Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

3D Prostate Boundary Segmentation From Ultrasound Images Using 2D Active Shape Models

Adam C. Hodge, and Hanif M. Ladak, *Member, IEEE*

***Abstract*—Boundary outlining, or segmentation, of the prostate is an important task in diagnosis and treatment planning for prostate cancer. This paper describes an algorithm for semi-automatic, three-dimensional (3D) segmentation of the prostate boundary from ultrasound images based on two-dimensional (2D) active shape models (ASM) and rotation-based slicing. Evaluation of the algorithm used distance- and volume-based error metrics to compare algorithm generated boundary outlines to gold standard (manually generated) boundary outlines. The mean absolute distance between the algorithm and gold standard boundaries was 1.09±0.49 mm, the average percent absolute volume difference was 3.28±3.16%, and a 5x speed increase as compared manual planimetry was achieved.**

I. Introduction

PROSTATE cancer is a commonly occurring disease in men with an estimated 20,700 new cases and 4,200 deaths in Canada [1], and 234,460 new cases and 27,350 deaths in the U.S.A. [2]. If diagnosed early prostate cancer is curable, and even at later stages treatment may still be effective. However, prostate cancer has a considerable impact on the quality of life of adult and elderly men.

Outlining, or segmentation, of the prostate boundary from ultrasound (US) images plays a crucial role in both diagnosis and therapy planning procedures of prostate cancer treatment. US is commonly employed in urology clinics as it is a fast, portable, and cost-effective medical imaging modality. However, US images tend to have low contrast, and suffer from imaging artifacts such as speckle noise and shadowing of tissue boundaries. A significant challenge facing urological clinicians is the identification of the prostate boundary in these images for the purpose of determining the prostate's shape and volume. Segmentation done using manual planimetry, in which a clinician draws a contour on the prostate boundary in a series of US images, is time-consuming, tedious, and subject to inter- and intra-observer variability, as well as observer bias.

Computer algorithms aim to aid clinicians by semi- or fully-automatically drawing a contour on the prostate boundary in an US image. These algorithms are used to decrease clinicians' workload and reduce segmentation time. A statistics-based deformable model called the Active Shape Model (ASM) has been shown to be a promising method of segmenting medical images [3]. The ASM was previously applied to two-dimensional (2D) prostate segmentation in trans-urethral ultrasound (TUUS) images by Arambula-Cosio and Davies [4]. However, the previous algorithm was applied only to 2D TUUS images. The objective of this paper is to apply the 2D ASM algorithm to prostate boundary segmentation of three-dimensional (3D) trans-rectal ultrasound (TRUS) images.

Manuscript received April 3rd, 2006. This work was supported in part by the Institute for Robotics and Intelligent Systems, and the Natural Sciences and Engineering Research Council of Canada.

A. C. Hodge is with the Department of Medical Biophysics, The University of Western Ontario, London, Ontario, N6A 5C1 Canada (e-mail: ahodge3@uwo.ca).

H. M. Ladak is with the Department of Medical Biophysics and the Department of Electrical and Computer Engineering, The University of Western Ontario, London, Ontario, N6A 5C1 Canada (phone: +1-519-661-2111 x86551; fax: +1-519-661-2123; e-mail: hladak@uwo.ca).

II. Methods

A. Active Shape Models

A general description of the ASM algorithm is given here, and complete details may be found in [3;5]. ASMs are based on two separate components: a shape descriptor function called a Point Distribution Model (PDM), and an image search procedure which uses the PDM to locate an object in an image.

