A Comparison of US versus MR Based 3D Prostate Shapes Using Radial Basis Function Interpolation and Statistical Shape Models

Ran Tao, Student Member IEEE, Mahdi Tavakoli, Ph.D., Member, IEEE, Ron Sloboda, Ph.D. and Nawaid Usmani, M.D.

Abstract—This paper presents a comparison of 3D segmentations of the prostate, based on 2D manually-segmented contours, obtained using ultrasound (US) and magnetic resonance (MR) imaging data collected from 40 patients diagnosed with localized prostate cancer and scheduled to receive brachytherapy treatment. The approach we propose here for 3D prostate segmentation first uses radial basis function interpolation to construct a 3D point distribution model for each prostate. Next, a modified principal axis transformation is utilized for rigid registration of the US and MR images of the same prostate in preparation for the following shape comparison. Then, statistical shape models are used to capture the segmented 3D prostate geometries for the subsequent cross-modality comparison. Our study includes not only cross-modality geometric comparisons in terms of prostate volumes and dimensions, but also an investigation of interchangeability of the two imaging modalities in terms of automatic contour segmentation at the pre-implant planning stage of prostate brachytherapy treatment. By developing a new scheme to compare the two imaging modalities in terms of the segmented 3D shapes, we have taken a first step necessary for building coupled US-MR segmentation strategies for prostate brachytherapy pre-implant planning, which at present is predominantly informed by ultrasound images only.

Index Terms—Prostate brachytherapy, radial basis function, rigid registration, segmentation, statistic shape model

I. INTRODUCTION

Prostate cancer has been identified as the most common non-dermal cancer in North American men [1], [2]. Amongst various treatment options including surgery, external beam radiotherapy and hormone therapy, brachytherapy has emerged as a popular minimally invasive approach with excellent long-term results. Consensus guidelines developed by the American Brachytherapy Society suggest using biplanar transrectal ultrasonography (TRUS) for pre-operative planning and intra-operative stages, with computed tomography (CT) or magnetic resonance (MR) imaging utilized for post-operative dosimetry [3]. While ultrasound (US) imaging is predominantly the imaging modality of choice in pre-implant dosimetry, magnetic resonance (MR) imaging presents certain advantages over it such as improved tissue contrast. An important question is, can MR imaging replace or complement US imaging in prostate brachytherapy pre-implant planning? A few studies have been carried out to investigate the feasibility of replacing US imaging with MR in pre-implant planning [4] and intraoperative stages [5]-[7].

Given that traditional manual segmentation of the prostate, which is a prerequisite for pre-implant planning, is time-consuming and tedious for clinicians and vulnerable to inter- and intra-observer variability, automatic or semi-automatic prostate segmentation techniques have been developed in 2D and 3D for alleviating clinicians’ workload and reducing contour segmentation times. Such techniques include edge and texture based segmentation [8]-[10], deformable model based segmentation [11], [12], and ellipsoid fitting [13], [14]. Among these techniques, one method that has seen increasing success in medical image segmentation and that we will use is the statistical shape model (SSM) based segmentation [15], lately applied to prostate segmentation [8], [9], [16]. It uses prior statistical information to create a model for the spatial shapes of prostate boundaries.

