

A Multi-Modal Prostate Segmentation Scheme by Combining Spectral Clustering and Active Shape Models

Robert Toth^a, Pallavi Tiwari^a, Mark Rosen^b, Arjun Kalyanpur^c, Sona Pungavkar^d, Anant Madabhushi^a

^aRutgers, The State University of New Jersey, NJ, USA. ^bUniversity of Pennsylvania, Philadelphia, PA, USA. ^cTeleradiology Solutions, Bangalore, India. ^dDr. Balabhai Nanavati Hospital, Mumbai, India.

ABSTRACT

Segmentation of the prostate boundary on clinical images is useful in a large number of applications including calculating prostate volume during biopsy, tumor estimation, and treatment planning. Manual segmentation of the prostate boundary is, however, time consuming and subject to inter- and intra-reader variability. Magnetic Resonance (MR) imaging (MRI) and MR Spectroscopy (MRS) have recently emerged as promising modalities for detection of prostate cancer *in vivo*. In this paper we present a novel scheme for accurate and automated prostate segmentation on *in vivo* 1.5 Tesla multi-modal MRI studies. The segmentation algorithm comprises two steps: (1) A hierarchical unsupervised spectral clustering scheme using MRS data to isolate the region of interest (ROI) corresponding to the prostate, and (2) an Active Shape Model (ASM) segmentation scheme where the ASM is initialized within the ROI obtained in the previous step. The hierarchical MRS clustering scheme in step 1 identifies spectra corresponding to locations within the prostate in an iterative fashion by discriminating between potential prostate and non-prostate spectra in a lower dimensional embedding space. The spatial locations of the prostate spectra so identified are used as the initial ROI for the ASM. The ASM is trained by identifying user-selected landmarks on the prostate boundary on T2 MRI images. Boundary points on the prostate are identified using mutual information (MI) as opposed to the traditional Mahalanobis distance, and the trained ASM is deformed to fit the boundary points so identified. Cross validation on 150 prostate MRI slices yields an average segmentation sensitivity, specificity, overlap, and positive predictive value of 89%, 86%, 83%, and 93% respectively. We demonstrate that the accurate initialization of the ASM via the spectral clustering scheme is necessary for automated boundary extraction. Our method is fully automated, robust to system parameters, and computationally efficient.

Keywords: Prostate segmentation, Active shape models (ASMs), Magnetic resonance spectroscopy (MRS), *in vivo* MRI, prostate cancer, segmentation, manifold learning, hierarchical clustering, non-linear dimensionality reduction, spectral clustering, graph embedding.

1. INTRODUCTION

Prostatic adenocarcinoma (CaP) is the second leading cause of cancer related deaths among men in America, with an estimated 220,000 new cases every year (Source: *American Cancer Society*). The current standard for detection of prostate cancer is transrectal ultrasound (TRUS) guided symmetrical needle biopsy which has a high false negative rate associated with it.¹ Recently, multi-modal magnetic resonance (MR) imaging (MRI) comprising both structural MRI and MR Spectroscopy (MRS) have emerged as important tools for early detection of prostate cancer. It has been clinically shown that the integration of MRI and MRS could potentially improve sensitivity and specificity for CaP detection.² Our previous work has consisted of generating computer-aided detection (CAD) systems for detection of prostate cancer. In [3], CAD was performed on *ex vivo* prostate MR images by analyzing 3D texture features. In [4], CAD was performed on *in vivo* prostate MRS data using spectral clustering. In these systems, as well as many other CAD systems, segmentation of the object of interest is a necessary first step, yet segmenting the prostate from *in vivo* MR images is a particularly difficult task. The prostate is especially difficult to see in *in vivo* imagery because of poor tissue contrast on account of MRI related artifacts such as background inhomogeneity. While the identification of the prostate boundary is critical for calculating the prostate volume, for creating patient specific prostate anatomical models, and for CAD systems, accurately identifying the prostate boundaries on an MR image is a tedious task, and manual segmentation is not only laborious but also very subjective.

Previous work on automatic or semi-automatic prostate segmentation has primarily focused on TRUS images.⁵⁻⁸ Ladek et al.⁶ have proposed a scheme in which 4 points are manually selected to initialize a discrete dynamic contour, and manual intervention is used to guide the segmentation. Pathak et al.⁷ presented an algorithm to detect prostate edges which were then used as a guide for manual delineation. In [8], deformable ellipses were used as a shape model to segment the prostate from TRUS images. Only a few prostate segmentation attempts for MR images currently exist. Klein et al.⁹ have proposed segmenting prostate MR images based on nonrigid registration, where they register a series of training ‘atlas’ images to the test image, and the set of atlas images are chosen that best match (based on mutual information) the test image. The selected atlas images are then averaged to achieve a segmentation of the prostate in the test image. Costa et al.¹⁰ have presented a 3D method for segmenting the prostate and bladder simultaneously to account for inter-patient variability in prostate appearance. In [11], polar transformations were used in addition to edge detection techniques such as non-maxima suppression to segment the prostate. A major limitation with previous prostate segmentation attempts is that the prostate varies widely between patients in size, shape, and texture.

A popular segmentation method is the Active Shape Model (ASM), a statistical scheme first developed in the mid-90’s. ASMs use a series of manually landmarked training images to generate a point distribution model, and principal component analysis is performed on this point distribution model to generate a statistical shape model.¹² Then, the Mahalanobis distance (MD) between possible gray values and the mean gray values (determined from the training images) is minimized to identify the boundary of an object.¹²⁻¹⁴ The MD is a statistical measurement used to determine similarity between sets, and is valid as long as the gray values of the training data have a normal distribution. At this point, the shape model is deformed to best fit the boundary points, and the process is repeated until convergence. While ASMs have been used for different applications, results of prostate segmentation yielded highly variable results in [15], with overlap coefficients ranging from about 0.15 to about 0.85. One of the main shortcomings of the ASM is that it requires careful initialization. This is usually done by manual intervention, which can be tedious, and subject to operator variability. Another method is to start at a very low resolution of the image, overlay the starting shape on it, and then increase the resolution, performing the segmentation at each resolution.¹⁶ While this can work, it is not always guaranteed to work. Also, as Ginneken et al. pointed out,¹⁶ since the object of interest could be anywhere in the image, very computationally expensive searches could be required for initialization, contributing to a slow overall convergence time. One such search was proposed by Brejl et al.,¹⁷ who used a shape-variant Hough transform to initialize the segmentation. Seghers et al.¹⁸ presented a highly promising segmentation scheme which searches the entire image, yet state that properly initialized regions of interest (ROIs) would greatly improve their algorithm’s efficiency. Another limitation of ASMs lies with the inherent limitations in using the MD to find the object of interest. Normally, the point with the minimum MD between its surrounding gray values and the mean gray values is assumed to lie on the border of the object.¹²⁻¹⁴ To compute the MD, a covariance matrix of the training gray values is constructed during the training phase, and the MD calculation uses the inverse of that covariance matrix. However, if the covariance matrix is sparse, then the inverse matrix will be undefined, and consequently the MD will be undefined. If the number of pixels sampled is greater than the number of training images, the covariance matrix will become sparse. For example, at least 225 training images would be required to sample a 15×15 window of gray values. Therefore, either having limited training data or attempting to sample a large number of pixels would prove problematic. Also, the MD assumes that the texture training data has an underlying Gaussian distribution, which is not always guaranteed.

