

## Segmentation of pelvic structures for planning CT using a geometrical shape model tuned by a multi-scale edge detector

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 Phys. Med. Biol. 59 1471

(<http://iopscience.iop.org/0031-9155/59/6/1471>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.20.25.168

This content was downloaded on 12/03/2014 at 17:39

Please note that [terms and conditions apply](#).

# Segmentation of pelvic structures for planning CT using a geometrical shape model tuned by a multi-scale edge detector

Fabio Martínez<sup>1,2</sup>, Eduardo Romero<sup>1</sup>, Gaël Dréan<sup>2,3</sup>,  
Antoine Simon<sup>2,3</sup>, Pascal Haigron<sup>2,3</sup>,  
Renaud de Crevoisier<sup>2,3,4</sup> and Oscar Acosta<sup>2,3</sup>

<sup>1</sup> CIM&Lab, Universidad Nacional de Colombia, Bogota, Colombia

<sup>2</sup> INSERM, U1099, Rennes, F-35000, France

<sup>3</sup> Université de Rennes 1, LTSI, Rennes, F-35000, France

<sup>4</sup> Département de Radiothérapie, Centre Eugène Marquis, Rennes, F-35000, France

E-mail: [edromero@unal.edu.co](mailto:edromero@unal.edu.co)

Received 5 September 2013, revised 17 January 2014

Accepted for publication 23 January 2014

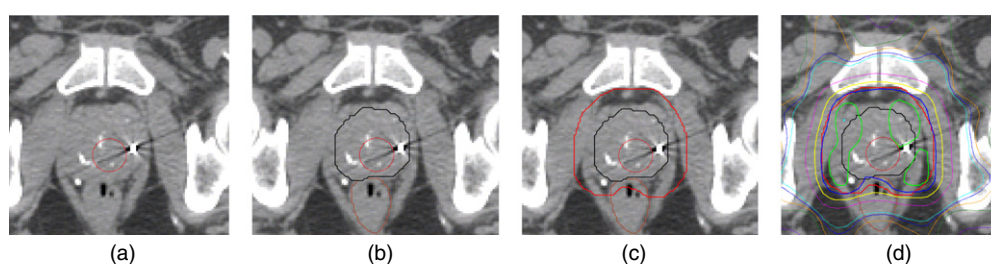
Published 5 March 2014

## Abstract

Accurate segmentation of the prostate and organs at risk in computed tomography (CT) images is a crucial step for radiotherapy planning. Manual segmentation, as performed nowadays, is a time consuming process and prone to errors due to the a high intra- and inter-expert variability. This paper introduces a new automatic method for prostate, rectum and bladder segmentation in planning CT using a geometrical shape model under a Bayesian framework. A set of prior organ shapes are first built by applying principal component analysis to a population of manually delineated CT images. Then, for a given individual, the most similar shape is obtained by mapping a set of multi-scale edge observations to the space of organs with a customized likelihood function. Finally, the selected shape is locally deformed to adjust the edges of each organ. Experiments were performed with real data from a population of 116 patients treated for prostate cancer. The data set was split in training and test groups, with 30 and 86 patients, respectively. Results show that the method produces competitive segmentations w.r.t standard methods (averaged dice = 0.91 for prostate, 0.94 for bladder, 0.89 for rectum) and outperforms the majority-vote multi-atlas approaches (using rigid registration, free-form deformation and the demons algorithm).

**Keywords:** CT segmentation, radiotherapy planning, prostate cancer, Bayesian models, geometrical multi-scale descriptors

(Some figures may appear in colour only in the online journal)



**Figure 1.** Example of planning dose on manually segmented CT. (a) Original CT scan, (b) prostate and rectum delineations, (c) PTV is obtained by adding margins to the prostate and (d) resulting 3D planning dose. The presence of fiducial markers may hamper segmentation.

## 1. Introduction

Prostate cancer is one of the most commonly diagnosed male cancers worldwide, with 190 000 new cases diagnosed in the USA in 2010 (American Cancer Society) and 71 000 new cases in France in 2011 (INCa 2011). External beam radiation therapy (EBRT) is a commonly prescribed treatment for prostate cancer which has proven to be efficient for tumor control (D'Amico *et al* 1998). EBRT uses high energy beams from multiple directions in order to deliver the dose within the patient tumor region to destroy the cancer cells. Modern treatment techniques offer improved treatment accuracy through better planning, delivery, visualization and the correction of patient setup errors. Segmentation plays a major role during the whole therapy. During a standard protocol for EBRT planning, computed tomography (CT) images are manually delineated to define not only the clinical target volume (CTV), prostate and seminal vesicles, but also the neighbouring organs at risk (OARs), for instance the bladder and rectum. The CTV is then expanded to constitute the planning target volume (PTV) for treatment. These spatial margins between the organs and the PTV, as depicted in figure 1, allow for uncertainties in delineation, patient setup, motion and organ deformations (Guckenberger and Flentje 2007, Bayley *et al* 2010). The segmentation allows the definition of dose constraints according to certain recommendations. The International Commission on Radiation Units and Measurements (ICRU) 50 and 62 reports, for example, define and describe several target and critical structure volumes that aid in the treatment planning process and may provide a basis for comparison of treatment outcomes. Thus the directions, strengths and shapes of the treatment beams are computed in a planning platform to comply with them, following a particular dose prescription. The delineated structures may also be used in later stages and for different purposes during the treatment, such as the computation of cumulated dose when image guided radiotherapy (IGRT) is used or for toxicity population studies (Heemsbergen *et al* 2010, Söhn *et al* 2012, Acosta *et al* 2013).

