DIFFERENTIAL SEGMENTATION OF THE PROSTATE IN MR IMAGES USING COMBINED 3D SHAPE MODELLING AND VOXEL CLASSIFICATION

P. D. Allen, J. Graham, D. C. Williamson and C. E. Hutchinson

ISBE, University of Manchester, Oxford Road, M13 9PT, UK

ABSTRACT
Benign Prostatic Hyperplasia (BPH) is a non-cancerous expansion of the prostate, the progress of which can be quantified by measuring the relative volumes of the prostate’s peripheral zone and central gland. Here we describe a method of automatic segmentation of both regions of the prostate from MR images using a combination of grey-level voxel classification and 3D statistical shape modelling.

1. INTRODUCTION
Benign Prostatic Hyperplasia (BPH) is a non-cancerous enlargement of the prostate which can cause constriction of the urethra and therefore obstruction of urinary flow. It affects 70% of men between the ages of 61 and 70, rising to 80% for men over 80 [1]. In 25% of men aged 80 symptoms are sufficiently severe to require surgical transurethral resection of the prostate (TURP), however this treatment has a high cost, morbidity (16%) and mortality (2.01%) [2] and so alternative treatments are sought.

Drugs such as finasteride can be used to treat BPH by shrinking the prostate, and evaluation of such candidate treatments requires a method of quantifying its effect. The current standard is Transrectal Ultrasound (TRUS) in which three orthogonal dimensions are measured and the volume is estimated using the formula for a prolate ellipsoid [2, 3]. Anatomically the prostate is divided into a number of zones: Peripheral (PZ), Central (CZ), Transitional (TZ), and fibromuscular. BPH primarily affects the TZ and so both the total Prostate (TP) and TZ volumes are measured using TRUS.

Tewari et al [2] have shown that the reduction in volume due to finasteride treatment over 12 months for the total TP and TZ are 8% and 27% respectively, however only the change in TZ volume correlates with improvement in urinary flow. For TRUS, intra-observer variability has been shown to be −18% to +18% for the TZ volume and −21% to +30% for the total volume [3].

Magnetic Resonance Imaging (MRI) is an attractive alternative to TRUS as it offers better definition of the prostate and is non invasive. MRI offers the possibility of accurately segmenting the prostate rather than assuming its volume from three orthogonal measurements. However, manual segmentation is time consuming, error prone and subjective, and so the goal of this project is to investigate the possibility of automatic segmentation of the appropriate regions of the prostate.

In MRI only two regions can be distinguished: the PZ, and what is referred to as the Central Gland (CG) comprising the remaining anatomical zones [4]. In cases of BPH the CG is mostly comprised of TZ due to the latter’s expansion and so CG and TZ can be considered equivalent. Methods of segmentation [5] and registration [6] of just the outer prostate surface have been described for MR imaging, here we describe a method of both whole prostate and CG segmentation.

For this study we have used T2 weighted fat suppressed (T2FS) images as the CG/PZ contrast is enhanced in comparison with T2 or T1 weighting, and there is clearer separation of the prostate from surrounding tissue. The data were collected using a 1.5T Philips Gyroscan ACS MR scanner (software version NT5.3, Power Track 600, synergy body coil) from 22 patients with BPH. For each patient there are 50 axial slices with a thickness of 2mm and an in-plane resolution of 1.56mm.

Figure 1 shows a T2FS MR image of a prostate sliced in the axial, sagittal, and coronal planes. In T2 weighted images the PZ is generally brighter than the CG and in this case the two can be distinguished reasonably well.

Manual segmentation of the prostate is particularly difficult toward the superior portion where the seminal vesicles are very difficult to distinguish from the PZ, and toward the inferior portion where surrounding structures can become confused with the prostate and the prostate itself tends to bifurcate into two lobes. In the mid-section of the prostate blood vessels anterior to it can be confused with the PZ or CG depending on their relative intensity. The border between the PZ and CG can vary greatly from patient to patient depending on the severity of glandular enlargement. Figure 6 shows manual segmentation of the axial slice of the prostate, illustrating that the boundary between the regions is defined not only by the voxel values, but requires a model of expected shape.