A PDM uses a-priori statistical information to describe shapes of a modelled class of object. Statistics of average shape and the modes of variation (ways in which the shape tends to vary) are determined by applying Principal Component Analysis (PCA) to a training set of labelled example objects. Shape labelling is used to represent the original, continuous contours of the training set (drawn manually by an expert observer) suitably well using a fixed number of landmark points that are placed on strongly identifiable features. Since the prostate does not have well-defined structures that can be reliably and consistently landmarked, the Minimum Description Length (MDL) method of landmark placement for statistical shape models developed by Davies [6] is used for shape labelling. Each labelled shape is represented by a $2n$-dimensional vector $\mathbf{x} = [x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]^T$, where n is the number of landmarks and each point (x_i, y_i) corresponds to the same location on all training set contours. In order to minimise variation due to linear transformations (i.e., scale, rotation, and translation), the labelled training set is first aligned to a common local coordinate frame ('model space'). Alignment of the training set of shapes is an automated, iterative process that aligns each shape in the training set to a current estimate of the mean shape. Once all training set shapes are

aligned, a new estimate of the mean shape is computed, centered at the origin, and scaled to unity. The process is repeated until the mean shape does not change substantially. PCA is then applied to the labelled, aligned data to find the modes of variation, which together with the mean shape define the PDM. New shapes of the modelled class are generated using the PDM by adding a weighted combination of the modes of variation to the average shape, given by:

$$\mathbf{x} \approx \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} \tag{1}$$

where $\mathbf{P} = [\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_m]$ is the modes of variation matrix, and $\mathbf{b} = [b_1,b_2,\ldots,b_m]^T$ is the modes of variation weighting vector, also referred to as the shape parameters. Only a number of modes of variation, $m \le n$, that explain a suitable amount of variation are used. The shape parameter values are constrained such that $|b_k| \le 2\sqrt{\lambda_k}$, where λ_k is the eigenvalue for b_k, so only shapes similar to those in the training set are produced.

The image search procedure is an iterative process which deforms the PDM to best fit the object in the image. The first step of the algorithm is to have the user interactively scale, rotate, and translate the average shape from model space to 'image space' (i.e. as the shape appears in the coordinate frame of the image). These transformation parameters (s,θ,t_x,t_y) are referred to as the pose parameters. Normal profiles are computed and searched $L_{search} = 10$ pixels to either side of the model point. A number of pixels, $L_{sub} = 9$, to either side of each pixel p in the normal profile are used for computing greyscale transition. Candidate target points (where the model points should move to) are selected based on the minimal value of the following fit cost function, $C(p)$, given by:

$$\begin{aligned} C(p) = 255 &+ \text{mean}\{p - L_{sub},.., p-1\} \\ &- \text{mean}\{p+1,.., p+L_{sub}\} \end{aligned} \tag{2}$$

An estimate of the target points is then generated using the PDM by applying the inverse transformation to project the updated image space instance back to model space, and computing a new shape parameters vector:

$$\mathbf{b} = \mathbf{P}^T(\mathbf{x} - \bar{\mathbf{x}}) \tag{3}$$

Note that $\mathbf{P}^{-1} = \mathbf{P}^T$ because $\mathbf{P}$ is orthonormal. These steps are repeated until the convergence criterion is met; that is, the shape does not change substantially or a maximum number of iteration is reached.

B. 3D TRUS Images

Two-dimensional images were extracted from volumetric prostate images acquired using a 3D TRUS imaging system [7]. A total of 36 volumetric images were collected, each from a different patient admitted for needle biopsy assessment. The volumetric images were approximately 350x400x800 pixels, with nearly isotropic voxels of ~0.15-0.20 mm per edge. A series of 36 sequential 2D image slices (corresponding to an angular sampling of 5°) were extracted from each 3D image using the rotational-based slicing technique presented in [8].

C. Training Data Set

As manual outlines are prone to inter- and intra-observer variability, and as suggested by Chalana and Kim [9], a manual outline produced by a single observer is not a suitable gold standard. Instead, [9] suggests that to form a suitable gold standard, the average manual outline built from multiple observer manual outlines may be used. In this work, the prostate boundary in each 2D image slice was outlined three times, once each by 3 trained graduate student observers. The 3 manual outlines were then averaged using the method presented in [9] to define a gold standard outline for that 2D image slice.

D. 3D Segmentation Using 2D ASMs

The 2D ASM is applied to 3D segmentation via a slice-to-slice propagation methodology as described in [8]. A slice-based technique avoids the computational burden of 3D-space calculations and difficulties encountered when manually editing a surface. Note that a 7x7 median filter was used to smooth each 2D image to suppress speckle noise. A leave-one-out approach was used so as to not include any shapes from the 3D image being segmented into the training set.

