



Published in final edited form as:

Comput Vis Image Underst. 2013 September 1; 117(9): 1051–1060. doi:10.1016/j.cviu.2012.11.013.

Simultaneous Segmentation of Prostatic Zones Using Active Appearance Models With Multiple Coupled Levelsets

Robert Toth^a, Justin Ribault^b, John Gentile^c, Dan Sperling^c, and Anant Madabhushi^d

^aDept. of Biomedical Engineering, Rutgers University, Piscataway, NJ, 08854

^bMacNeal Hospital, Berwyn, IL, 60402

^cNew Jersey Institute of Radiology, Carlstadt, NJ, 07072

^dDept. of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, 44120

Abstract

In this work we present an improvement to the popular Active Appearance Model (AAM) algorithm, that we call the Multiple-Levelset AAM (MLA). The MLA can simultaneously segment multiple objects, and makes use of multiple levelsets, rather than anatomical landmarks, to define the shapes. AAMs traditionally define the shape of each object using a set of anatomical landmarks. However, landmarks can be difficult to identify, and AAMs traditionally only allow for segmentation of a single object of interest. The MLA, which is a landmark independent AAM, allows for levelsets of multiple objects to be determined and allows for them to be coupled with image intensities. This gives the MLA the flexibility to simultaneously segment multiple objects of interest in a new image.

In this work we apply the MLA to segment the prostate capsule, the prostate peripheral zone (PZ), and the prostate central gland (CG), from a set of 40 endorectal, T2-weighted MRI images. The MLA system we employ in this work leverages a hierarchical segmentation framework, so constructed as to exploit domain specific attributes, by utilizing a given prostate segmentation to help drive the segmentations of the CG and PZ, which are embedded within the prostate. Our coupled MLA scheme yielded mean Dice accuracy values of .81, .79 and .68 for the prostate, CG, and PZ, respectively using a leave-one-out cross validation scheme over 40 patient studies. When only considering the midgland of the prostate, the mean *DSC* values were .89, .84, and .76 for the prostate, CG, and PZ respectively.

Keywords

Active Appearance Models; Prostate Segmentation; Levelsets

© 2013 Elsevier Inc. All rights reserved.

robtoth@gmail.com (Robert Toth), anant.madabhushi@case.edu (Anant Madabhushi) .

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

1.1. Background and Motivation

Statistical Shape Models (SSMs) use shape information to yield an accurate, shape constrained, segmentation of an object of interest, and are extremely popular in medical image segmentation [1, 2, 3, 4, 5, 6]. A common SSM methodology is the Active Appearance Model (AAM) segmentation algorithm [7]. AAMs attempt to learn both the appearance, and the shape, of an object of interest. In addition, AAMs aim to learn the relationship between the shape and appearance. When using an AAM to segment a new image, the appearance of the new image is matched to the AAM, and the associated shape yields a segmentation of the object of interest.

AAMs achieve this by performing principal component analysis (PCA) on a set of intensities defining the object of interest to yield a low dimensional intensity projection. The intensity projections are then concatenated with the shape information, and PCA is performed a second time. The eigenvectors resulting from the second “coupled” PCA define the linear relationship between shape and appearance. Thus a given coupled projection defines both shape and appearance, and can be used to reconstruct the high dimensional intensity and shape information. However, traditional AAMs define the shape by a set of Cartesian coordinates defined from the landmarks of a single object. Yet there are several issues with using landmarks to construct statistical shape models,

1. 1. A large number of anatomical landmarks are required to accurately capture shape variations.
2. 2. Anatomical landmarks must be accurately aligned and landmark correspondences must be established on all training images [8].
3. 3. Automated method for landmark detection and alignment can be prone to errors [9].
4. 4. Landmarks require triangulation, and the triangulation algorithm may be prone to errors [10].

Leventon et al. [11] first proposed performing PCA on a series of signed distance maps (levelsets) to capture shape variations, to overcome the issues with landmark based SSMs. A levelset is defined as a set of positive values at every pixel outside the object of interest, and a set of negative values at every pixel inside the object of interest. Therefore, a value of 0 would represent the surface of the object. The simplest way to compute a levelset is at each pixel in the image, calculate the Euclidean distance to the closest border pixel, and negate that value if the pixel values within the object of interest. To define multiple levelsets, the signed distance to the border of *each* object is computed.

Tsai et al. [12] concatenated the levelsets of multiple objects prior to performing PCA, essentially “coupling” the levelsets. Hence a single set of low dimensional values (a “projection”) is used to represent the shape of multiple objects. This allows for simultaneous segmentation of multiple objects. In addition, coupling the individual SSMs allows one to take advantage of the inherent dependency between the spatial location of multiple adjoining organs. The technique used by Tsai et al. [12] was in 2D, and the SSMs included not just shape, but also pose information. Akhondi-Asl et al. [13] developed a coupled SSM in 3D, which only accounted for shape variations (but not pose) by first aligning the training shapes. Akhondi-Asl et al. [13] then explored whether coupling the SSMs actually improved segmentation accuracy over simply constructing individual SSMs. It was found that in most cases, shape coupling improves results only when the levelsets were first aligned prior to training.

1.2. Novel Contributions

In this work, we present a new 3D AAM framework which is landmark-free, and which can segment multiple objects simultaneously in a new image via coupling multiple levelsets. Our model, the multiple levelset AAM (hereafter denoted as MLA) uses multiple coupled level sets to model the 3D shapes, thereby helping to alleviate many of the issues facing traditional landmark-based AAMs. The MLA offers the advantage of (1) not having to deal with the landmark identification problem, and (2) not having to triangulate a series of landmarks to generate a 3D model. In addition, we take a similar approach that proposed by Leventon et al. [11] and Akhondi-Asl et al. [13], in that multiple levelsets are coupled to allow for simultaneous segmentation of multiple objects.

Previous work in coupling multiple levelsets involved concatenating the high dimensional levelsets of multiple objects [12, 13]. Yet this can be computationally infeasible when dealing with (a) multiple levelsets, (b) a large number of training images, or (c) very large images. To overcome these issues, the MLA performs PCA on each shape, prior to coupling the levelsets, similar to how AAMs perform PCA on the intensities prior to coupling.

In addition, the MLA can also use existing segmentations of one or more organs to generate more accurate segmentations of the adjacent organs (for example using the prostate segmentation to simultaneously segment the bladder and rectum [14]). This is accomplished by generating the coupled projections using both intensities and levelsets, whereas prior AAM models are only able to consider intensity information [7].

This approach also allows the MLA to be used hierarchically, in which one object can first be segmented, and then used to drive the segmentations of other embedded objects. For example, in the case of prostate MRI, the central gland (CG) and peripheral zone (PZ) are substructures of the prostate itself (see Figure 1). Using an existing prostate segmentation to segment the CG and PZ reduces the search space, which can help hone in on the embedded substructures within the gland. In addition, the coupled model allows for structural linking of the adjoining sub-structures, thereby permitting incorporation of anatomic constraints.

