A magnetic resonance spectroscopy driven initialization scheme for active shape model based prostate segmentation

Robert Toth a, Pallavi Tiwari a, Mark Rosen b, Galen Reed c, John Kurhanewicz c, Arjun Kalyanpur d, Sona Pungavkar e, Anant Madabhushi a,⇑

a Rutgers, The State University of New Jersey, Department of Biomedical Engineering, Piscataway, NJ 08854, USA
b University of Pennsylvania, Department of Radiology, Philadelphia, PA 19104, USA
c University of California, San Francisco, CA, USA
d Teleradiology Solutions, Bangalore 560 048, India
e Dr. Balabhai Nanavati Hospital, Mumbai 400 056, India

ABSTRACT

Segmentation of the prostate boundary on clinical images is useful in a large number of applications including calculation of prostate volume pre- and post-treatment, to detect extra-capsular spread, and for creating patient-specific anatomical models. Manual segmentation of the prostate boundary is, however, time consuming and subject to inter- and intra-reader variability. T2-weighted (T2-w) magnetic resonance (MR) structural imaging (MRI) and MR spectroscopy (MRS) have recently emerged as promising modalities for detection of prostate cancer in vivo. MRS data consists of spectral signals measuring relative metabolic concentrations, and the metavoxels near the prostate have distinct spectral signals from metavoxels outside the prostate. Active Shape Models (ASM’s) have become very popular segmentation methods for biomedical imagery. However, ASMs require careful initialization and are extremely sensitive to model initialization. The primary contribution of this paper is a scheme to automatically initialize an ASM for prostate segmentation on endorectal in vivo multi-protocol MRI via automated identification of MR spectra that lie within the prostate. A replicated clustering scheme is employed to distinguish prostatic from extra-prostatic MR spectra in the midgland. The spatial locations of the prostate spectra so identified are used as the initial ROI for a 2D ASM. The midgland initializations are used to define a ROI that is then scaled in 3D to cover the base and apex of the prostate. A multi-feature ASM employing statistical texture features is then used to drive the edge detection instead of just image intensity information alone. Quantitative comparison with another recent ASM initialization method by Cosio showed that our scheme resulted in a superior average segmentation performance on a total of 388 2D MRI sections obtained from 32 3D endorectal in vivo patient studies. Initialization of a 2D ASM via our MRS-based clustering scheme resulted in an average overlap accuracy (true positive ratio) of 0.60, while the scheme of Cosio yielded a corresponding average accuracy of 0.56 over 388 2D MR image sections. During an ASM segmentation, using no initialization resulted in an overlap of 0.53, using the Cosio based methodology resulted in an overlap of 0.60, and using the MRS-based methodology resulted in an overlap of 0.67, with a paired Student’s t-test indicating statistical significance to a high degree for all results. We also show that the final ASM segmentation result is highly correlated (as high as 0.90) to the initialization scheme.

1. Introduction

Prostatic adenocarcinoma (CaP) is the second leading cause of cancer related deaths among men in the United States, with an estimated 186,000 new cases in 2008 (Source: American Cancer Society). The current standard for detection of CaP is transrectal ultrasound (TRUS) guided symmetrical needle biopsy, which has a high false negative rate associated with it (Catalona, 1991). Recently, multi-modal Magnetic Resonance (MR) Imaging (MRI) comprising both structural T2-weighted (T2-w) MRI (Madabhushi et al., 2005; Zhu et al., 2003) and MR Spectroscopy (MRS) (Kurhanewicz et al., 1996, 2002; Kumar et al., 2008; Vilanova and Barcelo, 2007; Hom et al., 2006; Tiwari et al., 2009; Zaider et al., 2000; Kim et al., 2003; Coakley et al., 2003) have emerged as promising modalities for early detection of CaP (Kumar et al., 2008; Vilanova and Barcelo, 2007). MRS measures the relative concentrations of
different biochemicals and metabolites in the prostate, and changes in relative concentrations of choline, creatine, and citrate are highly indicative of the presence of CaP. It is important to note that MRS acquisition has a lower resolution than MRI acquisition, and thus each MRS metavoxel (containing a spectral signal) is approximately 13 times the size of an MRI voxel (containing a single intensity value).

An example of a MR spectra signature associated with a T2-w MRI image is shown in Fig. 1. The spectra corresponding to three metavoxels within the prostate are shown in red, and three spectra corresponding to metavoxels outside the prostate are shown in blue. The top of the choline and creatine region is indicated by the left-most purple circle in each of (b)–(d), and the top of the citrate peak is denoted by the right-most purple circle in each of (b)–(d). The mean MRS spectra from the prostate spectra (red) and the extra-prostatic spectra (blue) are shown. (i) Each datapoint represents an MRS spectra (prostate in red, extra-prostatic in blue), in which principal component analysis was used to reduce the 256-dimensional spectra to 3-dimensional spectra for visualization purposes.

As of 2009, there are approximately 16 ongoing clinical trials in the US aiming to demonstrate the role of MR in a diagnostic, clinical setting. Recent literature suggests that the integration of MRI and MRS could potentially improve sensitivity and specificity for CaP detection (Hom et al., 2006). In fact, when combined with MRI, using MRS data could yield prostate cancer detection specificity and sensitivity values as high as 90% and 88% respectively (Testa et al., 2007). Recently, computer-aided diagnosis (CAD) schemes have emerged for automated CaP detection from prostate T2-w MRI (Madabhushi et al., 2005; Chan et al., 2003) and MRS (Tiwari et al., 2007, 2008, 2009). In Tiwari et al. (2009), we showed that spectral clustering of the MRS data could be used to distinguish between prostatic and extra-prostatic voxels with accuracies as high as 98%. This paper improves upon the methodology presented in Tiwari et al. (2009) to drive a segmentation scheme for the prostate capsule on T2-w MRI.