In this paper we present a novel, fully automated prostate segmentation scheme that integrates spectral clustering and ASMs. The algorithm comprises 2 distinct stages: spectral clustering of MRS data, followed by an ASM scheme. For the first stage, we perform non-linear dimensionality reduction followed by hierarchical clustering on MRS data to obtain a rectangular ROI, which will serve as the initialization for an ASM. Several non-linear dimensionality reduction techniques were explored, and graph embedding was decided upon to transform the multi-dimensional MRS data into a lower-dimensional space. Graph embedding is a non-linear dimensionality reduction technique in which the relationship between adjacent objects in the higher dimensional space is preserved in the co-embedded lower dimensional space.¹⁹ By clustering of metavoxels in this lower-dimensional space, non-informative spectra can be eliminated. This dimensionality reduction and clustering is repeated hierarchically to yield a bounding box encompassing the prostate. In the second stage, the prostate bounding box obtained from the spectral clustering scheme serves as an initialization for the ASM, in which the mean shape is transformed to fit inside this bounding box. Nearby points are then search to find the prostate border, and the shape is updated accordingly. The afore-mentioned limitations of the MD led us to use mutual information (MI) to calculate the location of the prostate border. Given two images, or regions of gray values I_1 and I_2 , the MI between I_1 and I_2 is an indication of how well the gray values can predict one another.²⁰ It is normally used for registration and alignment

tasks,²⁰ but in this paper we use MI to search for the prostate boundary. For each training image, we take a window of intensity values surrounding each manually landmarked point on the prostate boundary. We then average those intensity values to calculate the mean ‘expected’ intensity values. The idea behind the method is that if a pixel lies on the prostate border, then the MI between its surrounding gray values the mean intensity values of the border will be maximum. The advantages of using MI over MD are the following: (1) the number of points sampled is not dependent on the number of training images, and (2) MI does not require an underlying Gaussian distribution, as long as the gray values are predictive of one another. Finally, once a set of pixels presumably located on the prostate border are determined (henceforth referred to as ‘goal points’), we update the shape to best fit these goal points. We introduce a weighting scheme for fitting the goal points, which was first proposed by Cootes et al.¹² The goal points are weighted using two values. The first value is the normalized MI value. MI is normalized by the Entropy Correlation Coefficient (ECC),²¹ which rescales the MI values to be between 0 and 1. The second weighting value is how well the shape fit each goal point during the previous iteration, which is scaled from 0 to 1. We call this the ‘outlier weight,’ where if the shape model couldn’t deform close to a goal point, it is given a value close to 0, and if the shape model was able to deform close to a goal point, it is given a value close to 1. These two terms are multiplied together to obtain the final weighting factor for each landmark point. It’s important to note that as in traditional ASM schemes, the off-line training phase needs to only be done once, while the on-line segmentation is fully automated.

The primary contributions and novel aspects of this work are:

- A fully automated scheme, by performing spectral clustering on MRS data to obtain a bounding box, used to initialize an ASM search.
- Using MI instead of MD to find points on the prostate border.
- Using outlier distances and MI values to weight the landmark points.

Finally, an exhaustive evaluation of the model via randomized cross validation is performed to assess segmentation accuracy against expert segmentations. In addition, model parameter sensitivity and segmentation efficiency are also assessed. The rest of the paper will be organized as follows. In Section 3, we present the methodology for initializing the bounding box by performing spectral clustering on available MRS data. Then, in Section 4, we present the methodology for performing the segmentation, which will include finding the prostate border using MI, and updating the shape based on a weighting scheme. This will be followed by Section 5, giving the results compared to an expert radiologist’s segmentations.

2. SYSTEM OVERVIEW

2.1 Notation

We define a spectral scene $\hat{C} = (\hat{C}, \hat{f})$ where \hat{C} is a 3D grid of metavoxels. For each spatial location $\hat{c} \in \hat{C}$ there is an associated 256-dimensional valued spectral vector $\hat{F}(\hat{c}) = [\hat{f}_j(\hat{c}) \mid j \in \{1, \dots, 256\}]$, where $\hat{f}_j(\hat{c})$ represents the concentration of different biochemicals (such as creatinine, citrate, and choline). We define the associated MR intensity scene $\mathcal{C} = (C, f)$ where C represents a set of spatial locations, and $f(c)$ represents a function that returns the intensity value at any spatial location $c \in C$. It’s important to note that the distance between any two adjacent metavoxels $\hat{c}, \hat{d} \in \hat{C}$, $\|\hat{c} - \hat{d}\|$, (where $\|\cdot\|$ denotes the L_2 norm) is roughly 16 times the distance between any two adjacent spatial voxels $c, d \in C$. We define a κ -neighborhood centered on $c \in C$ as $\mathcal{N}_\kappa(c)$ where for $\forall d \in \mathcal{N}_\kappa(c)$, $\|d - c\| \leq \kappa$, $c \notin \mathcal{N}_\kappa(c)$. In addition, the set of pixel intensities for the κ -neighborhood centered on c is defined as $F_\kappa(c) = [f(d) \mid d \in \mathcal{N}_\kappa(c)]$. Finally, for any set C , we define $|C|$ as the cardinality of C . Table 1 lists the notation and commonly used symbols for the paper.

2.2 Data Description

The spectral datasets (consisting of both MRI and MRS data) used for the study were collected during the ACRIN multi-site trial (http://www.acrin.org/6659_protocol.html). All the MRS and MRI studies were 1.5 Tesla. The MRI studies were axial T2 images obtained from 19 patients. The 19 3D studies comprised a total of 148 image sections. These sections correspond to either the base, midgland, or apex of the prostate. Three distinct 2D ASM models were constructed, one each for the base, midgland, and apex. The ground truth was determined by manual outlining of the prostate border on each section by an expert radiologist.

Table 1. List of notation and symbols used.

Symbol	Description	Symbol	Description
\mathcal{C}	MRI scene	$\hat{\mathcal{C}}$	MRS scene
C	Set of pixel coordinates in \mathcal{C}	\hat{C}	Set of metavoxel coordinates in $\hat{\mathcal{C}}$
c	A pixel in C	\hat{c}	A metavoxel in \hat{C}
$f(c)$	Intensity value at c	$\hat{F}(\hat{c})$	Spectral content at \hat{c}
$\mathcal{N}_\kappa(c)$	κ -neighborhood near c	$F_\kappa(c)$	Set of intensity values for $\forall d \in \mathcal{N}_\kappa(c)$
\mathbf{X}	Set of landmark points, where $\mathbf{X} \subset C$	$T(\mathbf{X})$	Affine transformations T applied to \mathbf{X}
$\tilde{\mathbf{X}}$	Set of goal points where $\tilde{\mathbf{X}} \subset C$	\mathbf{b}	Variable controlling the ASM shape
\bar{g}_m	Expected gray values for the m^{th} landmark point	K	Number of training images