2. AUTOMATIC SEGMENTATION
In seeking a method of automatic segmentation, our approach is to formalise the two level process behind manual segmen-
Fig. 1. The appearance of the prostate in T2 weighted Fat suppressed MRI sliced in three orthogonal planes.

2.1. Voxel Classification

Fig. 2. Axial T2FS MR image of a prostate.

Fig. 3. The result of applying grey-level tissue classification to figure 2.

Fig. 4. The grey-level histogram of figure 2.

Fig. 5. The mean shape of the double surface 3D PDM of the prostate

Fig. 6. Manual segmentation of an axial slice into PZ and CG.

Fig. 7. The result of applying grey-level tissue classification to figure 6.

2.2. Shape Modelling

We are interested in fitting two surfaces: the Total Prostate (TP) and the Central Gland (CG) (see figure 5). The 22 images have been manually segmented to provide examples of each of these surfaces. To build a PDM from these surfaces requires a set of points on each surface which correspond across the data set and to achieve this we employ a method of automatic correspondence optimisation [10].

Using a leave-one-out evaluation of the ability of the PDM to represent the training data, it was found that using two separate PDMs for the TP and CG surfaces gave considerably better representation of the observed shape variability than using a single PDM of the two surfaces. This is because the 22 manually-segmented examples are not sufficient to ade-
quately describe the variation in the spatial relationships between the two surfaces.

2.3. Model Fitting

2.3.1. TP Surface:

From the tissue classification (section 2.1) each voxel has three values associated with it - \( P_{\text{PZ}}, P_{\text{CG}}, P_{\text{B}} \) representing the probabilities that it belongs to PZ, CG, or B. For an example surface we can sum these probabilities for the voxels enclosed by the surface giving the quantities which we can call \( P_{\text{Z}_{\text{in}}}, C_{\text{G}_{\text{in}}}, \) and \( B_{\text{in}} \). For fitting the whole prostate surface a sensible objective function would then be:

\[
P_{\text{Z}_{\text{in}}} + C_{\text{G}_{\text{in}}} - B_{\text{in}}
\]

We initialise the search for fitting all of the shape parameters by first fitting the pose of the average shape. Using the objective function in equation 1, and pose parameters only, the search space is fairly smooth. We are able to use simplex to find an initial configuration for shape search. Optimisation of the surface shape however presents a far more complex search space with many local minima, and so here a genetic algorithm is used.

2.3.2. CG Surface:

To fit the CG surface it is the PZ/CG border that must be emphasised in the objective function and this can be done in the following way: Create a candidate CG surface \( C_1 \), then dilate that surface by one voxel to form a second surface \( C_2 \), which, for a correct surface \( C_1 \), should be outside the CG. If we sum the probabilities on surfaces rather than in them we can form the values \( P_{\text{Z}_{\text{on}C_2}} \) and \( P_{\text{Z}_{\text{on}C_1}} \). The difference between these values should be a maximum when the CG surface is on the PZ/CG border. As the PZ does not always extend round to the anterior of the prostate (see figure 1), we also need to find the CG/B border in this region. We therefore also calculate \( B_{\text{on}C_2} \) and \( B_{\text{on}C_1} \). From a search point of view this objective function is spiky as candidate solutions near but not at the correct position are no better than surfaces further away, and so counting the voxel probabilities inside the surface is also necessary. Thus the CG objective function becomes:

\[
(P_{\text{Z}_{\text{on}C_2}} - P_{\text{Z}_{\text{on}C_1}}) + (B_{\text{on}C_2} - B_{\text{on}C_1}) + C_{\text{G}_{\text{in}}} - B_{\text{in}}
\]

In this case the sums of probabilities in the surface are normalised by surface volume to make them the same order of magnitude as the sums on the surface which are normalised by surface area. The term \( 1 - P_{Z_{\text{in}}} \) is left out of the objective function because in some cases the there is considerable misclassification of CG voxels as PZ.