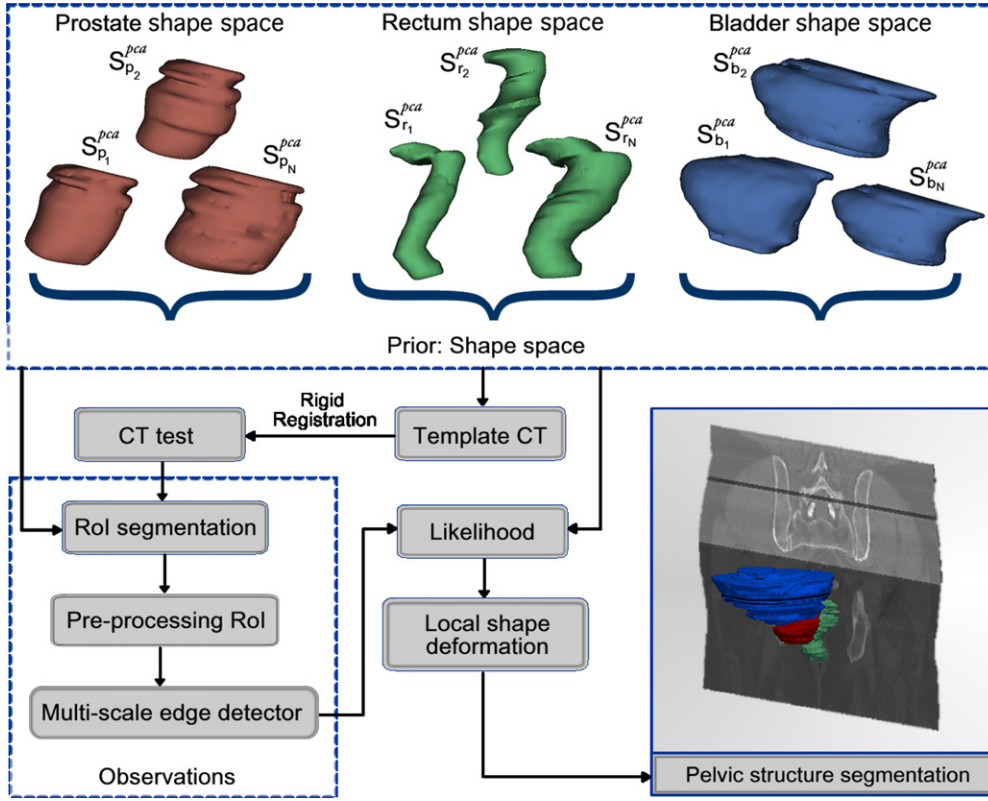
In current RT planning, the CT is still used to perform the segmentation since the dose computation relies on the CT electron density information. However, the CT offers a poor soft tissue contrast and therefore segmenting pelvic organs is highly time consuming (between 20–40 min to delineate each), and prone to errors, especially in the apical and basis regions (Collier *et al* 2003). Previous studies have reported variations in prostate delineation of about 20.60% of the organ volume (Gual-Arnau *et al* 2005). These uncertainties lead to large intra- and inter-observer variabilities and may impact the treatment planning (Fiorino *et al* 1998) and dosimetry (Al-Qaisieh *et al* 2002, Beckendorf *et al* 2002). Hence, reliable, efficient and reproducible (semi)-automatic segmentation techniques are required in treatment planning.

Several difficulties hamper automatic segmentation methods in this context: (i) the poor border contrast between adjacent organs, (ii) the high intra- and inter-individual shape variability, (iii) the inhomogeneities in the amount of bladder and rectum filling (Langen and Jones 2001) and (iv) the presence of fiducial markers used to guide dose delivery. Common automatic methods for organ segmentation of pelvic structures include deformable models (María Jimena *et al* 2007, Bystrov *et al* 2012), atlas-based methods (Acosta *et al* 2010, 2011, 2014) or statistical shape models (Hodge *et al* 2006, Pekar *et al* 2004). Atlas-based approaches use prior learning, not only for obtaining a final contour, but also to provide initial organ positions for further segmentation algorithms. In atlas-based methods, a pre-computed segmentation or prior information (i.e. probability map) in a template space is propagated to the image to be segmented via registration (Rohlfing *et al* 2004, Evangelia *et al* 2005, Acosta *et al* 2011, Rohlfing and Maurer 2005). Although atlas-based approaches may provide prior structural information, a high inter-individual variability and registration errors may lead this method to fail. Multi-atlas approaches can partly overcome some of these difficulties by selecting the most similar organ of interest among a large database, but the definition of proper similarity metrics between the available atlases and the individual query is still an open issue (Acosta *et al* 2011, Wu *et al* 2007, Acosta *et al* 2011). Statistical shape models may help to overcome atlas-based segmentation problems by introducing prior shapes, which in a further step may be locally adjusted (Hodge *et al* 2006, Pekar *et al* 2004). In general, these approaches define a cost function, encoding similarities that match a prior shape/appearance knowledge with a particular individual (test image) (Heimann and Meinzer 2009, Qianjin *et al* 2009). These approaches are relatively stable and robust to noise and voxel variations, however their accuracy may be limited by the set of features selected to represent the image test and the particular definition of the match function. Examples of prostate segmentation following this approach appear in Mikael *et al* (2005), where the segmentation problem was addressed without shape constraints by allowing flexibility to individual local variations, with some issues, however, in regions with poor contrast. More recently, Chen *et al* (2011a) proposed a statistical method that matches the prior shape information under a Bayesian hypothesis, incorporating anatomical constraints but without exploiting spatial information. Other approaches include learning strategies to segment the prostate, but within the context of intra-individual segmentations where previous organ delineations are required (Li *et al* 2012).

In this paper, we propose a novel method for automatically segmenting pelvic structures from CT scans based on a statistical shape model within a Bayesian framework and with local spatial variations. The proposed approach is appropriate for different types of organ shapes, as well as being stable to variations in contrast and bladder and rectum filling conditions. A method for removing CT artefacts was implemented which increases a contour's reliability using an adapted non-local filter, which was able to overcome the issues with fiducial markers. Since the prior is built from different available shapes, the approach is robust to large inter-individual variability. The likelihood matching includes a geometrical and topological analysis which allows one to set an optimal correspondence between a prior 3D shape model and the considered CT. Such a prior is built from an actual population while a multi-scale edge descriptor allows for the mapping between the prior and the CT observations. The yielded 3D organ segmentation is spatially coherent and voxels are connected together. In the remainder of this paper we first describe the method, then validate the accuracy of our approach on real CT data and make a comparison with baseline atlas-based segmentation methods.