With the recent advancements of prostate MR imaging, several prostate segmentation schemes have been developed (Zhu et al., 2003; Chiu et al., 2004; Costa et al., 2007; Ladak et al., 2000; Hu et al., 2003; Pathak et al., 2000; Gong et al., 2004; Cosio, 2008;}

---

**Fig. 1.** (a) 2D section of T2-weighted MR image with the prostate boundary in green. Metavoxels within the prostate are shown in red, and metavoxels outside the prostate are shown in blue. The top of the choline and creatine region is indicated by the left-most purple circle in each of (b)–(d), and the top of the citrate peak is denoted by the right-most purple circle in each of (b)–(d). (b) The mean MRS spectra from the prostate spectra (red) and the extra-prostatic spectra (blue) are shown. (i) Each datapoint represents an MRS spectra (prostate in red, extra-prostatic in blue), in which principal component analysis was used to reduce the 256-dimensional spectra to 3-dimensional spectra for visualization purposes.
Gao et al., 2010). Segmentation of the prostate is useful for a number of tasks, including calculating the prostate volume pre- and post-treatment (Hoffelt et al., 2003; Kaminski et al., 2002), for creating patient specific anatomical models (Nathan et al., 1996), and for planning surgeries by helping to determine just how far outside the capsule they might need to go in order to capture any possible extra-capsular spread of the tumor. Additionally, identifying the prostate capsule is clinically significant for determining whether extra-capsular spread of CaP has occurred. Manual segmentation of the prostate, however, is not only laborious, but is also subject to a high degree of inter-, and intra-observer variability (Warfield et al., 2002, 2004). The aim of this work is to automatically identify the spectra within the prostate in order to initialize a multi-feature active shape model (ASM) for precise segmentation of the prostate capsule.

2. Previous work and motivation

Previous work on automatic or semi-automatic prostate segmentation has been primarily for transrectal ultrasound (TRUS) images (Chiu et al., 2004; Ladak et al., 2000; Hu et al., 2003; Pathak et al., 2000; Gong et al., 2004; Cosio, 2008). Ladak et al. (2000) and Hu et al. (2003) presented semi-automated schemes in which several points on the prostate contour are manually selected to initialize a deformable model for prostate segmentation. Manual intervention is then used to guide the segmentation. Pathak et al. (2000) similarly presented an algorithm to detect prostate edges which were then used as a guide for manual delineation on prostate ultrasound images. In Gong et al. (2004), deformable ellipses were used as the shape model to segment the prostate from TRUS images. Recently, some researchers have attempted to develop prostate segmentation methods from in vivo endorectal prostate MR imagery (Costa et al., 2007; Gong et al., 2004; Flores-Tapia et al., 2008; Makni et al., 2008; Liu et al., 2009; Betrouni et al., 2008; Klein et al., 2008; Zwiggelaar et al., 2003; Gao et al., 2010). Klein et al. (2008) presented a segmentation algorithm in which the prostate boundary is obtained by averaging the boundary of a set of training images which are best registered to a test image. Costa et al. (2007) presented a 3D method for segmenting the prostate and bladder simultaneously to account for inter-patient variability in prostate appearance. In Zwiggelaar et al. (2003), polar transformations were used in conjunction with edge detection techniques such as non-maxima suppression to segment the prostate. A recent paper by Gao et al. (2010) uses a more advanced registration task to first align the prostate shapes, and hence creates a more accurate statistical shape model. The appearance of the prostate is learned using both the distance to the center of the prostate as well as the voxels' intensity information in a Bayesian framework, which showed very promising results on a number of studies. One possible limitation with some of these prostate segmentation schemes (Costa et al., 2007; Gong et al., 2004; Klein et al., 2008; Zwiggelaar et al., 2003) is in their inability to deal with variations in prostate size, shape, and across different patient studies. Further, some of these methods are often susceptible to MR image intensity artifacts including bias field inhomogeneity and image intensity non-standardness (Madabhushi et al., 2005, 2006; Madabhushi and Udupa, 2006).

A popular segmentation method is the active shape model (ASM), a statistical scheme that uses a series of manually land-marked training images to generate a point distribution model (Cootes et al., 1995). Principal component analysis (PCA) is then performed on this point distribution model to generate a statistical shape model (Cootes et al., 1995). A texture model is created near each landmark point of each training image, and the Mahalanobis distance between the test image and the training model is minimized to identify the boundary of an object (Cootes et al., 1995; Cootes and Taylor, 1994, 2004). The Mahalanobis distance is a statistical distance measurement and routinely used in conjunction with ASMs to estimate dissimilarity between training and test image intensities at user selected landmark points. After finding boundary points by minimizing the Mahalanobis distance, the shape model is deformed to best fit the boundary points, and the process is repeated until convergence.