1.3. Application to Prostatic Zone Segmentation

Our MLA is applied to the task of prostate segmentation from endorectal, 3 Tesla, T2-weighted (T2-w) MRI images. The prostate gland consists of internal structures including the peripheral zone (PZ), central zone (CZ), and transition zone (TZ), where the latter 2 structures are jointly referred to as the central gland (CG) [15] (see Figure 1). While most tumors are found in the PZ, tumors can also be found in the CG, and CG tumors can have drastically different appearance than PZ tumors [16, 17, 18]. In the PZ (where most cancers are found), tumors are typically characterized by hypointense regions on MRI images, in stark contrast to the usually hyperintense PZ regions. However, in the CG, tumors are typically noticeable due to their homogeneous texture, as compared to the traditionally heterogeneous texture in the CG. In recent years, several computer aided detection (CAD) systems have been developed for detecting tumors from prostate MRI imagery [19, 20, 21, 22, 23, 24, 25]. Since the tumors in the PZ can appear drastically different from tumors in the CG, CAD systems would invariably benefit from knowing where each internal prostate structure was located. In addition, treatment options can even be tailored to an individual patient, as CG tumors have been found to be significantly less aggressive compared to PZ tumors [26].

However, most extant prostate segmentation systems only consider the prostate capsule boundary [6, 27, 28, 29, 30, 31, 32]. Makni et al. [33] developed a system for distinguishing the internal prostate structures on multiparametric MRI, but assumed the prostate is already segmented. Additionally Makni et al. [33] required the use of both T2-w and DCE-MRI to drive the segmentation, as opposed to just a single protocol; imaging with additional

parameters (over and above standard T2-w) leads to an increased imaging time and hence cost of exam. Liu et al. [25] circumvented the issue of zonal segmentation within the prostate by proposing a spatially aware CAD system for cancer detection, to automatically identify in which part of the prostate the tumor was located.

In this work, we aim to use our MLA to automatically and simultaneously segment the prostate, PZ, and CG from T2-weighted MRI alone. As with Liu et al. [25], our ultimate goal is to develop a spatially aware CAD system for prostate cancer detection, but by leveraging the explicit, automated segmentations of the different prostate zones. A second application is to create patient-specific treatment models based on the zonal location of the tumor.

The rest of the paper is organized as follows. Section 2 describes the methodology for training the MLA and using the MLA to segment a new image. Section 3 describes the experimental design and dataset. Section 4 presents the results. A brief discussion is presented in Section 5, while Section 6 presents concluding remarks and future directions.

2. Methodology

2.1. Overview of MLA

The MLA is comprised of distinct training and segmentation steps, which are summarized as follows. To train the MLA, a series of levelsets is first calculated, where negative values represent pixels inside the object, and positive values represent pixels outside the object. PCA is first performed on a set of image intensities and levelsets from multiple training images, resulting in a set of low dimensional intensity and levelset projections [34]. Coupling of the levelsets is achieved by performing PCA a second time on a concatenation of these projections [34].

In the segmentation phase, the intensities of a new image are used to drive the segmentations of the objects of interest. In addition, the MLA framework allows for one or more pre-segmented objects to help drive the segmentation of other objects in the new image. The first step of the MLA is to affinely transform the input intensities and levelsets. These transformations will be used to align the new image with the MLA. After transforming the inputs, a set of low dimensional projections of the inputs (intensities and, optionally, levelsets) are calculated using the trained MLA. The matrix of coupled eigenvectors is then used to reconstruct the high dimensional set of transformed image intensities, and levelsets for all objects in the image. The transformation for which the reconstructions are best correlated with the inputs is then calculated. Given the optimal transformation, the high dimensional levelsets of all objects are reconstructed and thresholded, yielding a set of segmentations for all objects in the image.

Figure 2 shows the entire segmentation process on a new image. A comparison between the MLA with a traditional AAM and the coupled levelsets proposed by Tsai et al. [12] is shown in Figure 3. Traditional AAMs couple intensities with landmarks, and traditional coupled levelsets involve concatenating a set of levelsets, which can be computationally infeasible. The MLA, by comparison, allows for coupling of the projections of multiple levelsets with the intensity projections, essentially allowing simultaneous segmentations in an AAM framework. This helps to avoid the computational cost of concatenating a series of high dimensional levelsets.

2.2. Notation

An image scene is defined as $C = (C, f)$, where C represents a set of spatial locations, and each $c \in C$ represents a set of Cartesian coordinates $c = (x, y, z)$. The intensity at each $c \in C$

is denoted as $f(c)$. The number of segmented objects is denoted as M and the number of training images is denoted as N . The number of pixels in each image is denoted as $P = |C|$. A given training object is denoted as $C_n = (C_n, f_n)$. A given segmentation for object $m \in \{1, \dots, M\}$ for image $n \in \{1, \dots, N\}$ is denoted as $C_{m,n}^{(in)} \subset C_n$, which defines the set of pixels inside the object of interest. A new image to be segmented is denoted as $C_\theta = (C_\theta, f_\theta)$. A full list of notation and symbols used throughout the paper is presented in Table 1.

2.3. MLA Training

Step 1. Compute Levelsets—All training images are first aligned, as described in [13]. A single training image is chosen as the template, and an affine alignment is applied to align all the segmentations to the template. The next step is to generate a levelset from each training object, such that negative levelset values represent pixels inside the object, and positive levelset values represent pixels outside the object.

The levelset for object m in image n is denoted as $L_{m,n} = \{l(c) \mid c \in C_n\}$ where $L_{m,n} \in \mathbb{R}^P$. The levelset is represented by the signed distances to the object's surface [11], and is calculated as,

$$l(c) = \begin{cases} + \min_{d \in C_{m,n}^{(in)}} \|c - d\|_2 & \text{if } c \notin C_{m,n}^{(in)} \\ - \min_{d \notin C_{m,n}^{(in)}} \|c - d\|_2 & \text{if } c \in C_{m,n}^{(in)}. \end{cases} \quad (1)$$

Step 2. Perform PCA on Intensities and Levelsets—The set of training levelsets for object m is denoted as S_m , where $S_m = \{L_{m,1}, \dots, L_{m,N}\}$. PCA is performed on each S_m , $m \in \{1, \dots, M\}$ which results in a mean levelset $\mu_m \in \mathbb{R}^P$ and a matrix of eigenvectors $\psi_m \in \mathbb{R}^{P \times p_m}$, where $p_m \leq P$ and each column of ψ_m is an eigenvector [11]. The number of eigenvectors p_m is selected to retain at least a percentage α of the variance in S_m .

The set of intensities for training image n is denoted as $F_n = \{f_n(c) \mid c \in C_n\}$ where $F_n \in \mathbb{R}^P$. The entire set of training intensities is denoted as $S_F = \{F_1, \dots, F_N\}$. PCA is performed on S_F resulting in a set of mean intensities $\mu_F \in \mathbb{R}^P$ and a matrix of eigenvectors $\psi_F \in \mathbb{R}^{P \times p_F}$, where $p_F \leq P$ and each column of ψ_F is an eigenvector [7]. The number of eigenvectors p_F is selected to retain at least a percentage α of the variance in S_F .

Step 3. Coupling Intensities and Levelsets—Once PCA has been performed, each $F_n, L_{m,n} \forall m \in \{1, \dots, M\}, n \in \{1, \dots, N\}$ must be projected down into the lower (p) dimensional space. Given an object $X \in \mathbb{R}^P$, the optimal (least squares) projection $\widehat{X} \in \mathbb{R}^p$ is calculated as,

$$\widehat{X} = \psi^+ \cdot (X - \mu), \text{ where } \psi^+ = (\psi^\top \psi)^{-1} \cdot \psi^\top, \quad (2)$$

ψ^+ is the Moore-Penrose pseudoinverse [35], and \top is the transpose operation.