Recently, dual modality segmentation methods based on SSM have been developed. For instance, in [17], a cardiac left ventricle SSM model was built based on MR images with superior tissue definition and then used for contour segmentation in CT images without the need to re-train the model with CT datasets. In [18], a scheme for concurrent segmentation of the prostate using MR and CT images via linked SSM was proposed, again leveraging the superior soft tissue definition of MR over CT to improve the performance of CT-based segmentation in prostate radiotherapy planning. Similar to these efforts concerning dual-modality segmentation in various medical applications, in prostate brachytherapy, coupled strategies involving both US and MR imaging can be devised for pre-implant planning where TRUS is predominantly the modality of choice at present.
Since the segmentation of prostate in clinical images provides the very foundation of brachytherapy pre-implant planning, the choice of imaging modality (e.g., US vs. MR) changes the perception of the prostate in terms of volume and shape and may have implications to current practices [4], [19]. Studies have been carried out to compare the prostate shapes obtained by different imaging modalities. In [4], MR-based and TRUS-based preplanning in prostate brachytherapy were compared using dosimetric results, and prostate volume measured in MR and TRUS were also compared. In [20], prostate shapes obtained from post-implant 3DTRUS, MR and CT images were compared in terms of volume and contouring variability. Both intra- and inter-modality variation was analyzed in 3D. In [19], pre-implant TRUS and MR images were used for cross-modality shape and contouring variability comparison. Pre-implant images are not subject to post-implant edema. Specific attention was given to the effect of the TRUS probe induced deformations in this study. Therefore, replacing or complementing US with MR imaging needs to be done only after careful studies. To address this need, in this paper we will compare US and MR based prostate segmentation in 3D from a new perspective that utilizes models built from statistical shape analysis. Our approach will not only allow for a statistical geometric comparison between the two imaging modalities in terms of segmented 3D shapes, but also provide insight into the influences and intrinsic differences of the two imaging modalities (including the effect of the transrectal ultrasound probe) on the SSM-based analysis results. As elaborated later, our analyses help to correlate different modes of the SSM to aspects of the prostate shape that are interesting clinically.

To be more specific about the steps taken in the paper, after first using manually delineated contour points in US and MR images of each prostate in a 40-patient database, radial basis function (RBF) is used to convert each set of 2D contours into a 3D surface model; RBFs have previously been successful in generating 3D models for prostate [21] and other organs from point clouds [22]-[24]. While in [21] a fan-shaped TRUS sweeping was used to obtain non-parallel prostate contours, our study acquired parallel prostate contours in both TRUS and MR under the conditions of prostate brachytherapy. While, to the best of the authors’ knowledge, using parallel prostate contours for RBF interpolation was without precedence in the literature, the success of RBF interpolation in similar contexts in the literature made it a promising choice for prostate modeling in our application.

Next, the paper investigates the differences between pre-implant US and pre-implant MR prostate contours via an analysis of SSM models built from both data groups; few studies have been conducted to directly compare US and MR imaging modalities in terms of prostate shape delineation and even fewer have used pre-implant images, which are particularly pertinent to our brachytherapy treatment planning.

On the other hand, before working towards the establishment of coupled US-MR pre-implant planning strategies, the SSMs built from the two individual modalities need to be examined for their interchangeability. To our knowledge, such comparisons based on SSMs built from pre-implant US and MR prostate contours do not exist in the literature and constitute one of the contributions of this paper; our approach provides a new method for image modality comparison and constitutes a first step towards a US-MR coupled prostate segmentation scheme. The proposed comparison method is completely open-source and reproducible as it does not rely on any commercial software particular to prostate brachytherapy.

This paper is organized as follows. Section II presents the methods and materials used for data collection (II.A), 3D prostate model development (II.B), and SSM implementation (II.C). The paper will continue to discuss the results of the comparison between models built from US and MR imaging modalities in Section III, including average shape comparisons (III.A), SSM principal component analysis (III.B) and interchangeability investigation (III.C and III.D). In Section IV, our results are compared to previous studies both from a geometric point of view (IV.A) and an SSM point of view (IV.B). Finally, concluding remarks are provided in Section V.

II. METHODS AND MATERIALS

A. Imaging and Contouring

The study used imaging data from 40 consenting patients diagnosed with localized prostate cancer, who were scheduled to receive 125I brachytherapy treatment. Prior to the implant procedure, each patient received a pre-treatment planning scan using a Sonoline Adara TRUS scanner with an Endo PII probe (Siemens Medical Solutions USA, Inc., Malvern, PA), yielding a set of parallel axial images spaced at 5 mm intervals from the base to the apex of the prostate. Four to six weeks after the US scan and still prior to the implant procedure, an MR scan was performed using a 1.5 T Gyroscan Intera Imager (Philips Healthcare, Andover, MA) with a 5-channel cardiac coil. The acquired MR image slices were spaced with spatial depth increments of 3-4 mm; the prostate boundaries were made clearly visible by applying a T2-weighted spin echo pulse sequence [25] and the consistency in pelvic orientation was assured by the use of an under-knee rest. A recapitulation of the patient group information is shown in Table I. In Table I PSA is the abbreviation for ‘prostate specific antigen’. Gleason score is the score a pathologist gives to a prostate cancer based on its microscopic features at diagnosis that help characterize the aggressiveness of the cancer. Clinical stage refers to the extent of disease within the prostate that is palpated by a physician during a digital rectal examination, with severity increasing from T1 to T4.