As a 2D ASM is trained such that it may only take on shapes similar to those of the training set of shapes, the inclusion of training shapes from the entire range of sampling angles will have a detrimental effect on PDM specificity, i.e. the ability of the model to generate new, plausible prostate shapes. In order to avoid this problem a multiple region approach was taken. The number of regions was selected to minimise the need to change models without greatly reducing PDM specificity. Three regions were used with the ranges: 0°-55°, 60°-115°, and 120°-175°, respectively. PDMs were built from shapes extracted every 20° within each respective region, e.g. if patient 36 is left out for segmentation, a PDM for the region 0°-55° is built using 105 training shapes: 3 shapes (corresponding to 0°, 20°, and 40°) from each of patient 1...35; PDMs are similarly built for the regions 60°-115°, and 120°-175°. During segmentation, the PDM corresponding to the particular region is applied to each image slice. In the special case that the proceeding image is in a new model region, a model matching approach is used. Given a current model i and the model j to be switched to, the process is given in pseudo-code as:

1. transform $\mathbf{X}_i$ (current image space shape) into model space j to get model space shape $\mathbf{x}_j$
2. solve $\mathbf{b}_j = \mathbf{P}_j^T(\mathbf{x}_j - \bar{\mathbf{x}}_j)$

3. compute model space shape estimate $\mathbf{x}'_j = \bar{\mathbf{x}}_j + \mathbf{P}_j \mathbf{b}_j$
4. transform $\mathbf{x}'_j$ to image space to get $\mathbf{X}_j$
5. continue image search process using model *j*.

E. Manual Correction

As in [8], manual correction of the algorithm segmentation result was allowed. For each 2D image slice the user was asked to accept or manually correct the current segmentation result. Manual corrections were done by having the user select points that he or she considered too far from the prostate boundary, and move the selected points to a correct position. After repositioning, the moved points were clamped (not allowed to move) and the algorithm re-deformed the model to better match the corrected points.

F. Evaluation of 3D Segmentation

Evaluation of the algorithm was done using both distance- and volume-based error metrics to compare the algorithm to the gold standard. Algorithm segmentation time (including manual corrections) and the number of manual corrections required were also recorded.

The mean distance (*MD*), mean absolute distance (*MAD*), and maximum absolute distance (*MAXD*) error metrics used in [10] are modified so as to compare multiple 2D image slices and provide 3D error values. The modified equations are:

$$MD = \frac{1}{36 \cdot 72} \sum_{j=1}^{36} \sum_{i=1}^{72} d(\omega_{ji}) \tag{4a}$$

$$MAD = \frac{1}{36 \cdot 72} \sum_{j=1}^{36} \sum_{i=1}^{72} \left| d(\omega_{ji}) \right| \tag{4b}$$

$$MAXD = \max_{i=1}^{36} \left\{ \max_{i=1}^{72} \left\{ \left| d(\omega_{ji}) \right| \right\} \right\} \tag{4c}$$

where $d(\omega_{ji})$ is the distance between the algorithm generated contour and the gold standard contour for each sampled vertex *i* on shape *j* in the set of sequential 2D image slices. Note that the summation is over the 36 slices comprising the 3D image.

The volume-based error metrics percent volume difference (*PVD*), and percent absolute volume difference (*PAVD)* are:

$$PVD = \left(\frac{V_m - V_a}{V_m} \right) \cdot 100\% \tag{5a}$$

$$PAVD = \left(\frac{|V_m - V_a|}{V_m} \right) \cdot 100\% \tag{5b}$$

where V_a is the algorithm generated volume, V_m is the manually generated volume, and each is computed as given in [8]:

$$V = \left(\sum_{k=1}^{36} \left(\sum_{i,j}^{I,J} u_{ij} r_{ij} a_{ij} \right)_n \right) \frac{\pi}{36} \tag{6}$$

where *I,J* are the respective number of rows and columns of the 2D image slice in pixels, $u_{ij} = \begin{cases} 0 & if\ (i,j)\ outside\ contour \\ 1 & if\ (i,j)\ inside\ contour \end{cases}$, r_{ij} is the perpendicular distance from (*i,j*) to the rotational axis for angular sampling, and a_{ij} is the area of pixel (*i,j*). As for 3D distance metrics, the additional summation is over the 36 slices comprising the entire 3D image.