2.3 Brief Outline of Methods

Figure 1 illustrates the modules and the pathways comprising our automated segmentation system. The first step consists of performing non-linear dimensionality reduction (graph embedding) on spectra $\hat{F}(\hat{c})$ to embed these spectra non-linearly in a lower dimensional space. The embedded spectra, represented now as their corresponding dominant Eigenvectors, are clustered into clusters corresponding to informative (within or around the prostate) and non-informative spectra (background spectra lying outside the prostate). The spectra in the dominant, non-informative cluster are eliminated. This process repeats until a bounding box containing the prostate is obtained. At this point, the set of prostate training images are manually landmarked, from which the shape and intensity information is obtained. This landmarking can be done off-line, and only needs to be done once to construct a trained ASM model consisting of both shape and intensity information. The mean shape is then transformed to the bounding box previously obtained from the MRS data, which serves as the landmark points for the first iteration. Then, for each iteration, a technique maximizing MI is performed to find pixels on the prostate border. Finally, these landmarks are weighted, and the shape is updated to best fit the points on the border, yielding the next iteration’s landmark points. The process is repeated until convergence is obtained, resulting in a segmentation of the prostate.

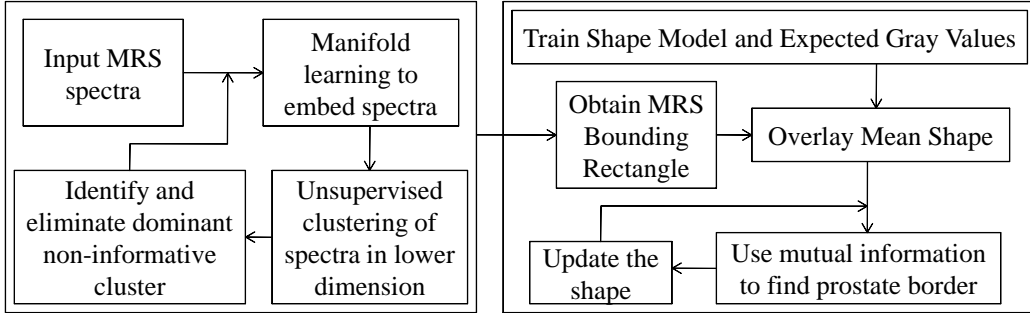


Figure 1. Flowchart for segmentation algorithm, with the MRS module shown on the left and the ASM module shown on the right.

3. MRS METHODOLOGY

3.1 Non-Linear Dimensionality Reduction via Graph Embedding

The spectra $\hat{F}(\hat{c})$, for $\hat{c} \in \hat{C}$, lie in a 256 dimensional space. Hence, our aim is to find an embedding vector $G(\hat{c})$ for $\forall \hat{c} \in \hat{C}$, and its associated class ω (informative or non-informative) such that if the distances between elements of ω are well preserved in the lower dimensional space. Hence if metavoxels $\hat{c}, \hat{d} \in \hat{C}$ both belong to class ω , then $\|G(\hat{c}) - G(\hat{d})\|$ should be small. To compute the optimal embedding, we first define a matrix $W \in \mathbb{R}^{|\hat{C}| \times |\hat{C}|}$ representing the similarity between all objects in \hat{C} . For $\forall \hat{c}, \hat{d} \in \hat{C}$, W is defined as

$$W(L(\hat{c}), L(\hat{d})) = e^{-\|\hat{F}(\hat{c}) - \hat{F}(\hat{d})\|} \quad (1)$$

where $L(\hat{c})$ represents a unique index position of $\hat{c} = (x, y, z)$ derived from its x, y, z coordinates in 3D space. The embedding vector G is obtained from the maximization of the function

$$E_W(G) = 2\gamma \frac{G^T(D - W)G}{G^T D G}, \quad (2)$$

where $\gamma = |\hat{C}| - 1$. In addition, D is a diagonal matrix where for $\forall \hat{c} \in \hat{C}$, the diagonal element $(L(\hat{c}), L(\hat{c}))$ is defined as $D(L(\hat{c}), L(\hat{c})) = \sum_{\hat{d} \in \hat{C}} W(L(\hat{c}), L(\hat{d}))$. The embedding space is defined by the Eigenvectors corresponding to the smallest β Eigenvalues of $(D - W)G = \lambda D G$. A matrix $M \in \mathbb{R}^{|\hat{C}| \times \beta}$ of the first β Eigenvectors is constructed, and for $\forall \hat{c} \in \hat{C}$, $G(\hat{c})$ is defined as row $L(\hat{c})$ of M . $G(\hat{c})$ is therefore a vector consisting of element number $L(\hat{c})$ from each of the first β Eigenvectors, which represents the β -dimensional Cartesian coordinates. The graph embedding algorithm is summarized below.

Algorithm *GraphEmbedding*

Input: $\hat{C}, \hat{F}(\hat{c})$ for $\forall \hat{c} \in \hat{C}$

Output: $G(\hat{c})$ for $\forall \hat{c} \in \hat{C}$

begin

0. Define a unique element mapping L , where $1 \leq L(\hat{c}) \leq |\hat{C}|$ and $L(\hat{c}) \neq L(\hat{d})$ for $\forall \hat{c}, \hat{d} \in \hat{C}$;
1. Define a $|\hat{C}| \times |\hat{C}|$ matrix W , where $W(L(\hat{c}), L(\hat{d})) = e^{-\|\hat{F}(\hat{c}) - \hat{F}(\hat{d})\|}$ for $\forall \hat{c}, \hat{d} \in \hat{C}$;
2. Define a $|\hat{C}| \times |\hat{C}|$ matrix D , where each diagonal element has a value of $D(L(\hat{c}), L(\hat{c})) = \sum_{\hat{d} \in \hat{C}} W(L(\hat{c}), L(\hat{d}))$;
3. Calculate the Eigenvalues and Eigenvectors of $(D - W)G = \lambda D G$;
4. Construct a $|\hat{C}| \times \beta$ matrix M of the first β Eigenvectors, where each column is an Eigenvector;
5. For $\forall \hat{c} \in \hat{C}$, define the vector $G(\hat{c})$ as row $L(\hat{c})$ of M ;