## 2. Methodology

The proposed method consists of several stages as depicted in figure 2. Let  $\hat{S}_o$ , the estimated organ shape (prostate, bladder or rectum) that optimizes a maximum *a posteriori*, under a



**Figure 2.** Proposed method for 3D segmentation. First, a shape space organ is built (PCA). The template is rigidly registered to the CT to be segmented followed by an automatic extraction of RoIs for pre-processing and multi-scale edge detection. A likelihood function matches the most similar PCA shape with the detected edges to finally be locally adjusted.

Bayesian framework be

$$\hat{S}_o = \max_{\hat{S}_o} \arg[P(\hat{S}_o | S_{o_1}^{pca}, S_{o_2}^{pca}, \dots, S_{o_N}^{pca})] \quad (1)$$

where  $\{S_{o_1}^{pca}, S_{o_2}^{pca}, \dots, S_{o_N}^{pca}\}$  is a collection of shapes (the shape space) of the organ  $o$ . The reference template serves as the coordinate framework of the principal component analysis (PCA) geometrical shape model and corresponds simply to a random selection of a single CT-image. The likelihood function (section 2.4), aimed at matching the most similar shape with the observations, extracts information from a multi-scale analysis in a region of interest (RoI) around the organs (section 2.3), after removing CT artefacts. Eventually, the most likely shape  $\hat{S}_o$  is locally deformed, driven by the multi-scale edge descriptor (section 2.5). These steps are described in the following sections.

### 2.1. Learning an organ shape model: the prior

A dimensionality reduction of the whole training data was firstly carried out by applying PCA (Sohn *et al* 2005) to the population of manually delineated organs (Lorenz *et al* 2000). Hence, a collection of shapes  $\{S_{o_1}^{pca}, S_{o_2}^{pca}, \dots, S_{o_N}^{pca}\}$  for each considered organ, was computed. A previous

rigid registration scales different volumes so that the number of slices is always the same for any organ. Then, a set of ‘360 landmarks’ was set in the polar space (a landmark per grade), using the centroid of every organ as reference. Each shape contour is the parametric curve defined by a set of landmarks among the training data lying on the contour. In consequence, correspondences are always one-to-one and the number of points of any contour is always the same. Thus, the first two contour moments are then computed, namely the mean shape vector  $\bar{s} \in R^{3M}$  and the covariance matrix  $\rho \in R^{3M \times 3M}$ , are computed as  $\bar{s} = \frac{1}{N} \sum_{i=1}^n s_i$  and  $\rho = \frac{1}{N-1} \sum_{i=1}^n (s_i - \bar{s})(s_i - \bar{s})^T$ , where the vector  $(s_i - \bar{s})$  characterizes the organ deviation with respect to the mean shape. A conventional spectral analysis allows diagonalization of this covariance matrix, thereby finding the directional gains or *eigenmodes*. Each eigenmode defines a 3D vector field of the correlated organ inter-patient variability. Thus, the organ statistical samples were generated by deforming the mean shape with a weighted linear combination of the  $L$  dominating eigenmodes:

$$S_{o_l}^{\text{pca}} = \bar{s} + \sum_{l=1}^L c_l q_l \quad (2)$$

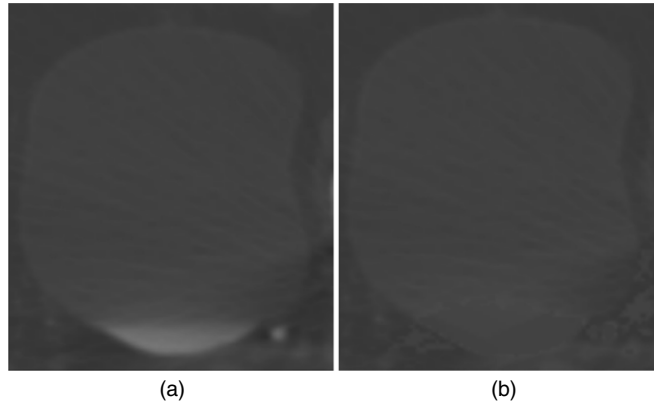
where the coefficients  $c_l$  follow a Gaussian distribution and the  $q_l$  are the eigenvectors or directions with a variance defined in the interval  $q_l \in \{+3\sqrt{\lambda}, \dots, -3\sqrt{\lambda}\}$ , accounting for 96% of the shape variability. This procedure was independently used for each organ, obtaining a family of shape models that were aligned with the previously chosen CT template. Examples of the different shapes obtained for each organ are shown in figure 2.

## 2.2. RoI extraction and pre-processing

During the segmentation of a given CT image, the previously computed model is rigidly registered toward the CT template from the training database, using a classical ‘block matching’ method (Ourselin *et al* 2001). Thus, a set of RoIs of size  $\{\bar{S}_o \pm \xi\}$ ,  $\bar{S}_o$  being the average organ size and  $\xi$  a tolerance value, allows one to define a particular partition for the test CT image. The computation of the organ boundaries is confined to this RoI, assuming that only two intensity classes exist, foreground (organs) and background. However, other classes may appear in the neighbourhood of the considered organ, mainly because of some acquisition artefacts, namely intensity inhomogeneities, noise and the presence or absence of fiducial markers. Then, we assume a two class  $RoI_o(\mathbf{x})$ : the organ of interest and some neighbouring tissues, which are randomly distributed. Every RoI is then modelled as a mixture of Gaussians (GMMs), aiming to approximate the organ of interest and the other contaminant tissues as a non uniform sum of Gaussians, each approaching a particular distribution of voxels. Such modelling is formulated as:

$$\psi(i) = \sum_{k=1,2,\dots,n} w_k N(i|\mu_k, \sigma_k^2) \quad (3)$$

where the two main distributions stand for the foreground tissues and any other structure (e.g. bones), respectively. Such an estimation is consistent because the different types of noises were herein modelled as additive, usually approached by a mix of Gaussians. A classical expectation-maximization algorithm is then applied to estimate the GMMs’ parameters. Let  $RoI_o(\mathbf{x})$  be the pixel tissue distributions, with the probability density function  $p(\mathbf{x}|\theta)$ , where  $\theta$  is an unknown vector of parameters  $(\mu_i, \sigma_i)$  and  $\mathbf{x}$  is every pixel of the RoI. Given an observation pixel  $\mathbf{x}$ , we aim at maximizing the likelihood function  $L(\theta) = p(\mathbf{x}|\theta)$  w.r.t. a given search space  $\Theta$ . This problem has been classically approximated by numerical routines, such as the well known expectation-maximization algorithm (Bilmes *et al* 1998). Once the mixture of