While ASMs have been employed for a variety of different segmentation tasks related to biomedical imagery (Zhu et al., 2003, 2005; Cootes et al., 1993; Smyth et al., 1997; Montagnat and Deligette, 1997; de Bruijne et al., 2003; Mitchell et al., 2002), they have a major shortcoming in that they tend to be very sensitive to model initialization (Wang et al., 2002) and sometimes fail to converge to the desired edge. Manual initialization can be tedious and subject to operator variability. Over the last few years, some researchers have been exploring schemes for accurate and reproducible initialization of ASM's (Cosio, 2008; Brejl et al., 2000; Cootes et al., 1994). Seghers et al. (2007) presented a segmentation scheme where the entire image is searched for landmarks. They however concede that accurately initialized regions of interest (ROIs) would greatly improve their algorithm's efficiency and accuracy. Gao et al. (2010) use a single user click to initialize their prostate segmentation scheme, and would undoubtedly benefit from an automatic method for localizing the prostate in MR imagery. van Ginneken et al. (2002) pointed out that without an a priori spatial knowledge of the ROI, very computationally expensive searches would be required for ASM initialization, contributing to a slow overall convergence time. Multi-resolution ASMs have also been proposed, wherein the model searches for the ROI in the entire scene at progressively higher image resolutions (Cootes et al., 1994). Brejl et al. (2000) presented a shape-variant Hough transform to initialize an ASM, but the scheme can be very computationally expensive. Cosio (2008) presented an ASM initialization method based on pixel classification which was applied to segmenting TRUS prostate imagery. The method employs a Bayesian classifier to discriminate between prostate and non-prostate pixels in ultrasound imagery. A trained prostate shape is then fit to the edge of the prostate, identified via the Bayesian classifier. A Genetic Algorithm (Mitchel, 1998) is employed to minimize the distance between the trained prostate shape and the edge of the prostate. The second major shortcoming of ASM's is their inability to converge to the desired object boundary in the case of weak image gradients. The appearance model normally uses the intensities of the image to learn a statistical appearance model. However, there have been several studies in which using statistical texture features and more advanced appearance models have shown to yield improved accuracy over just using image intensities (Zhu et al., 2005; Seghers et al., 2007; van Ginneken et al., 2002; Toth et al., 2008, 2009). In addition, since the ASM models the object border using a multi-dimensional Gaussian, a large number of training images are required for accurate model generation (Ledoit and Wolf, 2004).

In this paper we present a novel, fully automated ASM initialization scheme for segmentation of the prostate on multi-protocol in vivo MR imagery by exploiting the MR spectral data. Note that for the studies considered in this work, the MRS data was acquired as part of routine multi-protocol prostate MR imaging and not specifically for the purposes of this project. While the resolution of MR and MRS data are different, the identification of prostate spectra by eliminating non-informative spectra outside the prostate provides an initial accurate ROI for the prostate ASM. We leverage the idea first introduced in Tiwari et al. (2008, 2009), in which spectral clustering was employed to distinguish between prostatic and extra-prostatic spectra. We achieve this through replicated k-means clustering of the MR spectra in the midgland. Replicated k-means
clustering aims to overcome limitations associated with the traditional k-means algorithm (sensitivity to choice of initial cluster centers). For each slice, the largest cluster (identified as the non-informative cluster) is eliminated. The mean shape of the prostate is then transformed to fit inside the remaining spectra, which serves to provide the initial landmark points for a 2D ASM. In addition, since the spectral data severely degrades away from the mid-gland of the prostate, we limit the clustering to the midgland. The resulting initialized shape is then rescaled to account for the change in size of the gland towards the base and the apex.

The ASM is performed in 2D because our limited number of 3D studies (32 in total) would prevent accurate statistical models from being generated in 3D. However, while we only had access to 32 3D studies, this constituted a total of 388 2D image slices. According to Ledoit and Wolf (2004), the ratio of the number of Gaussian dimensions (in our case 5) to the number of samples (in our case either 388 or 32) must be significantly less than 1. We therefore reasoned that while 32 3D samples would be insufficient to create an accurate 3D appearance model, the 388 slices would suffice to generate an accurate 2D statistical appearance model. In addition, we have opted to use multiple statistical texture features to better detect the prostate border. Statistical texture features have been shown to improve ASM accuracy (Seghers et al., 2007; van Ginneken et al., 2002; Toth et al., 2008, 2009). We calculate gray level statistics (such as mean and variance) of the neighborhood surrounding each landmark on the prostate border and use these statistics to generate our appearance model.

Note that as in traditional ASM schemes, the off-line training phase needs to only be done once. In this paper we compare the ASM segmentation performance using our MRS-based initialization scheme against corresponding results obtained via the initialization method recently presented by Cosio (2008). In Cosio (2008), only four patients were evaluated, with a total of 22 TRUS image slices, resulting in a mean MAD value of 1.65 mm. Note that while the Cosio method was originally presented for US data, in this paper we evaluate it in the context of prostate MR imagery. The popular Cootes et al. (1995) ASM was employed with both initialization schemes. Both schemes were rigorously evaluated via a 5-fold randomized cross validation system, in which the manual delineations of the prostate by an expert radiologist were used as the surrogate for ground truth. We also rigorously evaluated the results of the replicated k-means clustering algorithm with the results from two other popular clustering schemes – hierarchical and mean-shift clustering.

The rest of the paper is organized as follows. In Section 3 we provide a brief overview of our MRS-based ASM initialization scheme. In Section 4 we present the details of our methodology for identifying the prostate ROI via spectral clustering of MRS data. Section 5 describes the methodology for our ASM segmentation system. Section 6 describes the experimental setup and evaluation methods, in addition to our implementation of the ASM initialization method by Cosio (2008). In Section 7 we present both the qualitative and quantitative results, followed by a brief discussion of our findings. Finally, concluding remarks and future directions are presented in Section 8.