end

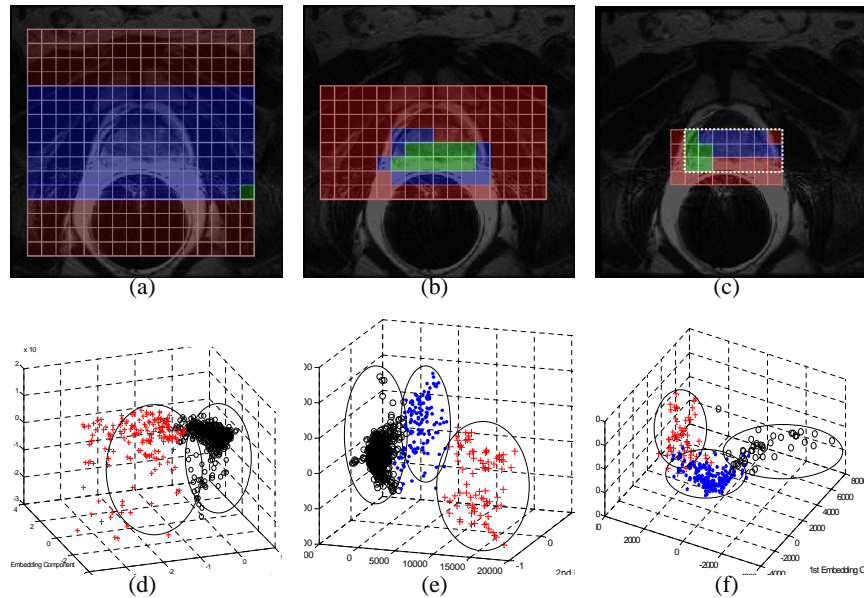


Figure 2. Spectral grids for a single slice within \hat{C} are shown superimposed on the T2 MRI section for (a) \hat{C}_0 , (b) \hat{C}_1 , and (c) \hat{C}_2 . Note that the size of the grid reduces from 16×16 metavoxels (a) to 6×3 metavoxels (c) by elimination of non-informative spectra on the 16×16 grid. Note the white box (shown with dashed lines) in (c), representing the rectangular ROI centered on the prostate. Also note the embedding plots of (d) \hat{C}_0 , (e) \hat{C}_1 , and (f) \hat{C}_2 clustered each into the 3 clusters V_h^1, V_h^2 , and V_h^3 in lower dimensional space.

3.2 Hierarchical Cascade to Prune Non-informative Spectra

At each iteration h , a subset of metavoxels $\hat{C}_h \subset \hat{C}$ is obtained by eliminating the non-informative spectra. The metavoxels $\hat{c} \in \hat{C}_h$ are aggregated into clusters V_h^1, V_h^2, V_h^3 by applying k -means clustering to all $\hat{c} \in \hat{C}_h$ in the low dimensional embedding $G_h(\hat{c})$. Initially, most of the locations $\hat{c} \in \hat{C}_h$ correspond to zero padded or non informative spectra. Therefore, while the unsupervised clustering results in 3 clusters, the dominant cluster (the cluster with the most number of elements) is non-informative and is eliminated. Figure 2(a) represents a 16×16 MRS grid, superimposed on the corresponding T2 MR image. After the first iteration, the outer voxels containing zero padded spectra are removed, and the grid becomes an 8×16 grid (Fig. 2(b)). The scheme is recursed until a final set of metavoxels \hat{C}_H is obtaining, serving as an ROI containing the prostate (Fig. 2(c)). While H can represent a constant number of iterations, the MR data actually contains the information about the size of the prostate, thereby allowing the desired value of $|\hat{C}_H|$ to be known beforehand. The algorithm below describes precisely how our methodology works. Also note that while it is presented with a constant H as an input, it is readily extendable to include prior information about the desired value of $|\hat{C}_H|$.

Algorithm *HierarclustMRSprostate*

Input: $\hat{F}(\hat{c})$ for $\forall \hat{c} \in \hat{C}$, H

Output: \hat{C}_H

begin

0. Initialize $\hat{C}_0 = \hat{C}$;
1. for $h = 0$ to $(H - 1)$ do
 2. Calculate $G_h(\hat{c})$ for $\forall \hat{c} \in \hat{C}_h$ as $G_h = \text{GraphEmbedding}(\hat{C}_h, \hat{F})$;
 3. Apply k -means clustering on $G_h(\hat{c})$ for $\forall \hat{c} \in \hat{C}$ to obtain clusters V_h^1, V_h^2, V_h^3 ;
 4. Identify largest cluster V_h^{\max} ;
 5. Create set $\hat{C}_{h+1} \subset \hat{C}_h$ by eliminating all $\hat{c} \in V_h^{\max}$ from \hat{C}_h ;
6. *endfor*

end

4. ASM METHODOLOGY

4.1 Training the Shape Model and Expected Gray Values

Step 1: The training phase begins by using training images of oblique axial slices and manually selecting M landmarks on the prostate boundary. The set of K training images is denoted as $S^t = \{C^\alpha \mid \alpha \in \{1, \dots, K\}\}$. The apex and the 2 posterolateral-most points on the prostate are landmarked, and the remaining $(M - 3)$ landmarks are equally spaced between these landmarks. Therefore, each training image $C^\alpha \in S^t$ has a corresponding set of landmarks $\mathbf{X}^\alpha \subset C^\alpha$, where $\mathbf{X}^\alpha = \{c_m^\alpha \mid m \in \{1, \dots, M\}\}$, and where $c_m^\alpha = (x_m^\alpha, y_m^\alpha)$ denotes the coordinates of the m^{th} landmark point in C^α .

Step 2: These shapes are aligned using a modified Procrustes analysis, as described in [14]. Figures 3(a) and 3(b) show the training landmark points before and after alignment. Once the training shapes have been aligned, the mean shape $\bar{\mathbf{X}}$ is defined as $\bar{\mathbf{X}} = \{\bar{c}_m \mid m \in \{1, \dots, M\}\}$, where $\bar{c}_m = (\bar{x}_m, \bar{y}_m)$ and $\bar{x}_m = \frac{1}{K} \sum_{\alpha} x_m^\alpha$ and $\bar{y}_m = \frac{1}{K} \sum_{\alpha} y_m^\alpha$. Then, principal component analysis is performed as per the technique described in [12], so that any valid prostate shape \mathbf{X} can be represented as

$$\mathbf{X} = T(\bar{\mathbf{X}} + \mathbf{P} \cdot \mathbf{b}) \quad (3)$$

where T is an affine transformation mapping, \mathbf{b} is the variable controlling the shape, and \mathbf{P} is a matrix with each column representing an Eigenvector. Out of the M Eigenvectors, only the first z are needed to explain most (98%) of the training variation, where if λ_r is the r^{th} Eigenvalue, z is as small as possible such that $(\sum_{r=1}^z \lambda_r) \geq (0.98 \cdot \sum_{r=1}^M \lambda_r)$. Therefore, $|\mathbf{b}| = z$, and \mathbf{P} is a matrix consisting of only the first z Eigenvectors. In our experiments, we represented valid prostate shapes as vector \mathbf{b} of length 18, with each element in \mathbf{b} constrained between ± 2.5 standard deviations, with one standard deviation for the r^{th} element in \mathbf{b} given as $\sqrt{\lambda_r}$.

Step 3: The mean gray values must now be calculated for each landmark point, which will represent the expected gray values for the prostate border. First, we take the κ -neighborhood centered on each c_m^α as $\mathcal{N}_\kappa(c_m^\alpha)$, where $\alpha \in \{1, \dots, K\}$. Then, for $\forall d_u^\alpha \in \mathcal{N}_\kappa(c_m^\alpha)$, we define $\bar{f}(d_u) = \frac{1}{K} \sum_{\alpha} f(d_u^\alpha)$, where $u \in \{1, \dots, |\mathcal{N}_\kappa(c_m)|\}$. Finally, the expected gray values are defined as $\bar{g}_m = \bar{F}_\kappa(c_m)$ where $\bar{F}_\kappa(c_m) = [\bar{f}(d_u) \mid u \in \{1, \dots, |\mathcal{N}_\kappa(c_m)|\}]$. This concludes the off-line training phase of the segmentation system, with the variables $\bar{\mathbf{X}}$, \mathbf{P} , and $\bar{g}_m, m \in \{1, \dots, M\}$ comprising a trained ASM model.