The 40 image volumes for each imaging modality were then imported into folders on VariSeed system (Varian Medical Systems, Inc., Palo Alto, CA) for contouring. All volumes were anonymized and randomized to minimize contouring bias associated with expectation of a given patient’s prostate shape. Then, three experienced oncologists performed contouring on all images. The contours, consisting of discrete points in parallel planes, were exported from the VariSeed planning system into MATLAB; see Fig. 1(a). In the orthonormal coordinate system shown, X and Y axes specified the image plane while the Z axis lied in the probe direction and oriented from the base to the apex of each prostate.
After importing these 2D contours into MATLAB, they were first interpolated in their own plane (i.e., XY plane) to decrease the spatial inter-sample gap; as a result, the total number of points per slice was increased to 60 (from an average approximately 22 points per slice) with the interpolated points being at equal angular distances from each other with respect to the center of mass (CoM) of each 2D contour. To do this, we used spline interpolation as it has been found well capable of representing prostate shapes [11]. Then, a cubic smoothing spline was applied to the interpolated contour in order to give it a smooth shape. The original (manually delineated) and the resampled contours of a typical slice are shown in Fig. 1.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>PATIENT POPULATION CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>40</td>
</tr>
<tr>
<td>Median age (yrs)</td>
<td>64 (range: 51 - 79)</td>
</tr>
<tr>
<td>Median pretreatment PSA (ng/mL)</td>
<td>6.3 (range: 2.0 - 14.0)</td>
</tr>
<tr>
<td>Gleason score:</td>
<td></td>
</tr>
<tr>
<td>≤ 5</td>
<td>2 (5 %)</td>
</tr>
<tr>
<td>6</td>
<td>36 (90 %)</td>
</tr>
<tr>
<td>7</td>
<td>2 (5 %)</td>
</tr>
<tr>
<td>Clinical stage:</td>
<td></td>
</tr>
<tr>
<td>T1c</td>
<td>24 (60 %)</td>
</tr>
<tr>
<td>T2a</td>
<td>9 (22 %)</td>
</tr>
<tr>
<td>T2b</td>
<td>7 (18 %)</td>
</tr>
<tr>
<td>Risk group:</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>26 (65 %)</td>
</tr>
<tr>
<td>Intermediate risk</td>
<td>14 (35 %)</td>
</tr>
</tbody>
</table>

Fig. 1. (a) Prostate gland image from TRUS. Manually-delineated contour points are in yellow squares. (b) Resampled points in white circles are superimposed on the original manually-delineated contour points.

B. Radial Basis Function (RBF) Interpolation

Noting that the outcome of contouring in Section II.A is a set of discrete contour points in parallel planes in the axial direction of the prostate and that our aim is to compare the prostate shapes obtained from US and MR imaging modalities in 3D, it is necessary to first fit a 3D surface to the 2D contour points for each prostate. The method applied here is the RBF interpolation, which is capable of effectively interpolating point data generated from a non-standard grid possibly with large data-free gaps [23]. In fact, since the gaps between our image slices range from 3 mm to 5 mm, RBF is ideal in our application for interpolating the sparse collection of contours obtained from TRUS and MR scans. We begin by briefly discussing the RBF interpolation method.