**Figure 3.** In (a) bladder RoI pre-processing before. The lower bladder region had a different appearance because of the filling bladder. (b) shows the processed RoI. The proposed filter reduces the appearance of difference making it possible to capture robust observations

Gaussians is determined, the two main distributions ( $\max_{k=1,2}(2\sigma_k)$ ) are chosen since they represent the two searched classes. Any other distribution is eliminated by setting its voxels to the closest main distribution, using an adapted non-local mean filter, as follows: a voxel  $\mathbf{x}$ , within the RoI, may represent any noise  $\{\mathbf{x} < \min_k(\mu_k - 2\sigma_k) \vee \mathbf{x} > \max_k(\mu_k + 2\sigma_k)\}$  and therefore is replaced by a weighted average of a neighbourhood with foreground/background information, satisfying a ‘non-local property’: weights depend on the pixel similarity in the image space as

$$\varrho(\mathbf{x}) = e^{\frac{-d(\mathbf{x})}{h^2}}, \quad (4)$$

where  $d(\mathbf{x}) = \sum_{i \in \phi(o)} \|RoI_o(\mathbf{x}) - N_{k1,k2}(\mu, \sigma^2)\|$ ,  $\phi(o)$  is the neighbourhood and  $h$  is a decay parameter. The underlying idea behind this filter is to replace a pixel belonging to a probable artefact by its nearest ‘foreground/background’ model. Examples of the RoI pre-processing are depicted in figure 3.

### 2.3. Multi-scale CT edge detector: the observations

The multi-scale image analysis was herein implemented using the first partial derivatives at each scale obtained by convolving each  $RoI_o(x, y)$  with a variable Gaussian kernel. Thus, the multi-scale edge estimation was obtained by a simple linear combination of the different gradient magnitudes at each scale, as follows:

$$S_o^{\text{edge}}(x, y; \sigma) = \sum_i RoI_o(x, y) * \frac{\partial G_{\sigma_i}}{\partial_x \partial_y} \quad (5)$$

where  $G_{\sigma_i}$  is the 2D Gaussian function with standard deviation  $\sigma_i$ . In this context, the Gaussian function is the unique kernel with an equivalent scale-space representation (linearity and shift-invariance in both frequency and space). A multi-scale image decomposition allows us to build a more stable set of observed edges (Lindeberg *et al* 1998) since true edges are coherently present at different scales. In this context, the detection of the boundaries is further pursued by applying a classical multi-scale non-maximum suppression strategy, consisting of detecting the maximum gradient magnitude for each particular direction as cited in Lindeberg *et al*



(1998). This multi-scale edge detection follows the universal principle of scale invariance and allows a robust edge description which is usually preserved through multiple scales (Romeny *et al* 1991).

#### 2.4. Computing the geometrical likelihood

Within a Bayesian framework, the devised likelihood function  $P(\hat{S}_{o_j}^{\text{pca}} | S_o^{\text{edge}})$  was defined to find the best geometrical matching between the sample shapes  $\hat{S}_{o_j}^{\text{pca}}$ , obtained from the learned model, and each multi-scale edge descriptor  $S_o^{\text{edge}}$  under a log-Euclidean metric (Vincent *et al* 2005). For doing so, every shape, from the organ space and the edges of each RoI, are characterized by the first seven Hu moments (Zhang 2004) as

$$P(S_{o_i}^{\text{PCA}} | S_{o_i}^{\text{edge}}) = \min_{S_{o_j}^{\text{pca}}} \left\{ \sum_{h_u (i=1 \dots 7)} \left\| m_i^{S_{o_i}^{\text{edge}}} - m_i^{S_{o_j}^{\text{pca}}} \right\| \right\} \quad (6)$$

where  $m_i^{S_{o_i}^{\text{edge}}} = \log |h_i^{S_{o_i}^{\text{edge}}}|$ , and  $m_i^{S_{o_j}^{\text{pca}}} = \log |h_i^{S_{o_j}^{\text{pca}}}|$  are the computed shape features for the edges and the PCA learned shapes, respectively, and  $h_i^{S_{o_i}^{\text{edge}}}$ ,  $h_i^{S_{o_j}^{\text{pca}}}$  are their Hu moments. Under a log-Euclidean metric, null values are excluded, and the metrics allow us to globally determine the most similar shape. This choice is also justified since the space of organs has a Lie group structure, that is to say, a similarity space in which continuity applies and algebraic operations are smooth mappings. This type of log-Euclidean metrics is invariant to orthogonal transformations, change of coordinates and scaling, and sets to an infinity distance those matrices with null or negative eigenvalues. The likelihood function reaches a maximal probability when a learned shape closely match the observations after the multi-edge-scale descriptor.

#### 2.5. Local shape deformation

In a final step, a 3D local deformation function was also introduced to eventually improve the local correspondence of the chosen organ, selected by the PCA model. Such deformation depends on two different measurements: (1) the radial distance of a particular shape point to the observed multi-edge-scale in the same slice and (2) the distance of the same particular organ point to the corresponding organ point in the upper and lower slices. These distances drive out the warping deformation of the organ either to the observed multi-edge-scale or to the upper and lower organ neighbouring points. This deformation is controlled by a  $\lambda$  term, which represents a trade-off between the observed edge and the prior organ shape

$$So_i(x, y) = \lambda \hat{S}_{o_i}^{\text{pca}}(x, y) + (1 - \lambda) \left( \min_{S_o} (\| \hat{S}_{o_i}^{\text{pca}} - \{ S_{o_i}^{\text{edge}}, \hat{S}_{o_{i\pm 1}}^{\text{pca}} \} \|) \right).$$

This local deformation tends to preserve a coherence between the prior shape and the observed edges, a compact representation of the shape given by the  $\lambda$  term and the nearest edge criterion. In this work, the best performance was obtained from a training database with  $\lambda = 0.6$ .