3. System overview

3.1. Notation

We define a spectral scene \( \tilde{C} = (\tilde{C}, \tilde{F}) \) where \( \tilde{C} \) is a 2D grid of metavoxels. Note that a metavoxel is a voxel at the lower spectral resolution. For each spatial location \( \tilde{c} \in \tilde{C} \), there is an associated 256-dimensional valued spectral vector \( \tilde{F}(\tilde{c}) = [f_j(\tilde{c})]_{j \in \{1, \ldots, 256\}} \), where \( f_j(\tilde{c}) \) represents the concentration of different biochemicals (such as creatine, citrate, and choline). We define the associated T2-w MR intensity image scene \( C = (C, f) \), where \( C \) represents a set of spatial locations (voxels), \( f(c) \) is the MR image intensity function associated with every \( c \in C \), and \( c = (x, y, z) \). The distance between any two adjacent meta-voxels, \( \tilde{c}, \tilde{d} \in \tilde{C}, \|\tilde{c} - \tilde{d}\|_2 \) (where \( \|\cdot\|_2 \) denotes the \( L_2 \) norm) is roughly 13 times the distance between any two adjacent spatial voxels \( c, d \in C \). We define \( X = \{c_1, \ldots, c_N\} \subset C \) as the set of \( N \) landmarks used to define a given prostate shape. The mean landmark coordinates across all training images is given as \( \bar{X} = \{c_1, \ldots, c_N\} \). The \( \kappa \)-neighborhood of pixels surrounding each \( c \in C \) is denoted as \( N\kappa(c) \), where for \( \forall d \in N\kappa(c) \), \( \|d - c\|_2 \leq \kappa, c \notin N\kappa(c) \). Finally, we denote as \( \Lambda \in (MRS, Cos) \) the MRS-based ASM presented in this work and the Cosio method (Cosio, 2008) respectively. A table of commonly used notation and symbols employed in this paper is shown in Table 1.

3.2. Data description

Our data comprises 32 multi-protocol clinical prostate MR datasets including both MRI and MRS endorectal in vivo data. These were collected during the American College of Radiology Imaging Network (ACRIN) multi-site trial (Mr imaging, xxxx) and from the University of California, San Fransisco. The MRS and MRI studies were obtained on 1.5 Tesla MRI scanners, and the MRI studies were axial T2-w images. The 32 3D studies comprised a total of 388 2D slices, with a spatial XY resolution of 256 × 256 pixels, or 140 × 140 mm. The ground truth for the prostate boundary on the T2-w images was obtained by manual outlining of the prostate border on each 2D section by a solitary expert radiologist, one with over 10 years of experience in prostate MR imagery.

3.3. Brief outline of ASM initialization schemes

Fig. 2 illustrates the modules and the pathways comprising our automated initialization system. First, replicated k-means clustering is performed on the spectra in the midgland, identified as the region of interest. Since the spectral data severely degrades away from the mid-gland towards the base and the apex, the resulting initialized shape is then rescaled to account for the change in size of the gland towards the base and the apex.

Table 1

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c )</td>
<td>MRI scene</td>
<td>( \tilde{c} )</td>
<td>MRS scene</td>
</tr>
<tr>
<td>( C )</td>
<td>Set of spatial coordinates in ( c )</td>
<td>( \tilde{C} )</td>
<td>Set of metavoxel coordinates in ( \tilde{C} )</td>
</tr>
<tr>
<td>( \tilde{c} )</td>
<td>A spatial location in ( \tilde{C} )</td>
<td>( \tilde{e} )</td>
<td>A metavoxel in ( \tilde{C} )</td>
</tr>
<tr>
<td>( f(\tilde{c}) )</td>
<td>Intensity value at ( \tilde{c} )</td>
<td>( \bar{F}(\tilde{C}) )</td>
<td>Spectral content at ( \tilde{C} )</td>
</tr>
<tr>
<td>( S_\kappa \subset C )</td>
<td>Prostate spatial locations (from experts)</td>
<td>( \Lambda \subset (MRS, Cos) )</td>
<td>Initialization method employed</td>
</tr>
<tr>
<td>( X )</td>
<td>Set of landmark points ((X \subset C))</td>
<td>( X^{\text{final}} )</td>
<td>Final segmented landmarks</td>
</tr>
<tr>
<td>( X^i )</td>
<td>Initialized landmarks</td>
<td>( X^i )</td>
<td></td>
</tr>
</tbody>
</table>
### 4. Methodology for MRS-based ASM initialization scheme

#### 4.1. Clustering of spectra (calculation of $S_{\text{MRS}}$)

The crux of the methodology is to determine a set of prostate voxels ($S_{\text{MRS}}$) based on a clustering of the spectroscopic data. This algorithm is described in the form of a sequence of steps below.

1. For a given 2D MRS slice $\hat{C} = (\hat{C}, \hat{F})$, we first obtain the MR spectra

$$\tilde{F}(\hat{c}) = \left[\tilde{f}(\hat{c}) | j \in \{1, \ldots, 256\}\right].$$

2. The metavoxels $\hat{c} \in \hat{C}$, are aggregated into $k$ clusters $V_a \subset \hat{C}$, $a \in \{1, \ldots, k\}$, by applying k-means clustering to all $\tilde{F}(\hat{c})$, $\forall \hat{c} \in \hat{C}$. k-means clustering aims to minimize the sum of distances to the clusters' centroids, for all clusters. Formally, it iteratively estimates

$$\arg\min_{V_1 \ldots V_k} \sum_{a=1}^{k} \sum_{c \in V_a} \left| \tilde{F}(\hat{c}) - \frac{1}{|V_a|} \sum_{c \in V_a} \tilde{F}(\hat{c}) \right|^2.$$  \hspace{1cm} (1)

3. Since the k-means algorithm is dependent on the starting locations of the centroids (i.e. which $V_a$ each $\hat{c}$ initially belongs to), the result is sometimes a local minima instead of a global minima. To overcome this limitation, the clustering was repeated 25 times with random initial locations of the centroids, and the resulting clustering which yields the minimum value from Eq. (1) is selected. Repeating the clustering more than 25 times did not significantly change the results.