1) Radial Basis Function (RBF) Interpolation

The RBF interpolation approach, first used for medical 3D modeling by Carr et al. [22], [26], approximates an input function of unknown form with three variables \( f: \mathbb{R}^3 \rightarrow \mathbb{R} \) by an interpolation function \( s: \mathbb{R}^3 \rightarrow \mathbb{R} \) given the values \( \{f(x_i), i = 1, 2, ..., n\} \). Here, \( x_i = (x_i, y_i, z_i) \), is a set of distinct input points in the 3D space \( \mathbb{R}^3 \) called nodes of interpolation or radial centers. Also, an evaluation point \( x = (x, y, z) \) represents a point at which the function \( f \) is to be approximated. The following form of the interpolation function is considered:

\[
s(x) = p_1(x) + \sum_{i=1}^{n} \lambda_i \phi(||x - x_i||), \quad x \in \mathbb{R}^3, \quad \lambda_i \in \mathbb{R} \quad (1)
\]

where \( p_1 \) is a first order polynomial, \( \phi(r) = r \) for bi-harmonic splines, and \( ||\cdot|| \) refers to the Euclidean norm. Let \( \pi_3^n \) represent the space of all first-order polynomials with three variables. The coefficients \( \lambda_i \) in (1) are determined by satisfying the interpolation conditions (2) and the side conditions (3):

\[
s(x_i) = f(x_i), \quad i = 1, 2, ..., n \quad (2)
\]

\[
\sum_{i=1}^{n} \lambda_i q(x_i) = 0, \quad \text{for all} \quad q \in \pi_3^n \quad (3)
\]

Equations (2) and (3) can be put into the following matrix form for algorithm implementation [26]:

\[
\begin{pmatrix}
A & p \\
p^T & 0
\end{pmatrix}
\begin{pmatrix}
\lambda \\
0
\end{pmatrix}
= \begin{pmatrix}
f \\
0
\end{pmatrix}
\quad (4)
\]

where

\[
A = (a_{ij}) = \left( \phi(||x_i - x_j||) \right)
= ||x_i - x_j||, \quad i, j = 1, 2, ..., n
\quad (5)
\]

\[
p = \begin{bmatrix}
1 & x_1 & y_1 & z_1 \\
1 & x_2 & y_2 & z_2 \\
\vdots & \vdots & \vdots & \vdots \\
1 & x_n & y_n & z_n
\end{bmatrix}
\quad (6)
\]

\[
\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)^T
\quad (7)
\]

\[
c = (c_1, c_2, c_3, c_4)^T
\quad (8)
\]

\[
p_1(x) = c_1 + c_2 x + c_3 y + c_4 z
\quad (9)
\]

Solving (4) yields values of \( \lambda_i \) for each radial center and coefficients \( (c_1, c_2, c_3, c_4) \) for \( p_1 \). In this way, the value of \( f \) at the evaluation point \( x \) can be estimated from (1). When applied to the surface interpolation problem, \( f(x_i) \), which represent the input function values at the provided points, are assigned to two groups of points: on-surface points and off-surface points. On-surface points (existing data points) are assigned the value 0 while off-surface points, which are artificially introduced along the surface normal in order to avoid the trivial solution \( f = 0 \) everywhere, are assigned non-
zero values. Typically, as suggested in [21], [26], [27], two layers of off-surface points, one on each side of the surface, are added. These points are given a positive value if on the outer surface and a negative value if on the inner surface. If at an evaluation point \( x \), \( s(x) \) is evaluated to be 0 then this point belongs to the surface.

Directly solving the linear system (4) is not practical as the matrix becomes very ill-conditioned as the number of interpolation nodes increases. Therefore, the RBF algorithm is combined with the partition of unity algorithm [27]-[29]. The idea is that RBF can be applied to interpolation problems involving a large number of data points by breaking down the large-scale interpolation problem to several small-scale interpolation problems. After solving (4) with the partition of unity algorithm, a mesh will be created in the entire volume where the interpolation is planned and, at these evaluation points, the interpolation function \( s(x) \) is evaluated according to (1). Finally, MATLAB’s isosurface algorithm is called and the surface, which is defined by all points carrying a value of 0, is rendered in 3D.