The projections for each object in each image, $\widehat{L}_{m,n}$, and the intensities, \widehat{F}_n are calculated using Equation (2). To create the coupled model, the set of levelset and intensity projections are concatenated for training image n as

$$\Omega_n = \{\widehat{L}_{1,n} \cup \dots \cup \widehat{L}_{M,n} \cup \widehat{F}_n\}, \quad (3)$$

where $\Omega_n \in \mathbb{R}^q$. PCA is performed a second time, resulting in a matrix of eigenvectors $\psi_\Omega \in \mathbb{R}^{q \times \tilde{p}}$, where $\tilde{p} \leq q$ and each column of Ψ_Ω is an eigenvector. The means of each projection are 0, and therefore $\mu_\Omega = 0$. Hence Ψ_Ω defines the linear relationship between intensity and levelset projections for M objects.

A graphical display of the coupling process is shown in Figure 4. The boxes in the left column represent the high dimensional intensities and levelsets. The boxes in the middle column represent the projections of each intensity and levelset. Finally, the boxes in the right column represent the coupled matrix Ψ_Ω , where each column represents a single eigenvector. Figure 4 also illustrates the fact that each row of Ψ_Ω corresponds precisely to a specific projection.

2.4. MLA Segmentation

Step 1. Transform Intensities and Levelsets—For a new study $C\theta$ to be segmented, it is assumed that at least the intensities $F\theta$ are given. In addition, it is possible that one or more segmentations (and therefore levelsets) are either known or have been previously calculated, which are denoted as $\{L_{k,\theta} \mid k \in K\}$, where $K \subset \{1, \dots, M\}$. If just the intensities are known, $K = \emptyset$. Given a transformation T , $T(c)$ denotes transforming the spatial location $c \in C\theta$. The set of transformed intensities is denoted as $F_T = \{f(T(c)) \mid c \in C\theta\}$ and levelsets as,

$$L_{k,T} = \{l_k(T(c)) \mid c \in C\theta\}, k \in K.$$

Step 2. Project Inputs—We first calculate the projections $\widehat{F}_T, \widehat{L}_{k,T}, k \in K$ using Equation (2). The projections $\widehat{F}_T, \widehat{L}_{k,T}, k \in K$ are concatenated as Ω_θ . This is shown graphically in Figure 5(a).

Step 3. Reconstruct Inputs—Given a projection $\widehat{X} \in \mathbb{R}^p$, a reconstruction $R(\widehat{X}) \in \mathbb{R}^p$ is calculated as,

$$R(\widehat{X}) = \mu + \psi \widehat{X}. \quad (4)$$

The rows from the coupled matrix Ψ_Ω corresponding to Ω_θ are extracted, denoted as $\tilde{\psi}_\Omega$. Then, a set of coupled projections $\widehat{\Omega}_\theta \in \mathbb{R}^{\tilde{p}}$ is calculated using Equation (2) with $X = \Omega_\theta$ and $\tilde{X} = \widehat{\Omega}_\theta$. Extracting only the rows corresponding to our known inputs allows us to estimate the q -dimensional coupled projection vector from our inputs (see Figure 5(b)).

A reconstruction of the entire set of projections (the intensities and all levelsets) is calculated as $R(\widehat{\Omega}_\theta) \in \mathbb{R}^q$ using Equation (4). The individual projections $\widehat{F}_T, \widehat{L}_{1,T}, \dots, \widehat{L}_{M,T}$ are extracted from $R(\widehat{\Omega}_\theta)$. Finally, the reconstructions $R(\widehat{F}_T), R(\widehat{L}_{1,T}), \dots, R(\widehat{L}_{M,T}) \in \mathbb{R}^p$ are calculated using Equation (4) (see Figure 5(c)).

Step 4. Optimize Transformation—For a given transformation T , the coupled reconstructions $R(\widehat{F}_T), R(\widehat{L}_{k,\theta}), k \in K$ are computed. When the reconstructions are most correlated with the inputs, the correct transformation T^* is presumed to be found.

Normalized cross correlation is used as the metric to define the reconstruction accuracy, where for an object X and its reconstruction $R(\widehat{X})$, the NCC is defined as,

$$\text{NCC}(X, R(\widehat{X})) = \left\langle \frac{X - \bar{X}}{\|X - \bar{X}\|_2}, \frac{R(\widehat{X}) - \bar{R}(\widehat{X})}{\|R(\widehat{X}) - \bar{R}(\widehat{X})\|_2} \right\rangle, \quad (5)$$

where \bar{X} represents the mean, $\|\cdot\|_2$ represents the L_2 norm, and $\langle \cdot, \cdot \rangle$ represents the inner product.

To determine the best transformation T^* , the NCC between the reconstructions and the original data is maximized as,

$$T^* = \underset{T}{\text{argmax}} \text{NCC}(F_T, R(\widehat{F}_T)) + \sum_{k \in K} \text{NCC}(L_{k,T}, R(\widehat{L}_{k,T})), \quad (6)$$

Equation 6 must be optimized to determine the set of affine parameters for which the NCC is maximized. This is a crucial step, since the working hypothesis is that the NCC will be maximized if and only if the MLA is properly aligned with the new image, and hence the reconstruction will properly capture the desired segmentations. A global optimization is first performed, followed by a local optimization to properly hone in on the maximum NCC. To perform the global optimization, an initialization-biased particle swarm optimizer [36] is used, in which 100 random affine parameters (100 particles) are each allowed to converge independently on the maximum NCC, and the particle with the maximum NCC overall is chosen. Then, a local Powell optimization [37] is performed, in which each of the 12 affine parameters is optimized independently. This process of optimizing each parameter independently is repeated until convergence, thus driving the transformations to the maximum NCC value.

Step 5. Calculate Segmentations—Given the optimal transform parameters T^* , the reconstructions of all levelsets are calculated as $R(\widehat{L}_{1,T^*}), \dots, R(\widehat{L}_{M,T^*})$ using Equation (4). A segmentation result for object m , denoted as $C_{m,\theta}^{(in)}$ is defined as the set of pixels for which the levelset is negative. However, the reconstructed levelsets have been transformed, and to bring them into the coordinate frame of the image C_θ , the transformation must be inverted prior to thresholding.

$$C_{m,\theta}^{(in)} = \{c | l_{m,\theta}(T^{*-1}(c)) \leq 0\}, \quad (7)$$

where $R(\widehat{L}_{1,T^*}) = \{l_{m,\theta}(c) | c \in C_\theta\}$. A graphical representation of the segmentation steps is shown in Figure 2.

3. Experimental Design

3.1. Data Description

Our data consists of 40 prostate endorectal MR images, acquired using T2-weighting protocol and a 3.0 Tesla coil. Each image was 512×512 pixels in the x, y directions with a variable number of slices. The prostate capsule, PZ, and CG boundaries were manually segmented in 3D by an expert radiologist using the 3D Slicer software [38, 39, 40]. The raw data for each study was preprocessed to normalize the intensities and remove the bias field

[41]. In addition, the variance of each levelset and each intensity image was normalized to a value of 1. Due to the fact that MR imagery of the prostate is used for staging of prostate cancer in the US, and not for screening, all 40 studies have biopsy-confirmed prostate cancer present. A full description of our dataset and associated parameters is shown in Table 2.