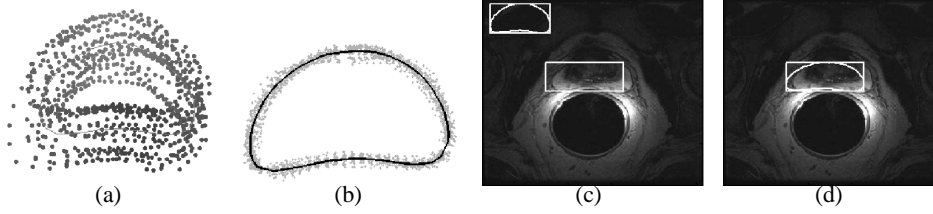


Figure 3. Training landmark points in (a), aligned in (b) with $\bar{\mathbf{X}}$ shown in black. $\bar{\mathbf{X}}$, encompassed by the smallest possible rectangle, is shown in the top left corner of (c). In (d), $\bar{\mathbf{X}}$ is seen highlighted on the MRS bounding box after being transformed to yield the initial landmarks \mathbf{X}^1 .

4.2 Initializing the ASM and Using Mutual Information to Find the Prostate Border

At this point, we define a new image to search for the prostate as the scene $\mathcal{C} = (C, f)$ where $\mathcal{C} \notin S^t$. The very first step of the system is to initialize the landmarks for the first iteration, \mathbf{X}^1 (where $\mathbf{X}^1 \subset C$). To do this, the smallest rectangle containing $\bar{\mathbf{X}}$, shown in Fig. 3(c), has its corners aligned to the corners of the bounding box obtained from the MRS data. The scaling and translation transformations T used to align the two rectangles are applied to $\bar{\mathbf{X}}$, yielding the initial landmark points as $\mathbf{X}^1 = T(\bar{\mathbf{X}})$. Figure 3(d) shows \mathbf{X}^1 initialized to the MRS rectangle.

A set of points presumed to lie on the prostate border must now be calculated. For the n^{th} iteration, the current landmark points $c_m^n, m \in \{1, \dots, M\}$ are denoted by the set \mathbf{X}^n and the set of points $\tilde{c}_m^n, m \in \{1, \dots, M\}$ presumed to lie on the prostate border is denoted by the set $\tilde{\mathbf{X}}^n$. The pixels surrounding each c_m^n , denoted as $\mathcal{N}_v(c_m^n)$, are all possible locations for the prostate border. For $\forall c_j \in \mathcal{N}_v(c_m^n)$, a set of nearby intensity values $F_\kappa(c_j)$ is compared with \bar{g}_m using mutual information (MI). One common way to calculate a normalized MI value is by calculating the Entropy Correlation Coefficient (ECC)²¹ between 2 sets of values. If a and b are 2 sets of intensity values, we denote the normalized MI value between them as $\text{ECC}(a, b)$. The location c_j with the highest mutual information (MI) value between $F_\kappa(c_j)$ and \bar{g}_m is denoted as \tilde{c}_m^n . Therefore, we now define the set of goal points as $\tilde{\mathbf{X}}^n = \{\tilde{c}_m^n \mid m \in \{1, \dots, M\}\}$ where

$$\tilde{c}_m^n = \underset{c_j}{\text{argmax}} (\text{ECC}(F_\kappa(c_j), \bar{g}_m)) \quad (4)$$

and $c_j \in \mathcal{N}_v(c_m^n)$. Figure 4(a) shows the MI values for each $c_j \in \mathcal{N}_v(c_m^n)$ as pixel intensities, with the brightest pixel representing \tilde{c}_m^n . Ideally, \tilde{c}_m^n would lie exactly on the border of the prostate. It is important to note that the number of pixels sampled ($|\mathcal{N}_v|$ and $|\bar{g}_m|$) need not necessarily equal the number of pixels to search ($|\mathcal{N}_v|$). In fact, the windows don't even have to be the same shape, although for our purposes we consistently used rectangles, as seen in Fig. 4(a).

4.3 Updating the Shape

\mathbf{X}^n must now be deformed to $\tilde{\mathbf{X}}^n$. At this point we introduce a weighting scheme for updating the shape. The vector of weights is denoted as $\Gamma^n = \{\Gamma_m^n \mid m \in \{1, \dots, M\}\}$, where each Γ_m^n is a weight associated with the landmark point c_m^n . First, the MI values for each goal point $\tilde{c}_m^n \in \tilde{\mathbf{X}}^n$ are compared to one another. The goal point with the highest MI value between $F_\kappa(\tilde{c}_m^n)$ and \bar{g}_m is given a weight of 1, the goal point with the lowest MI value is given a weight of 0, and the remaining points are linearly rescaled to lie between 0 and 1. The reasoning is that a high normalized MI value indicates that the intensity values surrounding the goal point are highly predictive of the mean of the training data, and should be weighted heavily. After the first iteration, each MI weight value is multiplied by an ‘outlier weight,’ which is based on how well the shape model was able to deform to the goal points during the previous iteration. Each MI weight value is multiplied by a number between 0 and 1, which is calculated as a linear rescaling of $\|\tilde{c}_m^{n-1} - c_m^n\|^{-1}$. This product is the final weight for each landmark point, Γ_m^n . Figure 4(b) shows several outlier weights where \tilde{c}_m^{n-1} are shown as white diamonds and all c_m^n are indicated by the white line. Most ASM systems don't allow the current iteration to deform completely to the goal points. Coates et al., for example, limit the number of pixels that \mathbf{X}^n can move to between 0 and 4 pixels per iteration.²² We have chosen instead to linearly rescale Γ so that $\min_m(\Gamma) = 0.25$ and $\max_m(\Gamma) = 0.75$. This will prevent the shape model from completely fitting the goal points, as each landmark point can only possibly move between 25% and 75% the distance to its corresponding goal point. These numbers don't necessarily have to be 0.25 and 0.75, but we felt these were reasonable constraints. Next, the landmarks \mathbf{X}^n are aligned to the goal points $\tilde{\mathbf{X}}^n$ using affine transformations. This is done using a weighted Procrustes analysis,¹⁴ with the weight for the m^{th} landmark defined as Γ_m^n , resulting in an affine