2) Radial Basis Function (RBF) Interpolation

Given that the purpose of this study is to investigate the differences in shape that the two imaging modalities introduce, a rigid-body registration accompanied the RBF interpolation. Shape is defined as a property that is invariant to similarity transformations [15] including translation, rotation, and scaling. In this case, the same prostate was not represented in the same coordinate system in the two imaging modalities; the TRUS probe and MR machine have their own coordinate systems. Thus, translations and rotations introduced during imaging must be eliminated in order for us to be able to compare pure shapes. Sometimes, the scaling is of interest in shape comparison [30], as is in our case since we want to investigate the differences in prostate volume introduced by different imaging modalities. Therefore, the scaling factor was not eliminated at this stage. On the next step, after obtaining the resampled prostate contour points in Section II.A, the CoM of each 3D volume was determined. Then, each 3D volume was translated so that its CoM was placed at the position \((0, 0, 0)\), eliminating the translation component. As for the rotation component, based on a modified version of the principal axis transformation [31] as described next, each prostate was rotated such that the cranial-caudal axes of all prostates were aligned to the same line in a uniform coordinate system. To this end, first a moment of inertia tensor was calculated for each prostate assuming the mass of each slice was concentrated at its CoM. The principal axis calculated from the inertia tensor pointed along the longitudinal direction of the prostate, thus defining the cranial-caudal axis. Then, all 2D contours of each prostate were pivoted as the same solid around the corresponding prostate CoM so that the cranial-caudal axis of the prostate coincided with the Z axis of the uniform coordinate system (Fig. 2). After this step, the orientation of the axial plane (perpendicular to the cranial-caudal axis) of the prostate is not yet the same across all prostates. To give every prostate model the same orientation in the axial plane, another principal axis analysis was carried out on the central three slices of each prostate, which best represent the orientation of the axial plane of the prostate. The second component of this analysis was aligned with the Y axis of the new uniform coordinate system (Fig. 3). This rotation sets the prostate “straight” on the X-Y plane. Note that although in Fig 2 and Fig. 3 the surface was fitted to the resampled contour points to better demonstrate these transformations, we only have the resampled points at this stage.

Subsequently the resampled and aligned contour points were stacked together in the same coordinate system (Fig. 4(a)). Next, off-surface points were added (Fig. 4(b)). Then, the RBF surface fitting algorithm was applied (Fig. 4(c)).
At the end of RBF interpolation aided with rigid-body registration, each prostate was represented by 30 evenly spaced slices with each slice containing points resampled at 6 degree intervals placed radially with respect to the CoM of each slice. Therefore, a total of 1,800 points were placed on each prostate model from either imaging modality.

The above-discussed interpolation process is crucial to building a statistical shape model and to the statistical shape analysis in the next subsection as it allows for representing each prostate with a point distribution model (PDM); this process has now allocated an equal number of points evenly distributed in the z direction and across each x-y plane to each prostate in each imaging modality. Thus, we did not perform the RBF interpolation aided with rigid-body registration for mere visualization purposes; rather, it has been used to generate methodically-placed and consistent contour points for each prostate in spite of the differences in orientation and slice gaps of the original manually-placed contour points. This processing greatly increases the data density along the z axis, making the subsequent 3D shape analysis feasible.

C. Implementation of Statistical Shape Models (SSM)

The SSM is a multivariate statistical approach to model data with a point distribution model (PDM) and analyze their variations with principal component analysis (PCA) [32]. It is often combined with image search algorithms to perform image segmentation [9], [16]. The details of implementation are given below.

In the Section II.B, a PDM was constructed for each prostate of both imaging modalities by using the RBF interpolation with rigid-body registration. The 1,800 points allocated to each prostate are regarded as corresponding landmarks. Normally, landmarks include points marking parts of the object with particular application-dependent significance such as the center of urethra in a prostate US scan, or points marking application-independent items such as the highest point of a prostate model, or other points that can be interpolated from points of the previous two types [32]. In medical applications, landmarks may be segmented manually by clinical experts but this procedure can be tedious and time consuming. Thus, in our work, the geometric correspondence between landmarks was decided automatically: points at the same radial position (in the XY plane) of 2D slices on the same length proportion (Z axis) of each prostate were considered to correspond to each other. A similar approach can be found in [20], in which the so-called solid angles were used to locate corresponding landmarks; points on the same solid angle from the center of prostate were assumed to correspond to each other. However, in our application, due to the lack of sample points at the base and at the apex of the prostate caused by the 2D nature of our imaging modalities, the prostate could not be put in such a spherical coordinate system. Instead, we adopted a cylindrical coordinate system and determined the correspondence of the landmarks through a similar parameterization. The scaling, directly associated with volume, was preserved at this stage because, as explained in Section II.B prostate volume provides an interesting parameter to compare between US and MR imaging modalities.