3.2. Implementation Details

The MLA was implemented in C++ using the ITK framework [42]. The MLA was run on a machine with 8 cores (each 2.67 GHz) and 32 GB of memory running Debian Linux, compiled using GCC (version 4.7.1). The segmentation process was performed in a multi-resolution fashion, with $P \approx 10^6$ at the coarsest resolution and $P \approx 10^7$ in the finest resolution. Segmenting the prostate, CG, and PZ on a 140 mm \times 140 mm \times 140 mm image took approximately 200 seconds. For all experiments, $\alpha = 0.95$ was used, similar to [7].

3.3. Hierarchical Prostate Segmentation

Two specific categories of experiments were performed: non-hierarchical experiments (E_1, E_3, E_5) and hierarchical experiments (E_2, E_4, E_6). The non-hierarchical experiments used only the imaging information to simultaneously segment the prostate, CG, and PZ. Therefore, for $K = \emptyset$, as only the intensities were used to segment the objects. The hierarchical experiments used the imaging information, as well as the known segmentation of the prostate, to segment the CG and PZ, and thus $K = \{\text{Prostate}\}$. Due to the fact that CG and PZ are embedded within the prostate itself, using a segmentation of the prostate boundary forces the MLA to only consider the desired region of interest. This is also similar to the approach taken in [33], which assumed the prostate was already segmented prior to segmenting the PZ and CG.

3.4. Cross Validation Experiments

For both the hierarchical experiments and the non-hierarchical experiments, both a leave-one-out cross validation, and a 30-run, 5-fold cross validation were performed. For the leave one out experiments (E_1, E_2), for each image C_m , the MLA was trained using the other 39 studies. For each run of 5-fold cross validation experiments (E_3 – E_6), the dataset was randomly split into 5 groups of 8 studies per group. Each study in a given group was segmented using an MLA trained from the 32 studies in the other 4 groups, resulting in a segmentation for each study. This was repeated 30 times, resulting in 30 segmentations for each study. The goal of the cross validation experiment is to determine the generalizability of the MLA to different training sets, by determining how well each study was segmented given different training sets, and the variance of the results over the various training sets.

The segmentation result $C_{m,\theta}^{(in)}$ was compared to the ground truth segmentation $C_{m,n}^{(in)}$ using the Dice similarity coefficient (DSC).

$$DSC(C_{m,n}^{(in)}, C_{m,\theta}^{(in)}) = 2 \cdot \frac{|C_{m,\theta}^{(in)} \cap C_{m,n}^{(in)}|}{|C_{m,\theta}^{(in)}| + |C_{m,n}^{(in)}|}. \quad (8)$$

In addition, the Mean Absolute Distance (MAD) between the surfaces was reported, calculated as,

$$MAD(C_{m,n}^{(in)}, C_{m,\theta}^{(in)}) = \frac{1}{|C_{m,\theta}^{(in)}|} \cdot \sum_{c \in C_{m,\theta}^{(in)}} \left(\min_{d \in C_{m,n}^{(in)}} \|c - d\|_2 \right) \quad (9)$$

where $C^{(on)}$ represents pixels on the surface of the object, and the MAD values are reported in mm.

The results from the prostate, PZ, and CG segmentations from the non-hierarchical experiments (E_1 , E_3 , E_5) are presented, in addition to the PZ, and CG segmentation results from the hierarchical experiments (E_2 , E_4 , E_6). The segmentation results in the midgland of the prostate are presented separately for the area-based *DSC* values (the boundary-based *MAD* is not easily defined for separate regions). This was done due to poor boundary contrast in the base and apex of the prostate, preventing accurate segmentations in these regions.

The cross validation experiments resulted in 30 values for each of the 40 studies (1200 total values). Experiments E_3 and E_4 present the results over all 1200 values for the non-hierarchical and hierarchical experiments respectively. To determine the generalizability over different training sets, the median value was computed over the 40 values for each run. The results over the 30 different median values (1 for each run) are presented as E_5 and E_6 . Table 3 summarizes the different experiments performed.

4. Results

The quantitative results for the prostate, CG, and PZ for experiments E_1 through E_6 are shown in Figure 6. The segmentation of the capsule boundary resulted in a mean *DSC* accuracy of 0.81, and a mean *MAD* value of 1.8 mm. When only considering the midgland of the prostate, the mean *DSC* value for the prostate increased to 0.89, reflecting the tapering off of the gland towards the base and apex.

The hierarchical segmentation results E_2 , in which it is assumed that the prostate segmentation already exists, as in [33], resulted in a mean *DSC* value of 0.79 for the CG, and 0.68 for the PZ, with mean *MAD* values of 1.4 mm and 1.0 mm for the CG, and PZ, respectively. When only considering the midgland, the mean *DSC* values were 0.84, and 0.76 for the CG and PZ, respectively. However, when only using the imaging information (non-hierarchical experiment E_1), the mean *DSC* values for the CG and PZ were 0.72 and 0.60 respectively.

Qualitative results from two studies are shown in Figures 7 and 8. The region in green represents the ground truth segmentations and red represents the segmentation results. The *DSC* values for the prostate, CG, and PZ, were .88, .86, and .76 respectively in Figure 7. The *DSC* values for the prostate, CG, and PZ, were .90, .71, and .73 respectively in Figure

8. In addition, the reconstruction $R(\widehat{F}_{r^*})$ is shown in Figures 7c and 8c, which demonstrates that the intensities in a previously unseen study can be reasonably well reconstructed.

5. Discussion

The closest related work is [33], which reported mean *DSC* values of 0.89 and 0.80 for the CG and PZ segmentations respectively on multi-spectral (T2-weighted and dynamic contrast enhanced), 1.5 Tesla prostate MRI. When using the ground truth prostate segmentations to drive the CG and PZ segmentations, similar to [33], resulted in *DSC* accuracies of 0.79 and 0.68 for the CG and PZ for the MLA. However, it should be noted that those reported *DSC* values in [33] were from a combined STAPLE segmentation of three expert ground truths. When considering only one of the ground truths in [33], the mean *DSC* values decreased to 0.82 and 0.71 respectively. This is a more appropriate and fairer comparison to the results presented in this work, and also reflects the difficulty of getting accurate expert

segmentations for the CG and PZ from prostate MRI. In addition, the data used in [33] contained 31 studies (as compared to the 40 we employed in this study).

Moreover, the algorithm in [33] was specifically designed to intelligently take into account data from multiple modalities such as T1 contrast enhanced and diffusion weighted images, to complement the T2-weighted MR imagery. This allowed the algorithm to extract more accurate CG and PZ boundaries by leveraging additional information that may not be present in T2-weighted MRI. It is not clear how well the algorithm in [33] would perform if only T2-weighted MRI images were available (as in the current dataset). In addition, it is also unclear how well the MLA would perform if other MRI protocols were used in addition to, or instead of, T2-weighted intensities, so a direct comparison is difficult. However, it is important to note that while T2-weighted MRI is routinely performed at all sites where prostate MRI is performed, multi-parametric MRI is only done in a subset of those imaging facilities. Hence our algorithm could be employed on data from a larger number of centers.