#### 2.6. Experimental setup

A quantitative comparison was performed between the manual organ delineations (prostate, bladder, rectum) and the computed segmentation, using two different measures: a dice score (DSC) and the Hausdorff distance. The DSC is an overlapping similarity measure, defined as  $\text{DSC}(A, B) = \frac{2|A \cap B|}{|A| + |B|}$ , where  $|\cdot|$  indicates the number of voxels of the considered



$A$  (gold standard) and  $B$  volumes. On the other hand, the Hausdorff measure  $H(A, B)$  computes the maximum distance between two set of points as  $\max(h(A, B), h(B, A))$  and  $h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|_2$ . In this case, each set of points represents the organ surface. This measure allows us to indirectly assess the segmentation compactness.

As mentioned before, the dataset was randomly split into: training (30 patients) and test (86 patients). The training set was used to build the organ shape space and the best representation was selected within a leave-one-out cross validation scheme. The performance of the proposed method during the training step was compared with three multi-atlas vote methods<sup>5</sup>, using a leave-one-out scheme. In a first multi-atlas method, the atlases were rigidly registered to the test volume. The other two methods followed two steps: (1) the atlases were rigidly registered to the whole set of volumes, using a ‘block matching’ strategy, and the transformed volumes were ranked according to the normalized cross-correlation (Ourselin *et al* 2001). (2) The delineations associated to each volume were propagated to a new test organ volume by non-rigidly propagating the whole set of organs. This non-rigid registration was carried out by either a standard free-form deformation (FFD) (Rueckert *et al* 1999) or by using the demons algorithm (Thirion 1998). Eventually, the majority-vote decision rule was applied to obtain a single segmentation for each considered organ. A second evaluation of the presented method was carried out by segmenting the test data (86 patients) that were not included in the training phase, using the best prior model constructed during the training step.

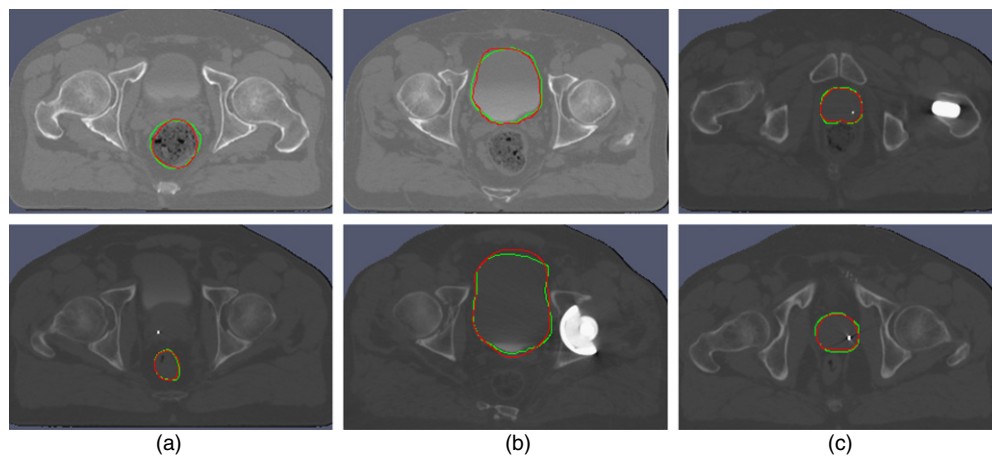
### 3. Data

Data used in this paper include a total of 116 patients, who underwent intensity modulated radiotherapy for localized prostate cancer between July 2006 and June 2007 in the same institution. The whole treatment (patient positioning, CT acquisition and volume delineations) and dose constraints complied with GETUG 06 recommendations, as previously reported (Beckendorf *et al* 2011). The size of the planning CT images in the axial plane was  $512 \times 512$  pixels, with 1 mm image resolution and 2 mm slice thickness. The treatment planning system used was Pinnacle V7.4 (Philips Medical System, Madison, WI). Each treatment plan used five field beams, in a step-and-shoot delivery configuration with gantry angles of 260°, 324°, 36°, 100° and 180°. The delivery was guided by means of an IGRT protocol (Perez *et al* 2005), with cone beam CT images or two orthogonal images (kV or MV imaging devices), using gold fiducial markers in 57% of patients. In this study, only the CT and the delineated prostate, bladder and rectum, were considered (Bagshaw *et al* 1990). For evaluation purposes, the data set was split into two parts : training (30 patients) and test (86 patients) datasets.

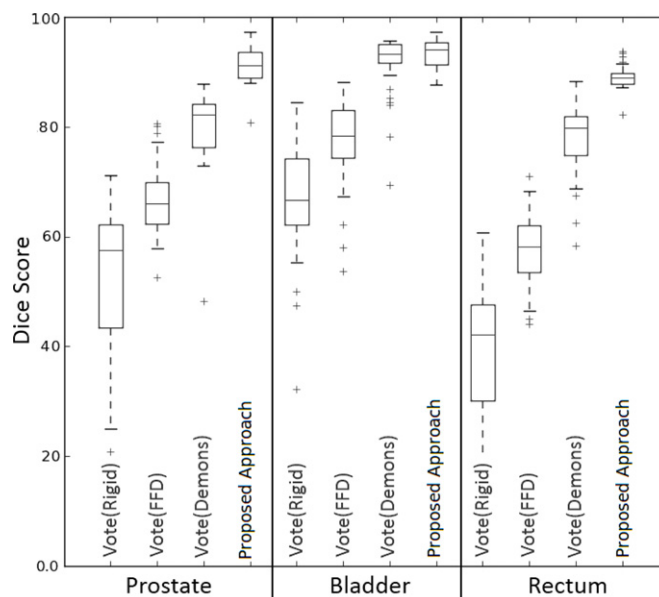
### 4. Evaluation and results

Figure 4 illustrates an example of segmentation obtained with the proposed approach (red), overlaid on an expert segmentation (in green). A major advantage of the proposed method is the adaptability of the prior shapes to the patient-specific organ, with local variations. As mentioned above, the approach performance was evaluated as follows.

<sup>5</sup> The database used corresponds to the 30 patients mentioned before.