After placing corresponding landmarks and aligning, we obtained the mean shape of all k prostates (with k =40) for each imaging modality by

$$\bar{x} = \frac{1}{k} \sum_{i=1}^{k} x_i$$  \hspace{1cm} (10)

with $x_i = (x_1, y_1, z_1, ..., x_j, y_j, z_j)^T$, which is the shape vector of each prostate shape, constructed from coordinates of the j = 1,800 landmark points.

Next, an eigen decomposition on the corresponding covariance matrix delivered 3j, principal modes of variation $\phi_m$ (eigenvectors) and their respective variances $\lambda_m$ (eigenvalues). Then, it was possible to approximate every valid shape by a weighted linear combination of the first c modes:

$$x_{estimate} = \bar{x} + Pb$$  \hspace{1cm} (11)

with $P= [\phi_1, \phi_2, ..., \phi_c]$ being the matrix composed of the first c eigenvectors and $b=[b_1, b_2, ..., b_c]^T$ being the weight vector. The value of c was chosen so that the accumulated variance $\sum_{m=1}^{c} \lambda_m$ reached a certain ratio $r$ of the total variance $\sum_{m=1}^{3j} \lambda_m$. Commonly used values for $r$ are between 0.9 and 0.98.

III. RESULTS

A. Average Shape Comparison

After performing RBF interpolation coupled with rigid-body transformation and by applying (10) on the 40 groups of contour data from each imaging modality, the average shapes for both US and MR were obtained. The resulting PDMs obtained from MR and US data are displayed together in Fig. 5; both models were translated so that their CoM were located at (0, 0, 0) in the same coordinate system. Four geometrical features were compared: average volume, average length, average height, and average width. For each feature, the ratio of the value corresponding to the US data to the value corresponding to the MR data was calculated as well. Length, height and width refer to the longest possible distance between any two points in the Z, Y, and X directions, respectively. The same coordinate system as defined in Section II.A was used for model construction. The paired Student’s t-test was applied to test for statistical significance of the four measures. The tests were two sided by default and the p-value threshold was set to be 0.05. The results are displayed in Table II.

It can be observed that the volume of the average US shape model is smaller than that of the MR average model by 1.4 ml. However, the difference in average volumes is not statistically significant. On the other hand, the US average model is longer than that of the MR average model by 2.1 mm, whereas in terms of width and height, the MR average model appears to be bigger than the US average model by 1.5 mm and 0.97 mm,
respectively. Although the differences in width, height and length are not statistically significant, the p value for length is considerably smaller than for the other two, implying a greater difference in length between US and MR average models than in width and in height. This appears to be logical as the ends of prostates in both imaging modalities are particularly difficult to determine, thus introducing a greater error. The US/MR average volume ratio is 0.96.

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>GEOMETRY COMPARISON BETWEEN AVERAGE US AND MR SHAPE MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Volume</td>
</tr>
<tr>
<td>US</td>
<td>36.5 ml</td>
</tr>
<tr>
<td>MR</td>
<td>37.9 ml</td>
</tr>
<tr>
<td>p value</td>
<td>0.55</td>
</tr>
<tr>
<td>US/MR ratio</td>
<td>0.96</td>
</tr>
</tbody>
</table>

The two average shape models were then compared by using a color map to illustrate shape differences at more local levels; Fig. 6 shows apex-anterior, base-anterior, and posterior views. The Euclidean distances from each landmark to its model CoM, which is located at (0,0,0) for both imaging modalities, were first calculated. Then, for each pair of corresponding landmarks, the distances calculated were compared and the differences were mapped onto the color map. The comparison was carried out with the MR model as the reference. Therefore, red, which represents positive values, indicates areas where the US model is larger compared to the MR model. Blue, which represents negative values, indicates areas where the US model is smaller compared to the MR model.