4. The dominant cluster is identified as being extra-prostatic (non-informative), and the metavoxels in this cluster are removed. This is akin to the approach used in Tiwari et al. (2009), in which it was found that the dominant cluster is the non-informative cluster. The set of remaining metavoxels is then defined as,

$$\hat{S}_{\text{MRS}} = \left\{ \hat{c} | \hat{c} \in V_a, \ a \neq \arg\max_a |V_a| \right\}.$$  \hspace{1cm} (2)

The set of MRI voxels corresponding to metavoxels in $\hat{S}_{\text{MRS}}$ are then identified. For our data, we found that $k = 3$ clusters yielded the best results. Fig. 3a shows the $V_1, V_2, V_3$ as colored metavoxels. The largest cluster is shown in green, and would be eliminated, yielding $S_{\text{MRS}}$ as the cluster comprising blue and red metavoxels. While $S_{\text{MRS}}$ denotes the set of metavoxels, the voxels associated with $S_{\text{MRS}}$ are denoted as $S_{\text{MRS}}$ and are shown in red in Fig. 3b.

#### 4.2. Fitting the prostate shape (calculation of $X^0$)

Cosio (2008) employed the Genetic Algorithm to optimize the pose parameters of the prostate shape to fit a given binary mask. We found that using the objective function described below yielded an accurate initialization for a given set of prostate pixels $S_{\text{MRS}}$. The mean shape $X$ constitutes a polygon, and the set of voxels inside this polygon is denoted as $S_p$. More generally, for a given set of affine transformations $T$, which represent scaling rotation and translation, the set of voxels within that polygon are denoted as $S_{\text{MRS}}$. The objective function we use aims to maximize the true positive ratio, so that the initialization is given as

$$X^0_i = T(\mathbf{X}), \quad \text{ where } T = \arg\max_{\mathbf{X}} \left[ \frac{S_p \cap S_{\text{MRS}}}{S_p \cup S_{\text{MRS}}} \right].$$  \hspace{1cm} (3)

where $A \in \{\text{MRS, Cos}\}$. The optimization of Eq. (3) is performed via the Genetic Algorithm (Mitechel, 1998). The entire initialization process for $A = \text{MRS}$ is shown in Fig. 3.

#### 4.3. Estimation of $X^0_{\text{MRS}}$ in the base and apex

The MRS spectra lose their fidelity towards the base and the apex of the prostate. This is demonstrated in Fig. 4a, in which the three resulting clusters ($V_1, \ldots, V_3$) are shown via red, green, and blue metavoxels respectively. It was found that the spectra in the midgland of the prostate yielded accurate estimations of $S_{\text{MRS}}$. For this reason, we perform our clustering algorithm in the midgland.

---

3 For interpretation of color in ‘Figs. 1–10’, the reader is referred to the web version of this article.
of the prostate, the results of which are then extended to the base and apex. Note that since the 2D T2-w MRI slices tend to cover the prostate from base to apex, we have found that in our studies, the middle slice typically corresponded to the midgland. Hence we used the middle slice as our anchor midgland slice. We observed that the area of the prostate decreases to 80% its size in the base, and 30% its size in the apex. Fig. 4b demonstrates this tapering off of the gland in the base and apex. Fig. 4b is a histogram showing the size of the prostate relative to the central slice for all ground truth segmentations. Hence, to calculate $X_{\text{SMRS}}^i$ for the remaining slices, $X_{\text{SMRS}}^i$ is first calculated for slice 5, and is linearly scaled down to 80% its size for the first third of the 2D sections, and to 30% its size for the latter third of the gland.

5. Multi-feature ASM based segmentation

5.1. Basic shape model

Following model initialization ($X_0^0$), an ASM search (Cootes et al., 1995) is performed to segment the prostate from a new image. An ASM is defined by the equation

$$X = X_0 + P \cdot b,$$  

(4)

where $X$ represents the mean shape, $P$ is a matrix of the first few principal components (Eigenvectors) of the shape, created using Principal Component Analysis (PCA), and $b$ is a vector defining the shape, which can range from between $-3$ and $+3$ standard deviations from the mean shape. Therefore, $X$ is defined by the variable $b$. Given a set of landmark points $X_a$, ($a \in \{\text{MRS, Cos}\}$) for iteration $i$, the goal is to find landmark points $X_i$ closest to the object border. The shape is then updated using Eq. (4) where

$$b = P^t \cdot (X^t - X^0).$$

(5)

where each element of $b$ can only be within ±3 standard deviations of the mean shape. The final ASM segmentation is denoted as $X_{\text{out}}^t$. The training of the ASM to determine $X$ and $P$ is performed by manual delineation of the prostate boundary followed by manual alignment of 100 equally spaced landmark points along this boundary ($N = 100$). It should be noted that this needs to only be performed once in an off-line setting, and once an ASM is trained it can be used for repeated segmentations without significant manual intervention.

5.2. Basic appearance model

We define the set of intensity values near $c$ as $g(c) = \{f(d)|d \in N_c(e)\}$. Over all training images, the mean intensity values for a given landmark point $c_n, n \in \{1, \ldots, N\}$, are denoted as $g_n$ with the covariance matrix denoted as $\Sigma_n$. We define the set of pixels near $c_n \in X^i$ along the normal to the shape as $N(c_n)$. The standard cost function for a given pixel to the training set is the Mahalanobis distance. Therefore, $X^i$ is defined as

$$X^i = \{d_n|n \in \{1, \ldots, N\}\}, \text{ where }$$

$$d_n = \arg \min_{e \in N(c_n)} (g(e) - g_n)^t \cdot \Sigma_n^{-1} \cdot (g(e) - g_n).$$  

(6)

