration from parts of atlases to target non-rigidly, which make final segmentation more robust.

## 3 Experiments and Results

The whole segmentation pipeline is implemented in C++ and the evaluation is run on quad 3.20GHz CPUs and 1GB of global memory, using parallel programming to calculate the patch weithing map and label fusion for each voxel simultaneously. The proposed method was evaluated on the 50 training MRI images including a transversal T2-weighted MR image of the prostate. For each training data, manual segmentation are provided.

A leave-one-out study has been implemented based on each of the training scans using the remaining 49 images as the atlas database. After initial denoising and inhomogeneity correction, the target image was transformed into the template space by affine registration, followed by intensity normalization compared to the histogram of the template image. Then, an appearance-specific atlas database was preselected. In the sub-database, the top 10 most similar atlases are selected by comparing the local prostate appearance according to equation 1. With these selected atlases, a pairwise rigid and affine registration was applied to transform atlas images to the target image in the template space, followed by a three-level non-rigid registration using free-form deformation [8], with b-spline control point spacings of 20 mm , 10mm and 5 mm.

In the multi-atlas fusion, the patch-based voxel weighting strategy in Eq.2 is applied to get the segmentation estimation, with a patch size of $5 \times 5 \times 5$ and search window size $9 \times 9 \times 9$.

For evaluation, we used the following metric compared with gold standard expert segmentation, and also, we compared our propose method with directed atlas fusion using majority voting:

1) Dice Coefficient (DC) : $\mathrm{DC} = 2\frac{A \cap B}{A+B}$.
2) Volumetric overlapping accuracy (VOA):
$\mathrm{VOA} = \min(\frac{A \cap B}{A} \times 100\%, \frac{A \cap B}{B} \times 100\%)$.

Table 1 shows the average value of these metric with their standard deviation among all the 50 MRI scan segmentation compared with manual segmentation. Final segmentation result for the prostate from MRI images exhibit an average DC of $0.8467 \pm 0.0435$ and an average VOA of $0.8259 \pm 0.0630$. In a fully automated way , the most traning data can be well segmented, but there are still a few examples leading to fault segmentation, as shown in Figure 2. This makes the average value of dice coefficient to be smaller.

| Segmentation Method | DC (average+sd) [%] | VOA (average+sd) |
|---|---|---|
| Directly atlas fusion | $0.82318 \pm 0.0456$ | $0.8015 \pm 0.0555$ |
| Our proposed method | $0.8467 \pm 0.0435$ | $0.8259 \pm 0.0630$ |

**Table 1.** Average metric compared with gold standard, including: volumetric overlapping accuracy (VOA) and dice coefficient (DC).

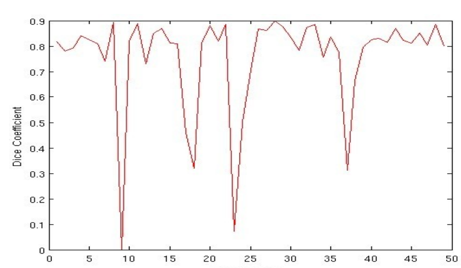


**Fig. 2.** Dice coefficient of each training image segmentation

As can been seen in Table1, the atlas fusion using patch-based voxel weighting outperforms that directly applying majority weighting, with higher dice metric and overlapping rate. Though the increase in the Dice metric is not huge, it is significantly produce better segmentation with smooth boundaries of the segmented prostate surface, as can be seen in Figure 3.

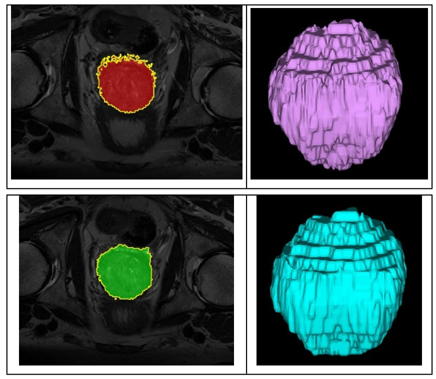

**Fig. 3.** The prostate segmentation result using multi-atlas segmentation with majority voting (Top) and with patch-based voxel weighting (Bottom)

## 4 Discussion and Conclusion

In this paper, we propose a novel and automatic multi-atlas segmentation using local appearance-specific atlases and patch-based local weighting. Among all the 50 training data with diverse intensity variation and prostate shape, most of the segmentation performed quite well in a fully automatic way, which could be useful in this clinical application. However, in some of the segmentation estimations, a small segmentation error will cause significant dice coefficient changes, since there is only a small part of the prostate structure contained in the MRI scan. This makes registration very difficult and there are matches to the wrong structures, which makes the average dice metric to be smaller, compared to the result for images of the whole prostate. In order to tackle this problem, landmarks could be introduced manually to get a better initial transformation.