From the apex-anterior view, it can be concluded that the US average shape is longer and flatter than the MR average shape. From the apex-anterior view and the base-anterior view, it can be observed that near the base and the apex of the prostate, US contours are generally larger than MR contours particularly on the anterior side, reaching a maximum difference of 2 mm. In the posterior areas, however, MR and US average models do not show significant differences.

B. Principal Component Analysis (PCA) of Shape Vectors

By continuing to apply the analysis procedure presented in Section II.C, a PCA decomposition was performed on the two sets of shape vectors. The first 15 eigen modes of each model were retained. While for the US shape model the first 15 modes represented 97.2% of its variance, for the MR shape model the first 15 modes represented 95.1% of its variance. The first five modes of each model are visualized in Fig. 7 in order to provide a more direct impression of what these modes represent. In order to visualize the influence of each mode on the shape model, the following operation was performed [16], [32]:

$$\mathbf{x} = \bar{\mathbf{x}} \pm 2\sqrt{\lambda_i} \mathbf{\phi}_i \quad \text{with } i = 1,2,...,5$$

In (12), $\lambda_i$ is the variance for each mode $i$ and $\mathbf{\phi}_i$ is the $i^{th}$ eigenvector. The original average shape is plotted in blue, $\mathbf{x} = \bar{\mathbf{x}} + 2\sqrt{\lambda_i} \mathbf{\phi}_i$ is plotted in red, and $\mathbf{x} = \bar{\mathbf{x}} - 2\sqrt{\lambda_i} \mathbf{\phi}_i$ is plotted in yellow.

In the US model (Fig. 7), the first mode represents variations in the diameter of the prostate shape in the XY plane, while the second mode represents the variations in the length along the X axis of the prostate. Modes 3, 4, and 5 appear to represent more localized effects on the shape: mode 3 and mode 5 introduce variations at the base and the apex of the prostate while mode 4 acts directly on the posterior surface of the prostate. It is interesting to note that with mode 4, the prostate model is able to exhibit different degrees of drooping; the presumed effect of insertion of the ultrasound probe into the rectum is well captured by mode 4.
In the MR model (Fig. 8) the first two modes appear to be less decoupled than those in the US model. Mode 1 appears to represent a combination of variations in length along the Z axis and a rotation around the X axis. Mode 2 seems to represent a combination of variations in diameter in the XY plane and also a rotation around the X axis. Modes 3, 4, and 5 show greater similarities with the results from the US model: modes 3 and 5 represent local variations close to the base and the apex of the prostate, while mode 4 once again corresponds to the different degrees of drooping. Since in MR imaging no probe was inserted in the rectum, the results in Fig. 8 imply that prostates naturally have different degrees of drooping.

C. In-group and Cross-group SSM Comparison Using the Generalization Measure

In the next two sections, the US SSM and the MR SSM are compared and evaluated using two measures, the generalization measure and the specificity measure, to examine their interchangeability. In order to take into account the inter-observer variability, a total of three oncologists were involved in the segmentation process on the same set of US/MR images. By doing this, we will see what happens if we use data from one imaging modality to build a SSM and then use it to recreate the shapes from another imaging modality. Therefore, four cases are considered: (1) use US contour data for SSM construction and approximate the shapes obtained from US imaging modality; (2) use MR contour data for SSM construction and approximate the shapes obtained from MR imaging modality; (3) use US contour data for SSM construction and approximate the shapes obtained from MR imaging modality; (4) use MR contour data for SSM construction and approximate the shapes obtained from US imaging modality.