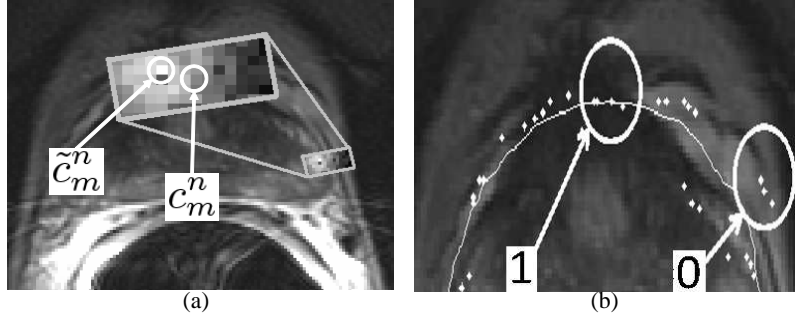


Figure 4. (a) Search area $\mathcal{N}_v(c_m^n)$ shown in the gray rectangle with the MI values for each $c_j \in \mathcal{N}_v(c_m^n)$ shown as pixel intensities. A brighter intensity value indicates a higher MI value between $F_\kappa(c_j)$ and \bar{g}_m^n . Also shown is the point with the highest MI value, \tilde{c}_m^n . Shown in (b) are outlier weights with \mathbf{X}^n shown as the white line and $\bar{\mathbf{X}}^{n-1}$ shown as white diamonds. Note that when the shape has deformed very closely to a goal point, the weighting is a 1 and when a goal point is too much of an outlier, the weighting is a 0.

transformation mapping T . However, only using a global affine transformation is not enough to optimally reach the goal points, which is the reason statistical shape models are generated. As described in [12], the shape parameter for the n^{th} iteration is defined as

$$\mathbf{b}^n = \begin{cases} \emptyset & \text{if } n = 0, \\ \mathbf{b}^{n-1} + \mathbf{P}^{-1} \cdot \mathbf{d}\mathbf{x} & \text{if } n > 0, \end{cases} \quad (5)$$

where $\mathbf{d}\mathbf{x} = \{dx_m \mid m \in \{1, \dots, M\}\}$ and $dx_m = \|\tilde{c}_m^n - T(c_m^n)\| \cdot \Gamma_m$, which allows the weights to be taken into account when updating the shape. Finally, the affine transformations T are applied to this new shape, yielding the next iteration's landmarks as

$$\mathbf{X}^{n+1} = T(\bar{\mathbf{X}} + \mathbf{P} \cdot \mathbf{b}^n). \quad (6)$$

The system then repeats searching for new goal points and updating the shape, until the mean Euclidean distance between \mathbf{X}^n and \mathbf{X}^{n+1} is less than 0.2 pixels, at which time convergence is assumed.

4.4 Evaluation Methods

Once the final landmarks have been determined, this shape is compared to a manually outlined expert segmentation. The area based metrics used to evaluate the segmentation system are positive predictive value (PPV), specificity, sensitivity, and global overlap coefficient, with values closer to 1 indicating better results.^{3,18} The edge based metric used to evaluate the system are mean absolute distance (MAD) and Hausdorff distance, which are given in units of pixels with values closer to 0 indicating better results.²³ To evaluate the system, the optimal parameters are first computed, which is followed by a randomized cross validation. To perform the cross validation, we let the variable q represents the number of images to test with, and P represent the total number of images in a certain group (either apex, midgland, or base). For each group, we randomly select q images to test with, and generate an ASM model with $(P - q)$ images. Averaging the metric results for those q images gives a set of mean values (μ) for sensitivity, specificity, overlap, PPV, MAD, and Hausdorff distance. We repeat this 50 times and report the mean metric values ($\bar{\mu}$'s), and the standard deviations (σ 's) over the 50 iterations.

4.5 Parameter Selection and Sensitivity Analysis

To evaluate the system, the optimal parameter values must be computed. First, the search area (\mathcal{N}_v) was set as a 15×3 rectangle, centered on the current landmark point and rotated to face in the normal direction of the shape at that point, as shown in Fig. 4(a). This was assumed to be large enough to encompass the prostate border given a decent initialization, without compromising computational costs. Using this search size, the best sampling sizes (\mathcal{N}_κ) were determined for each group of images. For sampling, we used rectangles rotated in the same direction as the search rectangles. The parameters that were tested were the sampling rectangle's height and width, each varied between 1 and 50 pixels. The best sampling rectangle sizes were determined to be 7×13 , 37×39 , and 33×31 pixels for the base, midgland, and apex respectively.

After performing the repeated segmentations with different sampling lengths and widths, we compared how many different values gave at least a minimum metric value, shown in Fig. 5. For example, 100% of the sampling sizes tested (y-axis) achieved sensitivity values of at least 0.60 (x-axis), while only 3% of the sampling sizes (y-axis) achieved sensitivity values of at least 0.88 (x-axis).

Once the optimal parameters were determined, we evaluated the importance of a correct initialization by randomly changing the initialization bounding box. We chose random ROIs to initialize the ASM system, and compared these to the ROIs obtained from the MRS data. Figure 6 shows the effects of changing the initialization, where the distance between the corners of the MRS ROI and the random ROI is plotted on the x-axes. The downward slopes in Fig. 6 show that the segmentation becomes worse as the initial ROI is moved further away from the MRS box, suggesting that a proper initialization is vital for an accurate segmentation, and that the spectral clustering of MRS data can provide this required initialization.

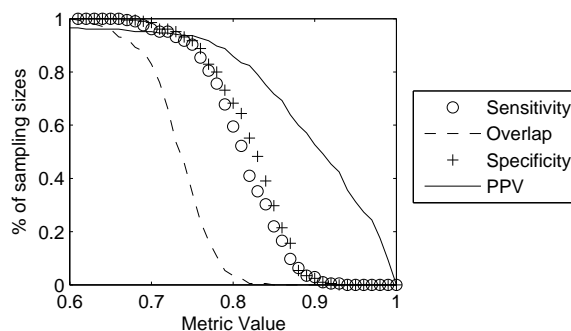


Figure 5. A graph of how many different sampling sizes (y axis) achieved a certain metric value (x axis).

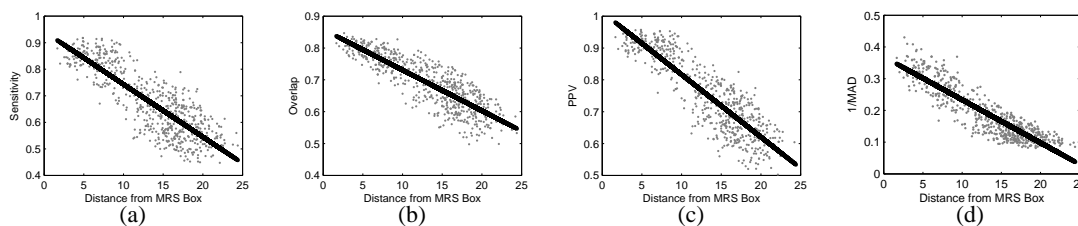


Figure 6. Effect of changing the initialization where the x-axis represents the distance (in pixels) between the corners of the initialization used and the corners of the MRS box. The y-axis represents the sensitivity values in (a), the overlap values in (b), the PPV values in (c), and the MAD values in (d). Note the downward slopes in (a)-(d), which suggest that on average the segmentation becomes worse as the initialization is moved away from the MRS box.

5. RESULTS AND DISCUSSION

5.1 Qualitative Results

Several qualitative results for the MRS initialization are shown in Fig. 7. In each of these images, the final bounding box from the MRS spectral clustering scheme is shown in white, overlaid on the original MR images. In all of the results, the bounding box obtained from MRS data accurately encompassed the prostate, validating the use of spectral clustering on MRS data for initializing a segmentation system. Qualitative results for the resulting segmentation are shown in Fig. 8. In these images, the ground truth (as determined by an expert) are shown in black, and the results from the automatic segmentation system are shown in white. The prostate images shown in Fig. 7 don't correspond to the prostate images in Fig. 8, but in both figures there is at least one image from the base, midgland, and apex.