III. Results and Discussion

An example segmentation result in 3D-space is shown in Fig. 1a with two reference images shown at 0° and 90°. Figs. 1b-f show segmentation results on 2D image slices sampled at 0°, 40°, 80°, 120°, and 160°, respectively. In each case, the solid contour represents the output of the

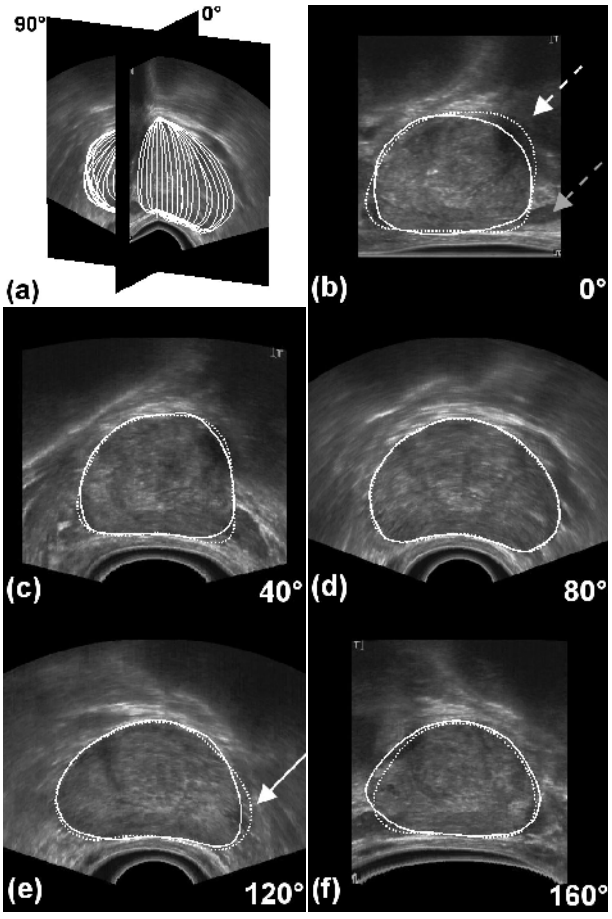

Fig. 1. Example 3D segmentation showing (a) algorithm segmented contours in 3D-space with 0° and 90° image orientation planes, and (b)-(f) sample images with algorithm generated outlines (solid) and manual outlines (dashed) for 2D image slices at 0°, 40°, 80°, 120°, and 160°, respectively. Note that the arrows indicate regions where the algorithm performed poorly: the bladder (dashed white arrow), seminal vesicle (dashed grey arrow), and areas where the US beam is parallel to the prostate boundary (solid white arrow).

algorithm, whereas the dotted contour represents the gold standard. The algorithm performs well in regions where the edges representing the prostate boundary in the image are clearly visible. However, as seen by the arrows in Figs. 1b,e, the algorithm performed poorly in areas where little to no edges were present. In the 2D slices, edges are generally weak where the prostate boundary is roughly parallel to the US beam (solid white arrow). Edges are completely missing near the bladder (dashed white arrow) and seminal vesicle (grey arrow).

TABLE I
AVERAGE DISTANCE- AND VOLUME-BASED ERROR METRICS

	MD (mm)	MAD (mm)	MAXD (mm)	PVD (%)	PVAD (%)
AVG	0.12	1.09	7.27	0.22	3.28
STD	0.45	0.49	2.32	4.58	3.16

Table I shows the average values for distance- and volume-based error metrics comparing the algorithm segmented contours to the gold standard manual outlined contours. The averages were taken over all 36 volumetric images.