5.3. Multi-feature appearance model

Our previous work and that of others in employing multiple image features to drive the ASM model has shown that multi-feature ASMs are more likely to latch on to weak edges and boundaries compared to the traditional intensity driven ASM (Seghers et al., 2007; van Ginneken et al., 2002; Toth et al., 2008, 2009). In this work, in addition to using image intensity values $g(c)$, we also extracted the mean, standard deviation, range, skewness, and kurtosis of intensity values with local neighborhoods associated with every $c \in C$. If $E$ represents the expected value of an image feature, then the feature vector $(G(c))$ associated with each $c \in C$ is defined as,

$$G(c) = \left\{ g(c), E[|g(c) - \bar{g}(c)|^4]^{1/2}, \max g(c) \right\}$$

$$- \min g(c), \frac{E[|g(c) - \bar{g}(c)|^3]}{E[|g(c) - \bar{g}(c)|^2]^{3/2}},$$

$$\frac{E[|g(c) - \bar{g}(c)|^4]}{E[|g(c) - \bar{g}(c)|^2]^{2}}\right\}. $$

(7)

Our cost function in Eq. (6) thus uses $G$ instead of $g$. 

![Fig. 3. (a) MRS metavoxels $C$ (shown as colored boxes) overlaid onto the T2-w image $c$ where each color represents a different class resulting from the replicated k-means clustering scheme. The voxels associated with the informative metavoxels $S_{\text{MRS}}$ are shown as a red overlay in (b), with the resulting shape initialization $X_{\text{SMRS}}$ shown as a green line in (b) and (c).](image-url)

![Fig. 4. (a) Example of a clustering result in the base of the prostate, in which the three colors represent $V_a$, $a \in \{1, 2, 3\}$. The results demonstrate the degradation of quality of MR spectra in the prostate near the base, a phenomenon which also occurs near the apex. (b) A 2D histogram of the relative size of the prostate as a function of the slice index from base to apex reveals that the prostate is largest in the center and tapers off towards the extrema.](image-url)
6. Experiments and performance measures

6.1. Performance measures

For each image, a single expert radiologist segmented the prostate region, yielding ground truth landmarks \( X_f \). For a given shape \( X \), the set of pixels contained within the shape is denoted as \( S_X \). We employ the performance measures shown in Table 2.

Performance measures 1–4 are area based performance measures, in which a higher value indicates a more accurate segmentation, while performance measures 5 and 6 are edge based performance measures which evaluate proximity of the ASM extracted boundary compared to the manually delineated boundary. A lower value in measures 5 and 6 reflects a more accurate segmentation.

6.2. Comparison of clustering algorithms

We compared the efficacy of the replicated k-means clustering scheme with hierarchical clustering (Hastie et al., 2009) and mean-shift clustering (Comaniciu and Meer, 2002). Hierarchical clustering generates a dendrogram based off the Euclidean distance between spectra, and hierarchically combines spectra which have a low distance between them into a single cluster. The process is repeated until a pre-specified number of clusters remain (Hastie et al., 2009). Mean-shift clustering attempts to iteratively learn the density of the feature space and yields a clustering result based on the manifold instead of a pre-specified number of clusters (Comaniciu and Meer, 2002). Each methodology was used to calculate \( S_{MB} \) in the midgland, and these estimations of \( X_{MB} \) in the midgland were compared to the ground truth segmentations \( X_f \) for all studies. Table 3 summarizes the clustering experiments performed.

6.3. Comparison of initialization methods

6.3.1. Cosio based initialization (calculation of \( X_{MB}^{\text{Cosio}} \))

Cosio developed an automated ASM initialization scheme (Cosio, 2008) to segment prostate ultrasound images. We extended the Cosio (2008) scheme to MR imagery in order to compare this scheme against our MRS-based initialization method.4 The crux of the methodology is to classify the pixels in the images based on a Bayesian classification scheme using a mixture of Gaussians to represent the distribution of prostate and non-prostate pixels.

1. A set of 10 images was expertly segmented, resulting in a set of voxels within the prostate, denoted as \( S_c \subseteq C \) and a set of pixels outside the prostate, denoted as \( S_c = C - S_c \).
2. Each voxel \( c = (x_c, y_c) \) had a three dimensional vector constructed from its x and y coordinates and its intensity value \( F_{xy}(c) = [x_c, y_c, f(c)] \).

3. The distribution for prostate pixels \( (D_p) \) and non-prostate pixels \( (D_n) \) based on the spatial-intensity vector \( F_{xy}(c) = [x_c, y_c, f(c)] \) was then estimated by using the Expectation–Maximization function to generate a mixture of 10 Gaussians. The distribution is given as \( D = \sum_{b=1}^{10} N (\mu_b, \sigma_b^2) \), where \( N \) is a 3-D normal distribution such that \( \mu_b \in R^3 \) and \( \sigma_b^2 \in R^{3x3} \).

4. For each of the images considered, Bayes’ law (Duda et al., 2001) was then used to determine the posterior conditional probabilities of a given voxel \( c \in C \) belonging to the prostate class \( (P(w|c, D_p)) \) and the non-prostate class \( (P(w|c, D_n)) \), where \( w \) denotes the class. The set of pixels \( S_{MB} \) represents those pixels in which the likelihood of belonging to the prostate class was greater than the likelihood of being part of the non-prostrate class, so that

\[
S_{MB} = \{c | \ln[P(w|c, D_p)] + \ln[P(w|c)]) > \ln[P(w|c, D_n)] + \ln[P(w|c)]\}.
\]

The methodology presented in Section 4.2 was then used to calculate \( X_{MB}^{\text{Cosio}} \).