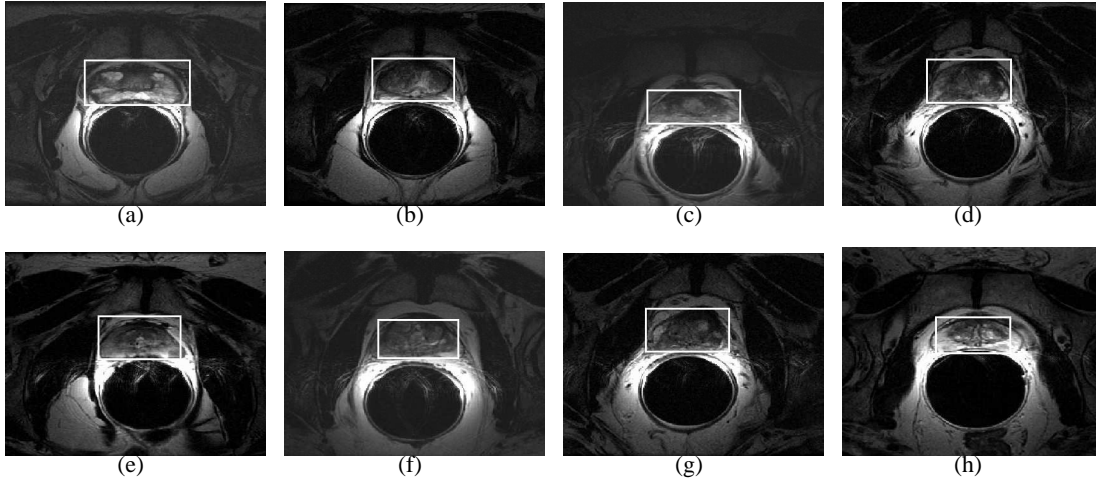


Figure 7. (a) - (h) show the 8 different initialization bounding boxes obtained from the MRS data in white, overlaid on the original MR images. Note that in each case the bounding box contains the prostate.

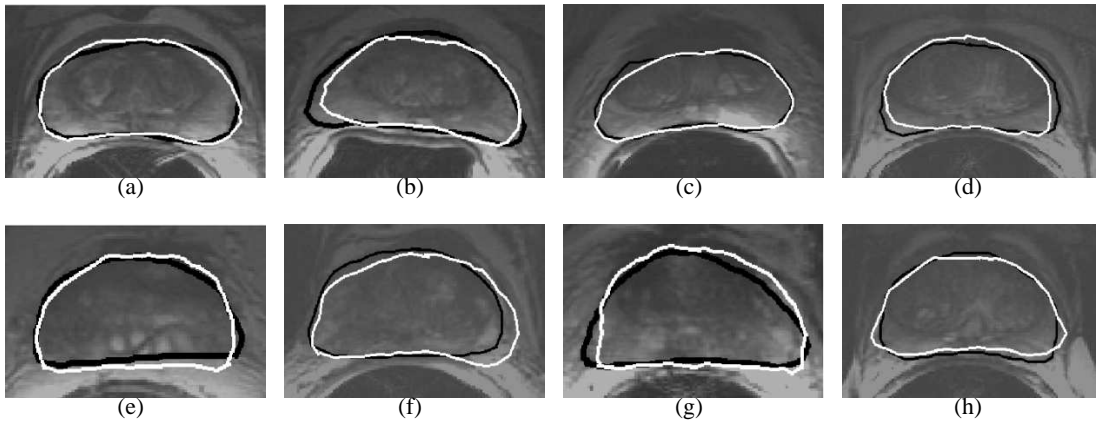


Figure 8. (a) - (h) show qualitative results with ground truth in black, and results from automatic segmentation of the prostate in white.

5.2 Quantitative Results

Table 2 shows the results from the randomized cross validation described in Section 4.4. We also calculated a confidence interval for the mean values, which is derived from the Student's t -distribution and the standard error. So for any of the metric values, after repeating 50 times, the 99% confidence interval is given as $\bar{\mu} \pm 0.3787\sigma$. The images from the base had the best results, yielding sensitivity, specificity, overlap, and PPV values as high as 0.89, 0.86, 0.83, and 0.92 respectively, and MAD and Hausdorff distance as low as 2.6 and 8.2 pixels respectively. The apex, however, yielded the worst results due to high variability in appearance between images, as well as the close proximity to surrounding tissues. Also, since q represents the number of images to test with ($q \in \{3, 5, 10\}$), a higher value of q indicates less images to train with, which was shown to slightly decrease accuracy for each group. Finally, the PPV values were consistently the highest values, suggesting that the segmentations tended to have minimal false positive area.

Table 2. Results over 50 trials, where q represents the number of images to leave and out and test with for each trial, $\bar{\mu}$ represents the mean metric value over all 50 trials, and σ represents the standard deviation over all 50 trials.

Group	q	Sensitivity		Specificity		Overlap		PPV		MAD		Hausdorff	
		$\bar{\mu}$	σ	$\bar{\mu}$	σ	$\bar{\mu}$	σ	$\bar{\mu}$	σ	$\bar{\mu}$	σ	$\bar{\mu}$	σ
Base	3	0.892	0.048	0.863	0.056	0.826	0.036	0.924	0.035	2.625	0.599	8.192	1.912
	5	0.884	0.052	0.852	0.050	0.818	0.047	0.917	0.044	2.624	0.551	8.298	1.727
	10	0.883	0.031	0.855	0.032	0.815	0.028	0.920	0.024	2.804	0.371	8.791	1.074
Midgland	3	0.846	0.073	0.863	0.061	0.785	0.062	0.925	0.037	3.467	1.222	9.594	2.597
	5	0.872	0.038	0.864	0.045	0.810	0.034	0.926	0.029	3.012	0.641	9.008	1.646
	10	0.851	0.030	0.862	0.030	0.789	0.027	0.924	0.019	3.367	0.543	9.556	1.451
Apex	3	0.847	0.075	0.817	0.079	0.749	0.064	0.881	0.067	3.821	1.085	11.06	3.106
	5	0.824	0.056	0.823	0.059	0.733	0.050	0.887	0.046	3.927	0.758	11.06	1.863
	10	0.833	0.050	0.810	0.049	0.730	0.037	0.876	0.039	4.004	0.581	11.24	1.514

6. CONCLUDING REMARKS

In this paper, we have presented a scheme that integrates dimensionality reduction and hierarchical clustering on MRS data to yield a fully automatic and accurate ASM based prostate segmentation method on *in vivo* MR data. Performing spectral clustering on MRS data to obtain an initial ROI is a significant novel aspect of our work, as the method is fully automated and does not require any manual initialization. In addition, we have also shown how MI could be successfully used as part of an ASM system, providing yet another novel aspect to our scheme. Finally, we have performed an exhaustive evaluation to demonstrate the accuracy, efficiency, and robustness of our system. Comparison of our segmentation system with other prostate segmentation schemes show that our system performs at least as well, if not better, than other systems. Klein et al.⁹ have reported a mean overlap of 0.82, and Zhu et al.¹⁵ have overlap coefficients ranging from about 0.15 to about 0.85, while our mean overlap coefficients range from 0.730 to 0.826. Costa et al.¹⁰ result in mean sensitivity and PPV values of 0.75 and 0.80 respectively for the prostate, while we achieved mean sensitivity and PPV values of 0.892 and 0.926 respectively.