Convergence of the CG fitting was not always successful starting from the mean shape. This difficulty was overcome by starting the search from a shape specified by a straightforward user interaction: the user marks four points on the middle slice roughly equally spaced around the CG, and selects the slices corresponding to the inferior and superior limits of the CG. This information is used to initialise the mean shape by stretching it in the X,Y, and Z axes and adjusting position in a simplex optimisation until the surface is as close to the user defined points as possible. The resulting surface is then used as a start point for a full GA optimisation of pose and shape.

2.4. Results

Prior to tissue classification and PDM fitting the image data was cropped manually around the prostate, as the full axial slices encompass the entire pelvic area in which the prostate is only a small region. The model fitting was tested in a series of leave-one-out experiments in which a surface model was built from the set of examples excluding the current example. GA optimisation was performed using MATLAB’s genetic algorithm toolbox.

The results of fitting the TP and CG surfaces to each of the 22 patients are shown in table 1 as a mean point distance and percentage volume error. The effects of the anatomical ambiguity in the superior and inferior portions of the prostate (see section 1) can be reduced by only considering the mid-third of the cropped volume during shape and pose optimisation and the results of this are also included in table 1. Naturally the TP and CG volumes for the mid-third of the prostate are meaningless in themselves, however the CG/TP ratio calculated for this region from manual segmentation has a strong correlation (\( r=0.97 \)) with the CG/TP ratio for the whole gland (figure 8). This suggests that to measure the CG/TP ratio, segmentation of the more clearly defined mid-section of the prostate may be sufficient.

2.4.1. Repeatability:

Since the GA includes a stochastic element the same fit given the same data is not guaranteed. The magnitude of this var-

<table>
<thead>
<tr>
<th>Surface</th>
<th>Point Diff (( \mu(\sigma) ))</th>
<th>Volume Diff (( \mu(\sigma) ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>4.1 (1.1)</td>
<td>11.1 (9.5)</td>
</tr>
<tr>
<td>TP (mid)</td>
<td>2.8 (0.82)</td>
<td>6.5 (5.4)</td>
</tr>
<tr>
<td>CG</td>
<td>3.1 (2.5)</td>
<td>11.9 (8.9)</td>
</tr>
<tr>
<td>CG (mid)</td>
<td>2.0 (0.6)</td>
<td>6.8 (8.5)</td>
</tr>
</tbody>
</table>

Table 1. Fit Results (see text).
Fig. 8. The CG/TP ratio for the mid-gland plotted against the CG/TP ratio for the whole prostate.

ability can be estimated by repeating the fitting process 10 times and observing the variation in measured volume. On a subset of 10 of the patient group the results of this suggest standard deviations of 3% and 2% for the whole and mid-gland fits respectively.

3. DISCUSSION AND CONCLUSIONS

Table 1 demonstrates that in the majority of cases automatic segmentation results in Total Prostate and Central Gland surfaces that correspond accurately to the manual segmentation ‘ground truth’. The key measure in this case is volume and even in cases where there is the greatest difference between automatic and manual segmentation these are comparable with the variation in volume estimates using TRUS. Automatic segmentation from MR images clearly has the potential to deliver precise estimates of volume change. Much of the discrepancy between manual and automatic segmentation arises in places where the boundary location is genuinely unclear, and the absolute nature of the ground truth is questionable. We intend to investigate the variability in manual measurement in these regions.

One of the more important measures to be derived is the ratio of volume of the Central Gland to Total Prostate. We have observed that this can be estimated reliably by only segmenting the more clearly defined central portion of the prostate. In a practical situation this may provide a solution to the more difficult cases.

4. REFERENCES