6.3.2. Comparing ASM segmentations using different initialization schemes

To test the ASM, each initialization method \((A \in \{MRS, Cosio\})\) resulted in an initialization \( X_A \) for each image, and a final ASM segmentation \( X_{MB}^{\text{final}} \). To evaluate ASM performance in the absence of any automated initialization, we placed the mean prostate shape in the center of the image yielding a third initialization, \( A = \text{Mid} \). A 5-fold cross validation was performed as follows. The 32 3D studies were randomly split into five groups of about six studies. For each group of studies, the remaining 26 were used for training the ASM, and the segmentation was performed for each slice in each of the six studies. This was repeated for each group resulting in segmentations for all 388 images. The segmentations were compared to the ground truth \( X_f \) in terms of the six performance measures in Table 2. A paired Student’s t-test was then carried out for each of the performance measures to determine the level of statistical significance. The t-test was performed over all 388 images (or, stated differently, with 387 degrees of freedom). Finally, the correlation (\( R^2 \) value) between each final segmentation result \( X_{MB}^{\text{final}}, A \in \{\text{MRS, Cosio, Mid}\} \) and the initialization result was calculated for each performance measure, to determine the correlation between the final segmentation with respect to the initialization. A high \( R^2 \) value would indicate that the model is very sensitive to the initialization.

6.4. Comparison of multi-resolution ASM’s

For each of the initialization methods \((A \in \{MRS, Cosio, Mid\})\), a multi-resolution ASM was also performed. A Gaussian image pyramid was constructed with 4 image resolution levels (ranging from \( 32 \times 32 \) pixels to \( 256 \times 256 \) pixels), and at each image level a distinct ASM model is constructed. The result from the final pyramid (the full-resolution image) was used as \( X_{MB}^{\text{final}} \), and 5-fold cross validation and statistical significance tests were performed as described above. Finally, the correlation (\( R^2 \) value) was also calculated, the hypothesis being that if the multi-resolution

### Table 2

<table>
<thead>
<tr>
<th>Measure</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap</td>
<td>(</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>(</td>
</tr>
<tr>
<td>Specificity</td>
<td>(</td>
</tr>
<tr>
<td>Positive predictive value (PPV)</td>
<td>(</td>
</tr>
<tr>
<td>Mean absolute distance (MAD)</td>
<td>( \sum_{i=1}^{n} ((c_i - d_i) \in X_f, d_i \in X_c) )</td>
</tr>
<tr>
<td>Hausdorff distance (HAD)</td>
<td>max_{i=1}^{n} ({c_i - d_i} \in X_f, d_i \in X_c)</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon_1 )</td>
<td>Hierarchical clustering</td>
</tr>
<tr>
<td>( \varepsilon_2 )</td>
<td>Mean-shift clustering</td>
</tr>
<tr>
<td>( \varepsilon_3 )</td>
<td>Replicated k-means clustering</td>
</tr>
</tbody>
</table>
framework was able to overcome poor or lack of initialization, then the $R^2$ values would be lower for the multi-resolution experiments. In essence, we were aiming to explore whether a multi-resolution approach would make redundant the need for initialization. The entire set of ASM experiments performed are summarized in Table 4.

7. Results and discussion

7.1. Comparison of clustering algorithms ($E_1 - E_3$)

Fig. 5 shows some qualitative results from $E_1 - E_3$ for a mid-gland slice, which was used to calculate $S_{MRS}$. In all the images,
the red cluster indicates the largest cluster which is removed, so that SMRS corresponds to all metavoxels that are not red. While all the clustering methodologies found a set of metavoxels within the prostate, it is obvious that only the replicated $k$-means clustering algorithm was able to properly identify most of the prostate metavoxels. The superior performance of replicated clustering over

![Graph](image)

**Fig. 7.** Quantitative results for $E_1/C_0E_3$ are shown for the six performance measures over 32 midgland slices, in which the blue bar represents $E_1$, the red represents $E_2$, and the yellow represents $E_3$. The standard deviations over 32 studies are shown with black bars.

![Graph](image)

**Fig. 8.** (a) Results from the initialization schemes comparing $X_0^D$ to $X_0^E$ for each slice, in which $E_4$ is shown in blue, $E_7$ is shown in red, and $E_{10}$ is shown in yellow. The statistical significance results are shown in Table 5. (b) Results from the ASM's comparing $X_{final}^D$ to $X_0^E$ for each slice, in which $E_5$ is shown in blue, $E_8$ is shown in red, and $E_{11}$ is shown in yellow. The statistical significance results are shown in Table 6. Standard deviations over 388 slices are shown as black bars in both (a) and (b).
the two other algorithms might have to do with the fact that we were unable to identify the optimal parameter values for the mean-shift and hierarchical clustering schemes. Although we experimented with several different parameter settings, the hierarchical and mean-shift clustering algorithms ended up clustering most of the background metaxoxels with the prostate metaxoxels. The replicated k-means clustering, however, was able to group most of the prostate metaxoxels together for a value of k = 3. Note that k = 3 appeared to work better than k = 2, owing perhaps to some heterogeneity of tissues (cancerous and benign areas) within the prostate.

For each of the 32 studies, \(X^0\) was calculated for the midgland slice, and compared to \(X_p\). The quantitative results for a single midgland slice are shown in Fig. 7 over all 32 studies. The most important thing to note is the extremely high sensitivity of the replicated k-means compared to the other clustering algorithms, suggesting that the mean-shift and hierarchical clustering tended to under-segment the prostate voxels. In addition, the overlap measure of the replicated k-means, which takes into account both false positive and false negative pixels, was higher than any of the other algorithms, suggesting a more accurate overall initialization.