However, there are certain cases in which our automatic segmentation performs poorly (Fig. 9). Figure 9(a) shows the poor initialization resulting in the segmentation shown in Fig. 9(b). Figure 9(c) shows a proper initialization, which still failed because of the lack of a clear prostate edge at the bottom right corner of the prostate (Fig. 9(d)). Limitations include high variability in prostate appearance between patients, as well as problems such as unclear edges, or edges that are simply not predictive of one another. Future work will attempt to overcome these limitations.

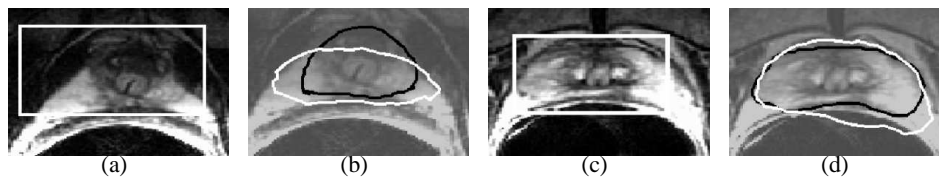


Figure 9. Two different initializations are shown in (a) and (c), with their respective resulting segmentations shown in (b) and (d). In (a) and (c), the rectangular bounding box, calculated from the MRS data, is shown as a white rectangle. In (b) and (d), the ground truth is shown in black, and the results from the automatic segmentation is shown in white. Note that the MR images shown in (b) and (d) are the same images shown in (a) and (c) respectively.

ACKNOWLEDGMENTS

This work was possible due to grants from the Coulter Foundation (WHCF4-29349, WHCF 4-29368), the Busch Biomedical Award, the Cancer Institute of New Jersey, the New Jersey Commission on Cancer Research, the National Cancer Institute (R21CA127186-01, R03CA128081-01), and the Aresty Foundation for Undergraduate Research. The authors would also like to acknowledge Johnson Henry from Teleradiology Solutions, Bangalore, India for his assistance in the expert segmentations. Finally, the authors would like to acknowledge the ACRIN database for the MRI/MRS data.

REFERENCES

1. W. Catalona *et al.*, "Measurement of Prostate-Specific Antigen in serum as a Screening Test for Prostate Cancer," *J.Med* **324**(17), pp. 1156–1161, 1991.
2. J. J. Hom, F. V. Coakley, J. P. Simko, Y. Lu, A. Qayyum, A. C. A. Westphalen, L. D. Schmitt, P. R. Carroll, and J. Kurhanewicz, "High-Grade Prostatic Intraepithelial Neoplasia in Patients with Prostate Cancer: MR and MR Spectroscopic Imaging Features—Initial Experience," *Radiology* **242**(2), pp. 483–489, 2006.
3. A. Madabhushi, M. Feldman, D. Metaxas, J. Tomaszewski, and D. Chute, "Automated detection of prostatic adenocarcinoma from high-resolution ex vivo mri," *IEEE Trans Med Imaging* **24**, pp. 1611–1625, Dec. 2005.
4. P. Tiwari, A. Madabhushi, and M. Rosen, "A hierarchical unsupervised spectral clustering scheme for detection of prostate cancer from magnetic resonance spectroscopy (MRS)," in *MICCAI*, **4792**, pp. 278–286, 2007.
5. B. Chiu, G. H. Freeman, M. M. A. Salama, and A. Fenster, "Prostate segmentation algorithm using dyadic wavelet transform and discrete dynamic contour," *Phys Med Biol* **49**, pp. 4943–4960, Nov 2004.
6. 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* **27**, pp. 1777–1788, Aug 2000.
7. S. D. Pathak, V. Chalana, D. R. Haynor, and Y. Kim, "Prostate segmentation algorithm using dyadic wavelet transform and discrete dynamic contour," *IEEE Trans Med Imag* **19**, pp. 1211–1219, Dec. 2000.
8. L. Gong, S. D. Pathak, D. R. Haynor, P. S. Cho, and Y. Kim, "Parametric shape modeling using deformable superellipses for prostate segmentation.," *IEEE Trans Med Imaging* **23**, pp. 340–349, Mar 2004.
9. S. Klein, U. van der Heidi, B. Raaymakers, A. Kotte, M. Staring, and J. Pluim, "Segmentation of the prostate in mr images by atlas matching," *Biomedical Imaging: From Nano to Macro*, pp. 1300–1303, Apr 2007.
10. J. Costa, H. Delingette, S. Novellas, and N. Ayache, "Automatic segmentation of bladder and prostate using coupled 3d deformable models.," in *MICCAI*, 2007.
11. R. Zwigelaar, Y. Zhu, and S. Williams, *Pattern Recognition and Image Analysis*, ch. Semi-automatic Segmentation of the Prostate, pp. 1108–1116. 2003.
12. T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding* **61**, pp. 38–59, Jan 1995.
13. T. Cootes and C. Taylor, "Using grey-level models to improve active shape model search," *Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on 1*, pp. 63–67, Oct 1994.
14. T. Cootes and C. Taylor, "Statistical models of appearance for computer vision." March 2004.
15. Y. Zhu, R. Zwigelaar, and S. Williams, "Prostate segmentation: a comparative study," in *Medical Image Understanding and Analysis*, pp. 129–132, 2003.
16. B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "Active shape model segmentation with optimal features.," *IEEE Trans Med Imaging* **21**, pp. 924–933, Aug 2002.
17. M. Brejl and M. Sonka, "Object localization and border detection criteria design in edge-based image segmentation: automated learning from examples.," *IEEE Trans Med Imaging* **19**, pp. 973–985, Oct 2000.
18. D. Seghers, D. Loeckx, F. Maes, D. Vandermeulen, and P. Suetens, "Minimal shape and intensity cost path segmentation.," *IEEE Trans Med Imaging* **26**, pp. 1115–1129, Aug 2007.
19. S. Yan, D. Xu, B. Zhang, and H.-J. Zhang, "Graph embedding: A general framework for dimensionality reduction," *cvpr* **2**, pp. 830–837, 2005.
20. J. Chappelow, A. Madabhushi, M. Rosen, J. Tomaszewski, and M. Feldman, "A combined feature ensemble based mutual information scheme for robust inter-modal, inter-protocol image registration," in *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pp. 644–647, Apr. 2007.
21. J. Astola and I. Virtanen, "Entropy correlation coefficient, a measure of statistical dependence for categorized data," in *Proc. Univ. Vaasa, Discussion Papers*, 1982.
22. T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," in *IPMI*, pp. 33–47, 1993.
23. A. Madabhushi and D. N. Metaxas, "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions," *IEEE Trans Med Imaging* **22**, pp. 155–170, Feb. 2005.