To the best of our knowledge, this is the first work exploring a fully automated CG and PZ segmentation algorithm, as [33] only reported results using the ground truth prostate segmentations as the inputs to the algorithm.

In all the cases, the 5-fold cross validation experiments performed worse than the leave-one-out experiments, mainly due to the 20% fewer studies used to train the MLA (32 versus 39 training studies), suggesting the need for a large training cohort. However, the median *DSC* and *MAD* values between cross validation runs were remarkably consistent, suggesting very little variance between different training sets.

6. Concluding Remarks

In this work we present a Multi-Levelset AAM (MLA), which offers several unique improvements over the traditional AAM algorithm. The use of anatomical landmarks, which can be prone to errors, is completely eliminated by casting the MLA shape model in a levelset framework. The MLA is able to couple multiple levelsets with image intensities efficiently and intelligently, to offer a coupled model resulting in simultaneous segmentations of multiple objects. The MLA presented uses the image intensities, as well as existing segmentations, to drive the simultaneous segmentations of multiple objects. This is accomplished using knowledge of the shapes of various objects, and how those shapes correlate with intensities.

The algorithm was tested on 40 T2-weighted, 3D, endorectal, 3.0 Tesla, prostate MRI images containing ground truth segmentations of the prostate, central gland (CG), and peripheral zone (PZ). Most existing prostate segmentation algorithms only segment the prostate boundary, and yet CG and PZ segmentations are critical for cancer detection and treatment planning. Future work will entail determining quantitatively how the presence and extent of disease affects the MLA.

When using the intensities and known prostate segmentations, mean *DSC* values of 0.79 and 0.68 were reported for the CG and PZ, respectively. We believe that with more training instances, and with inclusion of multi-parametric data, the segmentation results could be further refined. Overall, the MLA framework is a generalized extension to the AAM segmentation algorithm, and can be used to accurately and automatically segment multiple objects.

Acknowledgments

This work was made possible via grants from the National Institute of Health (Grant Nos. R01CA136535-01, R01CA140772-01, R21CA167811-01).

References

- [1]. Heimann T, Meinzer H. Statistical shape models for 3D medical image segmentation: a review. *Medical Image Analysis*. 2009; 13:543–563. [PubMed: 19525140]
- [2]. Cootes T, Taylor C, Lanitis A. Multi-resolution search with active shape models. *Computer Vision and Image Processing*. 1994; 1:610–612. vol.1.
- [3]. Jolesz F, Nabavi A, Kikinis R. Integration of interventional MRI with computer-assisted surgery. *Journal of Magnetic Resonance Imaging*. 2001; 13:69–77. [PubMed: 11169806]
- [4]. van Ginneken B, Frangi A, Staal J, Romeny B, Viergever M. Active shape model segmentation with optimal features. *Medical Imaging, IEEE Transactions on*. 2002; 21:924–933.
- [5]. Larsen R, Stegmann M, Darkner S, Forchhammer S, Cootes T, Ersboll B. Texture enhanced appearance models. *Computer Vision and Image Understanding*. 2007; 106:20–30.
- [6]. Toth R, Doyle S, Pungavkar S, Kalyanpur A, Madabhushi A. Weritas: Weighted ensemble of regional image textures for asm segmentation. *SPIE Medical Imaging*. volume 7260
- [7]. Cootes T, Edwards G, Taylor C. Active appearance models *Pattern Analysis and Machine Intelligence*. *IEEE Transactions on*. 2001; 23:681–685.
- [8]. Cootes T, Taylor C, Cooper D, Graham J. Active shape models - their training and application. *Computer Vision and Image Understanding*. 1995; 61:38–59.
- [9]. Styner, M.; Rajamani, K.; Nolte, L.; Zsemlye, G.; Székely, G.; Taylor, C.; Davies, R. *IPMI*. Vol. volume 2732. Springer; Berlin / Heidelberg: 2003. Evaluation of 3D correspondence methods for model building; p. 63-75.
- [10]. Morris D, Kanade T. Image-consistent surface triangulation. *Computer Vision and Pattern Recognition*. 2000; 1:1332.
- [11]. Leventon M, Grimson W, Faugeras O. Statistical shape influence in geodesic active contours. *Computer Vision and Pattern Recognition*. volume 1:316–323.
- [12]. Tsai A, Wells W, Tempany C, Grimson E, Willsky A. Mutual information in coupled multi-shape model for medical image segmentation. *Medical Image Analysis*. 2004; 8:429–445. [PubMed: 15567707]
- [13]. Akhondi-Asl A, Soltanian-Zadeh H. Effect of number of coupled structures on the segmentation of brain structures. *Journal of Signal Processing Systems*. 2009; 54:215–230.
- [14]. Chen S, Lovelock D, Radke R. Segmenting the prostate and rectum in ct imagery using anatomical constraints. *Medical Image Analysis*. 2011; 15:1–11. [PubMed: 20634121]
- [15]. Fitterer J, Barentsz J. 3T MRI of prostate cancer. *Applied Radiology*. 2009; 38:25–32.
- [16]. Akin O, Sala E, Moskowitz C, Kuroiwa K, Ishill N, Pucar D, Scardino P, Hricak H. Transition zone prostate cancers: Features, detection, localization, and staging at endorectal MR imaging. *Radiology*. 2006; 239:784–792. [PubMed: 16569788]
- [17]. Erbersdobler A, Augustin H, Schlomm T, Henke R. Prostate cancers in the transition zone: Part 1; pathological aspects. *BJU International*. 2004; 94:1221–1225. [PubMed: 15610093]
- [18]. Augustin H, Erbersdobler A, Hammerer P, Graefen M, Huland H. Prostate cancers in the transition zone: Part 2; clinical aspects. *BJU International*. 2004; 94:1226–1229. [PubMed: 15610094]
- [19]. Viswanath S, Bloch B, Rosen M, Chappelow J, Rofsky N, Lenkinski R, Genega E, Kalyanpur A, Madabhushi A. Integrating structural and functional imaging for computer assisted detection of prostate cancer on multi-protocol in vivo 3 tesla MRI. *SPIE Medical Imaging*. 2009; 7260
- [20]. Viswanath S, Bloch B, Chappelow J, Toth R, Genega E, Lenkinski R, Rofsky N, Madabhushi A. Central gland and peripheral zone prostate tumors have significantly different quantitative imaging signatures on 3 tesla endorectal, in vivo t2-weighted magnetic resonance imagery. *Journal of Magnetic Resonance Imaging (Accepted)*. 2012