### 7.2. Comparison of initialization schemes \((E_4, E_5, E_7, E_8, E_{10}, E_{11})\)

Fig. 6 shows \(S_4\) in red with \(X_i^0\) in white for \(\mathcal{A} \in \{\text{Cos, MRS}\}\) for two different studies, with \(X_p\) shown in the third column (\(C\)) and (\(I\)) for reference. It can be seen that both perform quite well in locating the prostate pixels. This was further tested quantitatively over all 388 slices, results of which are shown in Fig. 8. Fig. 8 illustrates that in many cases the MRS initialization yielded superior results. It should be noted that the specificity was extremely high in most cases due to the large size of the image compared to the size of the target object (\(|C| \gg |S_p|\)). Our model yielded an average overlap value of 0.66, while the Cosio method had a value of 0.62. Comparison of our segmentation system with other MR prostate segmentation schemes show that our system performs comparably to other related techniques. Zhu et al. (2003) have overlap coefficients ranging from about 0.15 to about 0.85, while our mean overlap is 0.66. Costa et al. (2007) used prior knowledge from the MRI scan to localize an ROI containing the prostate, resulting in a mean sensitivity and PPV values of 0.75 and 0.80 respectively for the prostate. Our model achieved mean sensitivity and PPV values of 0.81 and 0.79 respectively. Comparing the initialization schemes \((X_i, \mathcal{A} \in \{\text{MRS, Cos, Mid}\})\) revealed that all results were statistically significant to a high degree, except for the edge based performance measures between \(\mathcal{A} = \text{MRS}\) and \(\mathcal{A} = \text{Mid}\), shown in Table 5. When comparing the actual ASM segmentations, all results were statistically significant except for the Hausdorff distance between \(\mathcal{A} = \text{Cos}\) and \(\mathcal{A} = \text{Mid}\), shown in Table 6. In Table 8, it can be seen that the ASM is extremely sensitive to the initialization, with the lowest \(R^2\) value being 0.71.

### 7.3. Comparison of multi-resolution ASM’s \((E_6, E_9, E_{12})\)

Fig. 9 shows the result from the multi-resolution ASM’s over 388 images using a 5-fold cross validation scheme. These results were then compared to the values reported in Fig. 8b via the use of a statistical significance test, shown in Table 7. Apart from a few scenarios (sensitivity of \(\mathcal{A} = \text{MRS}\), specificity of \(A = \text{Cos}\), and overlap and edge performance measures of \(\mathcal{A} = \text{Mid}\)), the multi-resolution approach generally improved the ASM segmentation results significantly. Finally, the \(R^2\) values from the multi-resolution experiments were calculated, and they were all extre-
In this paper, we have presented a fully automated and accurate ASM initialization scheme for prostate segmentation from multi-protocol in vivo MRI/MRS data. With the increasing use of MR imaging of the prostate, several institutions are beginning to acquire multi-modal MR prostate data, including MR spectroscopy (Kurhanewicz et al., 1996; Kumar et al., 2008; Vilanova and Barcelo, 2007; Hom et al., 2006; Tiwari et al., 2009). The primary novel contribution of our work is in leveraging information from one imaging protocol (spectroscopy) to drive the segmentation of the prostate on a different protocol (T2-weighted structural MRI). To the best of our knowledge this is the first instance of multi-modal information being used in this fashion for ASM initialization. Our method uses replicated k-means clustering to cluster the MRS spectra in the midgland. For the studies we considered, the central slice was assumed to be the midgland. This may not always be the case, and future work will entail developing an automated, more intelligent scheme for selecting the midgland slice. We then eliminate the background spectra, fit the shape to the remaining spectra, and extend our initializations to the base and apex of the prostate. We employed a multi-feature ASM to perform the segmentation, wherein multiple statistical texture features were used to complement image intensities. We compared our MRS initialization method against a recent image feature driven ASM initialization method by Cosio (2008) and found that our scheme resulted in significantly better segmentations. Future work will entail evaluating our scheme on a larger cohort of data.

8. Concluding remarks

In this paper, we have presented a fully automated and accurate ASM initialization scheme for prostate segmentation from multi-protocol in vivo MRI/MRS data. With the increasing use of MR imaging of the prostate, several institutions are beginning to acquire multi-modal MR prostate data, including MR spectroscopy (Kurhanewicz et al., 1996; Kumar et al., 2008; Vilanova and Barcelo, 2007; Hom et al., 2006; Tiwari et al., 2009). The primary novel contribution of our work is in leveraging information from one imaging protocol (spectroscopy) to drive the segmentation of the prostate on a different protocol (T2-weighted structural MRI). To the best of our knowledge this is the first instance of multi-modal information being used in this fashion for ASM initialization. Our method uses replicated k-means clustering to cluster the MRS spectra in the midgland. For the studies we considered, the central slice was assumed to be the midgland. This may not always be the case, and future work will entail developing an automated, more intelligent scheme for selecting the midgland slice. We then eliminate the background spectra, fit the shape to the remaining spectra, and extend our initializations to the base and apex of the prostate. We employed a multi-feature ASM to perform the segmentation, wherein multiple statistical texture features were used to complement image intensities. We compared our MRS initialization method against a recent image feature driven ASM initialization method by Cosio (2008) and found that our scheme resulted in significantly better segmentations. Future work will entail evaluating our scheme on a larger cohort of data.

Acknowledgments

This work was made possible via grants from the Wallace H. Coulter Foundation, New Jersey Commission on Cancer Research, National Cancer Institute (Grant Nos. R01CA136535-01, ARRA-NCTI-3 R21CA127186, R21CA127186, R03CA128081-01, R01CA140772-01A1, and R03CA143991-01), the Cancer Institute of New Jersey, and the Life Science Commercialization Award from Rutgers University. The authors would like to thank Fernando Cosio in his assistance implementing his initialization methodology. Finally, the authors would like to acknowledge the ACRIN trial for providing the MRI/MRS data.

References