There is still some work to be done to achieve more accurate segmentation. We aim to introduce the CUDA programming to speed up the non-rigid registration, improving the efficiency of multi-atlas segmentation. And also, patch learning could also be introduced to make more convincing weighting of each voxel during the fusion process. We are also investigating methods to improve registration in scans where there is only small overlap with the region of interest.


## Acknowledgments

This research was funded by Cancer Research UK under project A8087/C24250. We are grateful to the radiology and theatre staff at the Imperial College Healthcare NHS Trust for their help and cooperation throughout this project. The work has ethical approval from the London - Dulwich research ethics committee.

# Bibliography


[1] K O Babalola, B Patenaude, P Aljabar, J Schnabel, D Kennedy, W Crum, S Smith, T F Cootes, M Jenkinson, and D Rueckert. Comparison and evaluation of segmentation techniques for subcortical structures in brain MRI. *MICCAI 2008*, 11(Pt 1):409–16, January 2008.

[2] Daniel Cohen, Erik Mayer, Dongbin Chen, Ann Anstee, Justin Vale, Guang-Zhong Yang, Ara Darzi, and Philip Edwards. Augmented reality image guidance in minimally invasive prostatectomy. In *Prostate Cancer Imaging. Computer-Aided Diagnosis, Prognosis, and Intervention*, volume 6367, pages 101–110. 2010. ISBN 978-3-642-15988-6.

[3] Olivier Commowick and Grégoire Malandain. Efficient selection of the most similar image in a database for critical structures segmentation. *MICCAI 2007*, 10:203–10, January 2007.

[4] Pierrick Coupé, José V Manjón, Vladimir Fonov, Jens Pruessner, Montserrat Robles, and D Louis Collins. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *NeuroImage*, 54(2):940–54, January 2011. ISSN 1095-9572.

[5] S Ghose, A Oliver, R Marti, X Llad, J Freixenet, J C Vilanova, F Meriaudeau, De Bourgogne, Lei-umr Cnrs, and Le Creusot. A probabilistic framework for atutomatic Prostate segmentation with a statistical model of shape and appearance. pages 725–728, 2011.

[6] Jyrki Mp Lötjönen, Robin Wolz, Juha R Koikkalainen, Lennart Thurfjell, Gunhild Waldemar, Hilkka Soininen, and Daniel Rueckert. Fast and robust multi-atlas segmentation of brain magnetic resonance images. *NeuroImage*, 49(3):2352–65, March 2010.

[7] L G Nyúl and J K Udupa. On standardizing the MR image intensity scale. *Magnetic resonance in medicine*, 42(6):1072–81, December 1999. ISSN 0740-3194.

[8] D Rueckert, L I Sonoda, C Hayes, D L Hill, M O Leach, and D J Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE transactions on medical imaging*, 18(8):712–21, August 1999. ISSN 0278-0062.

[9] J G Sled, a P Zijdenbos, and a C Evans. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE transactions on medical imaging*, 17(1):87–97, February 1998. ISSN 0278-0062.

[10] Hongzhi Wang, Jung Wook Suh, Sandhitsu Das, John Pluta, Murat Altinay, and Paul Yushkevich. Regression-based label fusion for multi-atlas segmentation. *CVPR 2011*, pages 1113–1120, June 2011.

[11] Nicolas Wiest-daesslé, Sylvain Prima, Pierrick Coupé, Sean Patrick Morrissey, and Christian Barillot. Rician noise removal by non-Local Means filtering for low signal- to-noise ratio MRI : applications to DT-MRI HAL-AO Author Manuscript. pages 1–12, 2008.