**Figure 4.** Axial segmentation illustration of the pelvic structures ((a) rectum, (b) bladder and (c) prostate). The delineation obtained by the presented method (red) and the expert reference (green).



**Figure 5.** Dice scores comparison for vote versus the proposed approach.

#### 4.1. Evaluation of training data

In a first experiment, by using the 30 patients randomly selected from the 116 available cases, a DSC leave-one-out cross validation exploratory analysis was performed. Since DSC estimates the percentage of overlapping area between the evaluated segmentations, the obtained score is strongly dependent on the volume of the organs to be evaluated. Figure 5 shows the results obtained when comparing the presented approach with the three atlas-based methods. The graph depicts the mean DSC and standard deviation for the four different automatic methods.

**Table 1.** The Hausdorff distances obtained with the multi-atlas majority-vote method using rigid, FFD or a demons registration and with the proposed approach

Segmentation methods	Hausdorff Distance (mm)		
	Prostate	Bladder	Rectum
Vote(Rigid)	16.61 $\pm$ 5.6	102.02 $\pm$ 26	66.87 $\pm$ 10.3
Vote(FFD)	14.27 $\pm$ 4.2	78.63 $\pm$ 20.1	65.22 $\pm$ 6.1
Vote(Demons)	9.33 $\pm$ 3.2	79.42 $\pm$ 18.2	61.44 $\pm$ 5.8
proposed approach	5.98 $\pm$ 2.2	19.09 $\pm$ 3.1	7.52 $\pm$ 2.3

**Table 2.** The Dice score and the Hausdorff distances obtained the proposed approach in using the test data

Macro name	Prostate	Bladder	Rectum
Hausdorff distance (mm)	9.98 $\pm$ 3.4	25.07 $\pm$ 4.6	13.52 $\pm$ 5.1
Dice score (DSC)	0.87 $\pm$ 0.071	0.89 $\pm$ 0.083	0.82 $\pm$ 0.062

The average score observed for the proposed method was  $0.91 \pm 0.033$  for the prostate,  $0.94 \pm 0.028$  for the bladder and  $0.89 \pm 0.022$  for the rectum. Overall, in terms of the DSC, the presented approach outperforms the atlas-based method (using the demons) for the prostate and rectum, by 9%, and 9.2% ( $p < 0.001$ ), respectively, while for the bladder a slight gain of 1.2% was observed. Likewise, it should be noted that in all cases the standard deviation of the proposed approach is much smaller than the atlas-based approaches.

An additional comparison was performed by computing the Hausdorff distance, which allows for the compactness of the segmentation to be compared since this measure penalizes the isolated voxels that are far from the ground truth. Table 1 summarizes the obtained performance for the three target organs, with smaller values for the proposed method in any case, indicating that the method coherently finds shapes compatible with the knowledge stored in the bank of shape organs. These results, in addition, illustrate that the presented approach systematically obtains shape segmentations with average distances of 5.98 for the prostate, 19.09 for the bladder and 7.52 for the rectum, while with the best atlas approach, the average distances were 9.33 for the prostate, 79.42 for the bladder and 61.44 for the rectum. It should be emphasized that the Hausdorff distance evaluates not only the superposition of two shapes but also the quantity of scattered voxels that are outside of the segmented prostate: the larger the measure the higher the number of outlier voxels.

The results obtained with this metric, point out the importance of using structural priors that define a particular topology. The bladder segmentation, calculated with the three atlas-based approaches and compared with the gold standard under the DSC metric, seemed to be appropriate, but when using the Hausdorff distance, larger differences appeared and produced a very large error of 79.42 mm in this study.

#### 4.2. Evaluation of test data

Once a particular organ model was set, its generalization capacity was also evaluated. The best prior, determined by the PCA in the previous experiment with 30 patients, was used in a second test group with 86 patients, a larger data set with different shape organs, presence of artefacts or contrast changes in CT. Quantitative results, with the previously introduced metrics, are reported in table 2. In general terms, the obtained results show a segmentation with an averaged

DSC of 0.86 and an averaged Hausdorff distance of 16.19. The obtained segmentations are consistent with what was observed in the training data, i.e. with the best score being achieved for the bladder. This fact can also be attributed to the overlapping measure since results depend on the organ volume. Although the obtained score errors are slightly larger for the three organ segmentations, the proposed approach properly captures the shape variability. The Hausdorff distance, on the other hand, shows that the proposed approach produces more stable segmentations, meaning that voxels belong mostly to a single shape since the measurement penalizes isolated or outlier voxels. Unlike the atlas-based approaches, this method reaches a tolerable margin of error, an important element when planning radiation therapy. Finally, in both quantitative metrics, the obtained standard deviation was lower, evidencing coherency in the obtained segmentations.

## 5. Discussion

An automatic method for the segmentation of pelvic organs in CT images was presented. Under a Bayesian framework, a geometrical likelihood function mapped a set of global observations, built from a structural multi-scale analysis (Lindeberg *et al* 1998), to a prior shape space that stored the shape organ knowledge. This prior captured the principal shape variation modes, constituting an organ space with samples that represent 96% of the shape variability (Heimann and Meinzer 2009, Cootes and Taylor 2001). Observations were automatically captured from RoIs around each specific organ, namely, the prostate, the bladder and the rectum. Each of these RoIs is pre-processed by an adapted non-local filter which regularizes the region by considering only two principal pixel distributions: the organ and the background. This filter allows for artefact reduction coming from gas in the rectum, several filling bladder states or bone parts randomly captured within the RoI. Afterward, a set of structural observations, the Hu moments, are calculated from a multi-edge-scale feature. The obtained Hu-based descriptor is a global shape representation that is used by a geometrical likelihood function to find the most likely organ shape given the set of observations. The likelihood function rules out the null Hu moments and redistributes the organ space using a log-Euclidean metric and thereby it matches the most similar.