- [21]. Madabhushi A, Feldman M, Metaxas D, Tomasezewski J, Chute D. Automated detection of prostatic adenocarcinoma from high resolution ex vivo MRI Medical Imaging. *IEEE Transactions on.* 2005; 24:1611–1625.
- [22]. Lopes R, Ayache N, Makni N. Prostate cancer characterization on MR images using fractal features. *Medical Physics.* 2010; 38:83–95. [PubMed: 21361178]
- [23]. Vos P, Hambrock T, Barenstz J, Huisman H. Computer-assisted analysis of peripheral zone prostate lesions using T2-weighted and dynamic contrast enhanced T1-weighted MRI. *Physics in Medicine and Biology.* 2010; 55:1719. [PubMed: 20197602]
- [24]. Ozer S, Langer D, Liu X. Supervised and unsupervised methods for prostate cancer segmentation with multispectral MRI. *Medical Physics.* 2010; 37:1873–1883. [PubMed: 20443509]
- [25]. Liu X, Yetik I. Automated prostate cancer localization without the need for peripheral zone extraction using multiparametric MRI. *Medical Physics.* 2011; 38:2986–2994. [PubMed: 21815372]
- [26]. Shannon B, McNeal J, Cohen R. Transition zone carcinoma of the prostate gland: a common indolent tumour type that occasionally manifests aggressive behaviour. *Pathology.* 2003; 35:467–471. [PubMed: 14660095]
- [27]. Klein S, van der Heide U, Lips I, vanVulpen M, Staring M, Pluim J. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Medical Physics.* 2008; 35:1407–1417. [PubMed: 18491536]
- [28]. Martin S, Daanen V, Troccaz J. Automated segmentation of the prostate in 3D MR images using a probabilistic atlas and a spatially constrained deformable model. *Medical Physics.* 2010; 37:1579–1590. [PubMed: 20443479]
- [29]. Pasquier D, Lacormerie T, Vermandel M, Rousseau J, Lartigau E, Betrouni N. Automatic segmentation of pelvic structures from magnetic resonance images for prostate cancer radiotherapy. *International Journal of Radiation Oncology and Biological Physics.* 2007; 68:592–600.
- [30]. Makni N, Puech P, Lopes R, Dewalle A. Combining a deformable model and a probabilistic framework for an automatic 3D segmentation of prostate on MRI. *International Journal of Computer Assisted Radiology and Surgery.* 2009; 4:181–188. [PubMed: 20033618]
- [31]. Toth R, Bloch B, Genega E, Rofsky N, Lenkinski R, Rosen M, Madabhushi A. Accurate prostate volume estimation using active shape models on T2-weighted MRI. *Academic Radiology.* 2011; 18:745–754. [PubMed: 21549962]
- [32]. Bulman J, Toth R, Patel A, Bloch N, C.J. M, L. N, Madabhushi A, Rofsky N. Automated computer-derived prostate volumes from MRI data: Comparison to radiologist-derived mri volumes and pathology specimen volumes. *Radiology.* 2012; 262:144–151. [PubMed: 22190657]
- [33]. Makni N, Iancu A, Puech P, Colot O, Mordon S, Betrouni N. Zonal segmentation of prostate using multispectral magnetic resonance images. *Medical Physics.* 2011; 38:6093–6105. [PubMed: 22047374]
- [34]. Toth R, Madabhushi A. Multi-feature landmark free active appearance models: Application to prostate MRI segmentation, *Medical Imaging. IEEE Transactions on.* 2012; 38:1638–1650.
- [35]. Albert, A. *Regression and the Moore-Penrose Pseudoinverse.* Academic Press; 1972.
- [36]. Wachowiak M, Smolikova R, Zheng Y, Zurada J, Elmaghraby A. An approach to multimodal biomedical image registration utilizing particle swarm optimization, *Evolutionary Computing. IEEE Transactions on.* 2004; 8:289–301.
- [37]. Powell M. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal.* 1964; 7:155–162.
- [38]. Pieper S, Lorensen B, Schroeder W, Kikinis R. The na-mic kit: Itk, vtk, pipelines, grids and 3D slicer as an open platform for the medical image computing community. *Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: From Nano to Macro.* :698–701.
- [39]. Pieper S, Halle M, Kikinis R. 3D slicer. *Proceedings of the 1st IEEE International Symposium on Biomedical Imaging: From Nano to Macro.* :632–635.
- [40]. Gering D, Nabavi A, Kikinis R, Grimsom W, Hata N, Everett P, Jolesz F, Wells W. An integrated visualization system for surgical planning and guidance using image fusion and interventional imaging. *MIC-CAI.* volume 2:809–819.

- [41]. Cohen M, Dubois R, Zeineh M. Rapid and effective correction of rf inhomogeneity for high field magnetic resonance imaging. *Human Brain Mapping*. 2000; 10:204211.
- [42]. Yoo, T.; Ackerman, M.; Lorensen, W.; Schroeder, W.; Chalana, AS.; V.; Metaxas, D.; Whitaker, R. Engineering and algorithm design for an image processing api: A technical report on itk - the insight toolkit. In: Westwood, e., editor. *Proc. of Medicine Meets Virtual Reality*. IOS Press; Amsterdam: 2002.

- Landmark free Active Appearance Model can segment multiple objects simultaneously
- Prostate, Central Gland, Peripheral Zone segmented with Dice = .81, .79, .68
- Hierarchical system uses existing segmentations to drive other segmentations

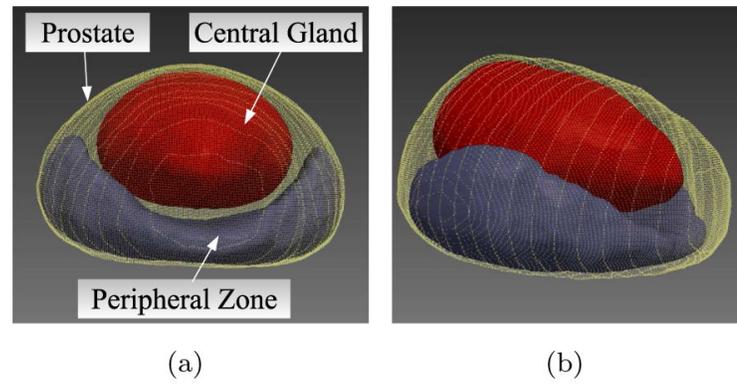


Figure 1. Two different 3D views of the prostate (yellow) with the central gland (CG) (red) and peripheral zone (PZ) (purple) segmented.

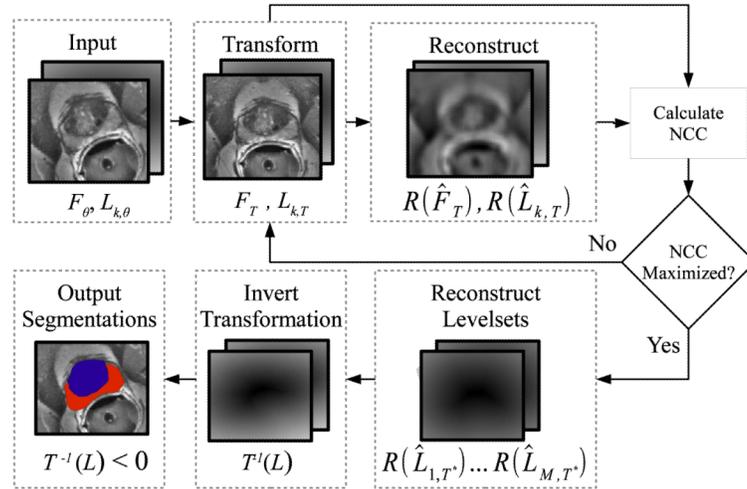
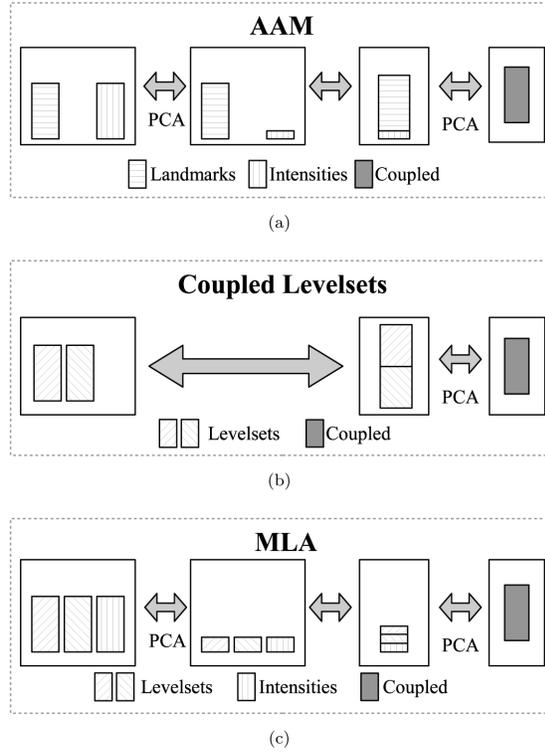