TABLE II
AVERAGE ALGORITHM RUN-TIME AND MANUAL CORRECTION

	Time (min)	No. Points Edited	No. Slices Edited	% Slices Edited
AVG	6.9	42.0	9.5	26.3
STD	2.1	20.2	3.3	9.3

Average algorithm runtime and manual correction rates are given in Table II. Listed are the average number of points edited per 3D image, and the average number of slices edited per 3D image as well as the percentage of slices this represents per 3D image. The algorithm is substantially faster than the 45 mins – 1.5 hrs that manual outlining required, and in many cases very little user intervention was required with on average roughly one quarter (26.3±9.28%) of the slices needing correction by the user.

Results from the rotational-based algorithm in [8] showed average error values of PVD = -1.74±3.51% and $PVAD$ = 3.13±2.06%, with an editing rate of 14.4±5.4% (distance-based error metric were not used). The 3D deformable surface used in [11] showed average error values of MD = -0.20±0.28 mm, MAD = 1.19±0.14 mm, $MAXD \approx$ 7.01±1.04 mm, and PVD = 7.16±3.45%, with ~33% of the 3D images requiring editing. Therefore, the performance of the present algorithm (as shown in Table I) is comparable to these previously published algorithms. However, it should be noted that in both [8] and [11] only 6 data sets were used.

IV. CONCLUSION

In the present work a 2D ASM for 3D prostate boundary segmentation from US images was implemented and evaluated. Evaluation of the 3D algorithm found that the algorithm and gold standard boundaries tended to be generally similar, and differences that occurred tended to be in regions of low edge information. Generally, the 3D algorithm was found to be comparable to those previously published.

ACKNOWLEDGEMENTS

Special thanks are extended to Dr. Hans Henrick Thodberg for generously supplying MATLAB code for the building of MDL-based models. Drs. Aaron Fenster and Dónal Downey are thanked for contribution of data and expert advice regarding prostate outlining and US image interpretation. Particular recognition is also extended to the graduate students who participated in the tedious manual outlining of so many prostate data sets. The collection of clinical data was funded by the Canadian Institutes of Health Research (CIHR).

REFERENCES

[1] Canadian Cancer Society/National Cancer Institute of Canada, "Canadian cancer statistics 2006," Toronto, Canada,2006.

[2] American Cancer Society, "Cancer facts & figures 2006," Atlanta, USA,2006.

[3] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models -- their training and application," *Comput. Vis. Image Und.*, vol. 61, no. 1, pp. 38-59, 1995.

[4] Arambula-Cosio F. and B. L. Davies, "Automated prostate recognition: a key process for clinically effective robotic prostatectomy," *Med. Biol. Eng Comput.*, vol. 37, no. 2, pp. 236-243, 1999.

[5] T. F. Cootes, "Statistical models of appearance for computer vision," Manchester, UK,2004.

[6] R. H. Davies, "Learning shape: optimal models for analysing shape variability." 2002.

[7] S. Tong, D. B. Downey, H. N. Cardinal, and A. Fenster, "A three-dimensional ultrasound prostate imaging system," *Ultrasound Med. Biol.*, vol. 22, no. 6, pp. 735-746, 1996.

[8] Y. Wang, H. N. Cardinal, D. B. Downey, and A. Fenster, "Semiautomatic three-dimensional segmentation of the prostate using two-dimensional ultrasound images," *Med. Phys.*, vol. 30, no. 5, pp. 887-897, 2003.

[9] V. Chalana and Y. Kim, "A methodology for evaluation of boundary detection algorithms in medical images," *IEEE Trans. Med. Imaging*, vol. 16, no. 5, pp. 642-652, 1997.

[10] H. M. Ladak, F. Mao, Y. Wang, D. B. Downey, D. A. Steinman, and A. Fenster, "Prostate boundary segmentation from 2D ultrasound images," *Med. Phys.*, vol. 27, no. 8, pp. 1777-1788, 2000.

[11] N. Hu, D. B. Downey, A. Fenster, and H. M. Ladak, "Prostate boundary segmentation from 3D ultrasound images," *Med. Phys.*, vol. 30, no. 7, pp. 1648-1659, 2003.