In general, state-of-the-art methods attempt to obtain more accurate segmentations by integrating a prior to the CT image information, for instance the CT or MRI manual expert delineations or CT salient morphological points. The Bayesian approach presented here may be included in the statistical shape models family since the organ shape prior is mapped to the CT through a geometrical likelihood. Several statistical models for segmenting CT registered structures have been proposed. Among them, a semi-automatic Bayesian method that matches a set of organ templates and finds the most probable template transformation using an iterative increasing and bending algorithm may be found in Joshi *et al* (2012). Each of these templates is characterized by a medial axis relationship at different resolution scales. However, the media axis result, in many cases, is a very simple shape representation that can lead to wrong segmentations since the growing iterative algorithm can easily overflow the pelvic organ boundary. Likewise, segmentation of the prostate and rectum has also been approximated using a histogram learning representation under a Bayesian framework (Chen *et al* 2011a), achieving an average volume overlapping of 0.89 and 0.71, for the prostate and rectum, respectively. These low figures can be attributed to the histogram representation since (1) it is not robust to contamination with any complex noise, (2) some bladder states and bone structures show practically the same grey level, (3) there exists a high inter patient variability regarding the bladder state and (4) it completely misses the spatial relationships.

Essentially, the proposed approach combines both a multi-scale (global) and derivative estimation (local) to obtain the real organ edges. Since organ boundaries in the CT images are usually very blurred and even experts can miss actual boundaries, the proposed approach sets a proper edge estimation using a likelihood function that uses global metrics (the Hu moment characterization) of the searched shape and then corrects for the possible edge mismatches. Another advantage of the proposed approach is the use of prior shapes by the likelihood function so that voxels belonging to any organ shape are always connected. This investigation has extended the usual prostate segmentation to the whole set of OARs, with very little change in the required parameters. The method is demonstrated to be robust to many types of noise and adaptable to different organs, with comparable accuracy values when segmenting the prostate, the bladder or the rectum. The approach was compared to three different atlas-based methods which also included geometrical prior information, using the DSC and Hausdorff distance calculated between the automatic organ segmentation and a manual delineation carried out by a radiologist expert. Patients were grouped in training and test groups, and in both scenarios the proposed approach outperformed the other methods.

Finally, several pre-processing strategies have been proposed to overcome the natural high intra- and inter-patient organ appearance variability, for instance, Davis *et al* (2005) deal with the bowel gas problem by defining a binary mask containing the regions with gas. Then, they apply a 'deflation' transformation, based on the flow computed or induced by the binary image mask, and in doing so this region tends to converge to the centre of these regions. These approximations allow a reduction in the bowel gas effect but without taking into account that very often this phenomenon appears as scattered in small regions. In the proposed approach, a new non-local filter which replaces the RoI artefacts was introduced, taking advantage of the redundancy and assuming that a segmented RoI is composed of two tissues: organ and background. This facilitates the organ segmentation, above all because some organs, like the bladder or rectum, are composed of different objects that result in different distributions, this filter allows us to homogenize the organ texture and to isolate the concept. After this filter, an organ and background pixel distribution may be facilitated and used thereafter. Last but not least, the obtained results are highly encouraging and demonstrate the method's usefulness in clinical scenarios as a support of the final delineation that dramatically decreases the expert burden in the daily RT planning.

## 6. Conclusions and perspectives

In this work we presented a new methodology to segment pelvic structures in CT scans. The Bayesian method combines a deformable prostate model, learned from examples, and a geometrical likelihood strategy that maps actual observation into the space of model organs and selects the best shape to superimpose it in the original CT image. The results suggest that this segmentation technique may be suitable to use as an oncologist's support. This approach may be easily extended to other structures in CT images. Future works include a more robust CT characterization by the association of edge observation and pixel distributions in order to find more likely shapes in the organ space and also develop a more reliable local shape adjustment.

## Acknowledgments

This work is partially supported by Rennes Metropole through a mobility scholarship and the project ANR TIGRE.

## References

- Acosta O *et al* 2010 Atlas based segmentation and mapping of organs at risk from planning CT for the development of voxel-wise predictive models of toxicity in prostate radiotherapy *Prostate Cancer Imaging: Int. Workshop in MICCAI 2010 (Lecture Notes in Computer Science vol 6367)* pp 42–51
- Acosta O *et al* 2011 Evaluation of multi-atlas-based segmentation of CT scans in prostate cancer radiotherapy *ISBI'11: IEEE Int. Symp. Biomed. Imaging* pp 1966–9
- Acosta O, Dowling J, Drean G, de Crevoisier R, Simon A and Haigron P 2014 *Abdomen and Thoracic Imaging, an Engineering and Clinical Perspective* ed A S El-Baz, L Saba and J S Suri (Berlin: Springer) pp 623–56
- Acosta O, Drean G, Ospina J D, Simon A, Haigron P, Lafond C and de Crevoisier R 2013 Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy *Phys. Med. Biol.* **58** 2581
- Al-Qaisieh B *et al* 2002 Impact of prostate volume evaluation by different observers on CT-based post-implant dosimetry *Radiother Oncol.* **62** 267–73
- Bagshaw M A, Cox R S and Ramback J E 1990 Radiation therapy for localized prostate cancer. Justification by long-term follow-up *Urol. Clin. North Am.* **17** 787–802
- Bayley A, Rosewall T, Craig T, Bristow R, Chung P, Gospodarowicz M, Ménard C, Milosevic M, Warde P and Catton C 2010 Clinical application of high-dose, image-guided intensity-modulated radiotherapy in high-risk prostate cancer *Int. J. Radiat. Oncol. Biol. Phys.* **77** 477–83
- Beckendorf V *et al* 2002 Target-volume and critical-organ delineation for conformal radiotherapy of prostate cancer: experience of French dose-escalation trials *Cancer Radiother.* **6** 78–92
- Beckendorf V *et al* 2011 70 Gy versus 80 Gy in localized prostate cancer: 5-year results of GETUG 06 randomized trial *Int. J. Radiat. Oncol. Biol. Phys.* **80** 1056–63
- Bilmes J 1998 A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models *Technical Report* Berkeley, CA, TR-97-021
- Bystrov D *et al* 2012 Simultaneous fully automatic segmentation of male pelvic risk structures *Estro Eur. Soc. Radiother. Oncol.* **103** S333–4
- Chen S, Lovelock D M and Radke R J 2011a Segmenting the prostate and rectum in CT imagery using anatomical constraints *Med. Image Anal.* **15** 1–11
- Chen S, Lovelock D M and Radke R J 2011b Segmenting the prostate and rectum in CT imagery using anatomical constraints *Med. Image Anal.* **15** 1–11
- Collier D *et al* 2003 Assessment of consistency in contouring of normal-tissue anatomic structures *J. Appl. Clin. Med. Phys.* **4** 17–24
- Cootes T and Taylor C 2001 Statistical models of appearance for medical image analysis and computer vision *Proc. SPIE Med. Imaging* **4322** 236–48
- Costa M J, Delingette H, Novellas S and Ayache N 2007 Automatic segmentation of bladder and prostate using coupled 3D deformable models *Med. Image Comput. Comput. Assist. Interv.* **10** 252–60
- D'Amico *et al* 1998 Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer *JAMA* **280** 969–74
- Davis B C, Foskey M, Rosenman J, Goyal L, Chang S and Joshi S 2005 Automatic segmentation of intra-treatment CT images for adaptive radiation therapy of the prostate *Med. Image Comput. Comput. Assist. Interv.* **8** 442–50
- Evangelia M-T, Torsten R, Robert B, Randolph M, Daniel R and Calvin M 2005 *Atlas-based Segmentation Handbook of Biomedical Image Analysis* (Berlin: Springer) pp 435–86
- Feng Q, Foskey M, Tang S, Chen W and Shen D 2009 Segmenting CT prostate images using population and patient-specific statistics for radiotherapy deformable models *ISBI'09: Proc. IEEE Int. Symp. on Biomedical Imaging* pp 282–5
- Fiorino C, Reni M, Bolognesi A, Cattaneo G M and Calandrino R 1998 Intra- and inter-observer variability in contouring prostate and seminal vesicles: implications for conformal treatment planning *Radiother. Oncol.* **47** 285–92
- Gual-Arnau X, Gual I, Lliso F and Roldán S 2005 Organ contouring for prostate cancer: interobserver and internal organ motion variability *Comput. Med. Imaging Graph.* **29** 639–47
- Guckenberger M and Flentje M 2007 Intensity-modulated radiotherapy (IMRT) of localized prostate cancer: a review and future perspectives *Strahlenther Onkol.* **183** 57–62
- Heemsbergen W D *et al* 2010 Urinary obstruction in prostate cancer patients from the Dutch trial (68 Gy versus 78 Gy): relationships with local dose, acute effects, and baseline characteristics *Int. J. Radiat. Oncol. Biol. Phys.* **78** 19–25