Figure 2.

The segmentation process begins with an intensity image F_{θ} and input levels $L_{k,\theta}$. The inputs are transformed as $F_T, L_{k,T}$. The coupled MLA is used to reconstruct the transformed inputs, as $R(\hat{F}_T), R(\hat{L}_{k,T})$. When the normalized cross correlation (NCC) between the reconstructions and transformed inputs is maximized, the MLA is used to reconstruct all levelsets. The levelsets are transformed back into the image's original coordinate frame, and thresholded to yield a final set of segmentations.

**Figure 3.**

(a) Traditional AAMs [7] perform PCA on a set of intensities, and couple the intensity projections with the shape by performing a second PCA. (b) Traditional coupled levelsets [12] concatenate the high dimensional levelsets, and couple the shapes by performing PCA on the result. However, performing PCA can be computationally infeasible with either a large number of levelsets, high dimensionality of the levelsets, or with a large number of training images. (c) The MLA projects each levelset to a low dimensional space prior to coupling. In addition, the levelsets are coupled with intensities, similar to a traditional AAM. Each step (PCA and concatenation) in the case of each of the 3 models is reversible, where training involves generating the final coupled model, and segmentation step involves reconstructing the original high dimensional data from the coupled model.

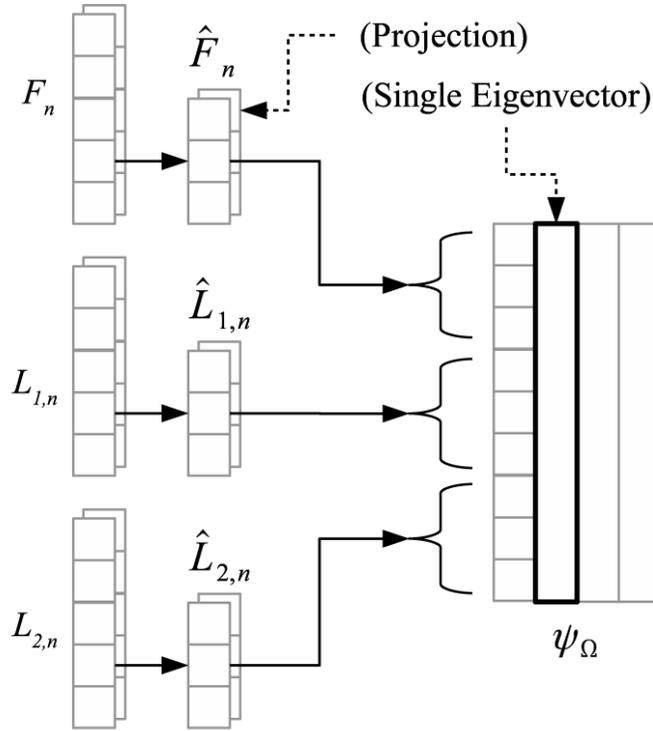


Figure 4. Coupling intensities with multiple levelsets for the MLA training. First, PCA is used to project each intensity F_n and levelset $L_{m,n}$ down into a lower dimensional space, denoted as $\hat{F}_n, \hat{L}_{m,n}$. Then, a second PCA is performed, creating a coupled matrix Ψ_Ω , where each row corresponds to a specific projection value, and each column represents a single eigenvector.

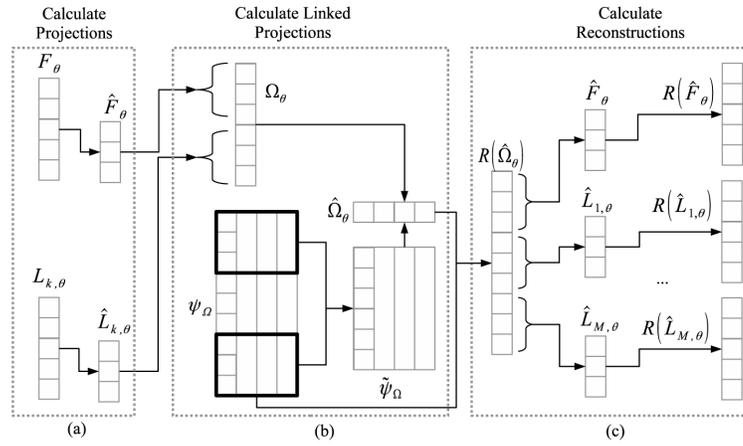
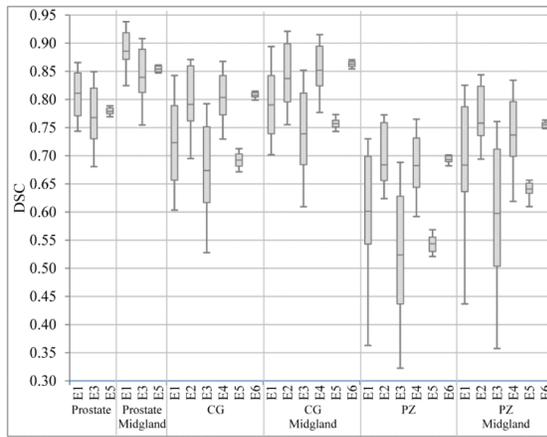
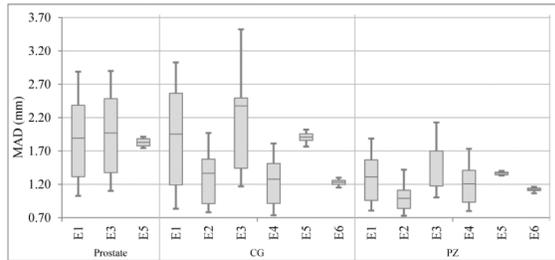


Figure 5. (a) The projections from the input intensities and levelsets are calculated using Equation (2). (b) The rows from the coupled matrix Ψ_{Ω} corresponding to the given inputs are extracted as $\tilde{\psi}_{\Omega}$. The projections are used to estimate a set of coupled projections $\widehat{\Omega}_{\theta} \in \mathbb{R}^{\tilde{p}}$. (c) The coupled projections $\widehat{\Omega}_{\theta}$ are used to reconstruct the entire set of projections $\widehat{F}_{\theta}, \widehat{L}_{1,\theta}, \dots, \widehat{L}_{M,\theta}$. Finally, a set of reconstructions $R(\widehat{F}_{\theta}), R(\widehat{L}_{1,\theta}), \dots, R(\widehat{L}_{M,\theta})$ are calculated using Equation (4).



(a) Prostate, CG, PZ DSC Values



(b) Prostate, CG, PZ MAD Values

Figure 6. Quantitative segmentation results from experiments E_1 through E_6 for 40 studies segmenting the prostate, CG, and PZ. The mean is given by a solid gray line, the 25th – 75th percentiles are shown as a shaded gray rectangle, and the 10th – 90th percentiles are shown as error bars.

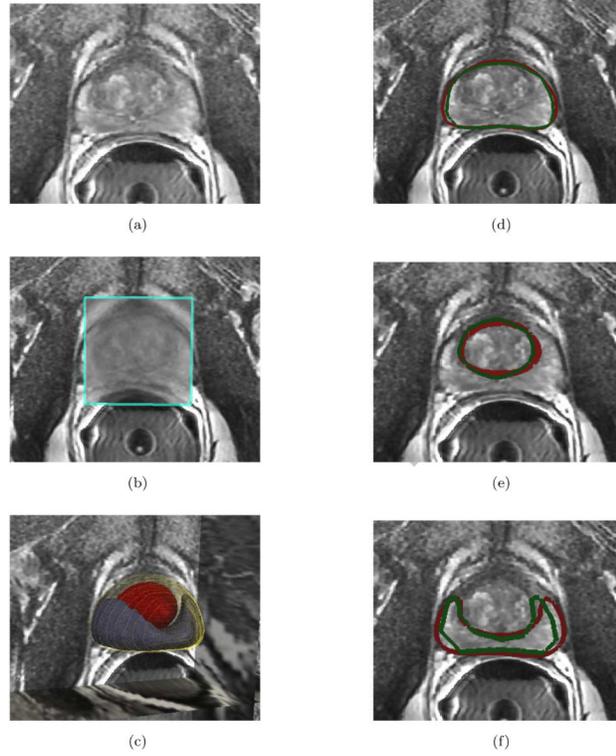


Figure 7.

(a) illustrates the intensities F_θ of a midland prostate slice from a 3D, T2-w, endorectal MR image. (b) represents the reconstruction $R(\widehat{F}_{T^*})$ resulting from the MLA. (c) illustrates a 3D rendering of the prostate in light yellow, CG in red, and PZ in dark purple. In (d), (e), and (f), the MLA segmentations are shown in red while the ground truth segmentations are shown in green. (d) illustrates the prostate, with $DSC = 0.878$. (e) represents the CG, with $DSC = 0.861$. (f) illustrates the PZ, with $DSC = 0.764$.

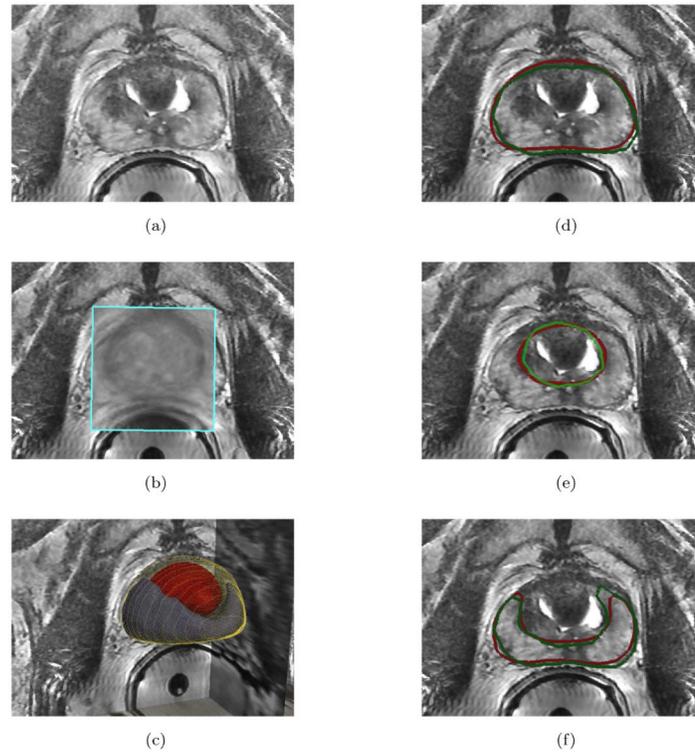


Figure 8.

a) illustrates the intensities F_θ of a midgland prostate slice from a 3D, T2-w, endorectal MR image. (b) illustrates the reconstruction $R(\widehat{F}_{T^*})$ resulting from the MLA. (c) illustrates a 3D rendering of the prostate in light yellow, CG in red, and PZ in dark purple. In (d), (e), and (f), the MLA segmentations are shown in red while the ground truth segmentations are shown in green. (d) illustrates the prostate, with $DSC = 0.90$. (e) illustrates the CG, with $DSC = 0.81$. (f) illustrates the PZ, with $DSC = 0.73$.

Table 1

List of commonly used notations and symbols in this paper

Symbol	Description	Symbol	Description
M	Number of objects in an image.	N	Number of training images.
C	Set of spatial locations, where $c = (x, y, z)$ for each $c \in C$	$C(in)$	Set of spatial locations inside an object, where $C(in) \subset C$.
F	Set of intensities, where $f(c)$ is the intensity at location c .	\hat{F}	Low dimensional projection of intensities, where $ \hat{F} \ll F $.
L	Levelset of an object, where $L = 0$ represents the object surface.	\hat{L}	Low dimensional projection of a levelset, where $ \hat{L} \ll L $.
S_m	Set of N levelsets for object $m \in \{1, \dots, M\}$ in the training set.	S_F	Set of N intensities in the training set.
Ω	Concatenation of multiple projections.	Ψ	Matrix of eigenvectors, where each column is an eigenvector.
T	A±ne transformation.	R	High dimensional reconstruction, calculated from a low dimensional projection.

Table 2

Detailed description of the data used to test the MLA

# of Studies	Protocol	MRI Acquisition
40 Studies	3.0 Tesla, T2-weighted	Fast Spin Echo, Endorectal Coil
Image Size (pixels)	Field of View (mm)	Resolution (mm)
512x512xZ, 20 < Z < 50	140 x 140 x Z, 60 < Z < 150	0:27 x 0:27 x 3:0

Table 3

Description of the leave-one-out and cross validation experiments performed to quantitatively test the MLA. The difference between E_3/E_4 and E_5/E_6 is that with E_3/E_4 we calculate the results over all 1200 trials (30 runs x 40 studies), while with E_5/E_6 we calculate the median value for each of 30 runs and calculate the results over all 30 runs to determine the generalizability of the MLA over different training sets

Experiment	Hierarchical	Runs	Folds	Total # of Trials
E_1	No	1	40	40
E_2	Yes	1	40	40
E_3	No	30	5	1200
E_4	Yes	30	5	1200
E_5	No	30	5	30
E_6	Yes	30	5	30