- Heimann T and Meinzer H P 2009 Statistical shape models for 3D medical image segmentation: a review *Med. Image Anal.* **13** 543–63
- Hodge A C, Fenster A, Downey D B and Ladak H M 2006 Prostate boundary segmentation from ultrasound images using 2D active shape models: optimisation and extension to 3D *Comput. Methods Programs Biomed.* **84** 99–113
- Joshi S, Pizer S, Fletcher P T, Yushkevich P, Thall A and Marron J S 2012 Multiscale deformable model segmentation and statistical shape analysis using medial descriptions *IEEE Trans. Med. Imaging* **21** 538–50
- Langen K M and Jones D T L 2001 Organ motion and its management *Int. J. Radiat. Oncol. Biol. Phys.* **50** 265–78
- Li W, Liao S, Feng Q, Chen W and Shen D 2012 Learning image context for segmentation of the prostate in CT-guided radiotherapy *Phys. Med. Biol.* **57** 1283
- Lindeberg T 1998 Feature detection with automatic scale selection *Int. J. Comput. Vision* **30** 79–116
- Lorenz C *et al* 2000 Generation of point-based 3D statistical shape models for anatomical objects *Comput. Vis. Image Underst.* **77** 175–91
- Mikael R, Ali K, Mamadou D, Juan C and Frank S 2005 Constrained surface evolutions for prostate and bladder segmentation in CT images *Computer Vision for Biomedical Image Applications* (Berlin: Springer)
- Ourselin S *et al* 2001 Reconstructing a 3D structure from serial histological sections *Image Vis. Comput.* **19** 25–31
- Pekar V, McNutt T R and Kaus M R 2004 Automated model-based organ delineation for radiotherapy planning in prostatic regions *Int. J. Radiat. Oncol. Biol. Phys.* **3** 973–80
- Perez C A, Michalski J and Mansur D 2005 Clinical assessment of outcome of prostate cancer (TCP, NTCP) *Rays* **30** 109–20
- Rohlfing T and Maurer C R 2005 Multi-classifier framework for atlas-based image segmentation *Pattern Recognit. Lett.* **26** 2070–9
- Rohlfing *et al* 2004 Multi-classifier framework for atlas-based image segmentation *CVPR'04: Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition* **1** 255–60
- Romeny T H *et al* 1991 Scale space: its natural operators and differential invariants *Information Processing in Medical Imaging* (Berlin: Springer) pp 239–55
- Rueckert D *et al* 1999 Nonrigid registration using free-form deformations: application to breast MR images *IEEE Trans. Med. Imaging* **18** 712–21
- Liao S, Gao Y and Shen D 2012 Sparse patch based prostate segmentation in CT images *Medical Image Computing and Computer-Assisted Intervention: MICCAI 2012* (Berlin: Springer) pp 385–92
- Sohn M *et al* 2005 Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation *Phys. Med. Biol.* **50** 5893
- Söhn M *et al* 2012 Dosimetric treatment course simulation based on a statistical model of deformable organ motion *Phys. Med. Biol.* **57** 3693
- Thirion J P 1998 Image matching as a diffusion process: an analogy with Maxwell's demons *Med. Image Anal.* **2** 243–60
- Arsigny V, Fillard P, Pennec X and Ayache N 2005 Fast and simple calculus on tensors in the log-Euclidean framework *Medical Image Comput. Comput. Assist. Interv.* **8** 115–22
- Wu M, Rosano C, Lopez-Garcia P, Carter C S and Aizenstein H J 2007 Optimum template selection for atlas-based segmentation *Neuroimage* **34** 1612–8
- Zhang D 2004 Review of shape representation and description techniques *Pattern Recognit.* **37** 1–19