In order to quantify and compare the performance of the two SSMs in question in their respective scenarios, we used the previously-mentioned two measures, which were first introduced in [33] as metrics to compare shape models obtained from different registration and alignment methods. Here, we use them to see how predictive and pertinent the shape models obtained from US and MR imaging modalities are, both within their own population group and within the other population group. First, the generalization measure was used, which measures the ability of a SSM model to provide a good fit under unseen circumstances. It follows a 'leave-one-out' approach where s-1 samples of a total of s samples are used to produce a SSM while the remaining sample is used as a test sample. From (11), we have

\[ \mathbf{b} = \mathbf{P}^T (\mathbf{x} - \bar{x}) \]  

(13)

as \( \mathbf{P} \) is orthonormal [16]. The estimate of the left-out shape is calculated with (11), while the vector \( \mathbf{b} \) is calculated in (13) and \( \bar{x} \) and \( \mathbf{P} \) are calculated from the s-1 samples. The vector \( \mathbf{x} \) in (13) represents the shape vector of the left-out sample. The average error between the estimate of the left-out shape and the true left-out shape over the entire set of trials is quantified by the measure G as

\[ G = \frac{1}{M} \sum_{i=1}^{M} \epsilon_{Gl}^2 \]  

(14)

and

\[ \epsilon_{Gl}^2 = |\mathbf{x}_1 - \mathbf{x}'_i|^2 \]  

(15)

where \( M \) is the total number of shapes (40 in this study), \( \mathbf{x}_1 \) the left-out training shape, and \( \mathbf{x}'_i \) is its estimate by the SSM built from the rest of the population, and \( \epsilon_{Gl}^2 \) represents the sum-of-squares of approximation errors. An SSM that has a lower G value is better in terms of approximating unseen shapes whereas an SSM that has a higher G value implies less capability of taking on unseen shapes [33].
Therefore, if we regard the entire 40 prostate samples as one population, the s-1 samples can come from either the US modality or the MR modality, and the same is true for the left-out sample. As a result, the aforementioned four scenarios were produced in order to see how the SSM of US and MR shapes were validated both in-group and cross-group. The measure G results from the three oncologists’ data are given in Table III, in which Model Imaging Modality denotes the imaging modality from which the s-1 samples were taken and Test Imaging Modality denotes the imaging modality from which the test sample (the left-out sample) was taken. The first 15 modes of each model were used.

<table>
<thead>
<tr>
<th>Case</th>
<th>Model Imaging Modality</th>
<th>Test Imaging Modality</th>
<th>G Measure Observer 1</th>
<th>G Measure Observer 2</th>
<th>G Measure Observer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>US</td>
<td>US</td>
<td>1.7</td>
<td>2.3</td>
<td>2.7</td>
</tr>
<tr>
<td>2</td>
<td>MR</td>
<td>MR</td>
<td>2.1</td>
<td>2.9</td>
<td>3.6</td>
</tr>
<tr>
<td>3</td>
<td>US</td>
<td>MR</td>
<td>2.5</td>
<td>3.6</td>
<td>4.2</td>
</tr>
<tr>
<td>4</td>
<td>MR</td>
<td>US</td>
<td>2.2</td>
<td>3.5</td>
<td>3.7</td>
</tr>
</tbody>
</table>

The results show that US models (Case 1) perform better than MR models (Case 2) in-group, with a smaller average G value across all three observers. This is not surprising as we showed in Section III.B that the first two US modes are much more decoupled than the first two MR modes. When it comes to cross-group approximation, MR models perform slightly better than US models. The MR-US (Case 4) averages are inferior to that of the US-MR (Case 3) consistently across all three observers.

Next, the squared errors $e_{G_{i}}^{2}$ in (15) for each prostate are broken down in Fig. 9, Fig. 10 and Fig. 11 for all four combinations. For brevity, we are only showing data from one of the observers, which is representative. The X axis represents the prostate number and the Y axis represents the corresponding $e_{G_{i}}^{2}$ values.

As displayed in Fig. 9 the MR prostate shapes have several samples that are associated with much higher $e_{G_{i}}^{2}$ values. The samples that present the highest $e_{G_{i}}^{2}$ values include prostates number 4, 17, 24 and 39. Although suspected to be outliers at first, we reperformed the segmentation of the prostate samples in question and the results were unchanged. Therefore in the MR imaging modality there exists shapes that are harder to be modeled by the obtained MR SSM. In the case of testing US samples with both US and MR models, the US and MR SSM perform similarly (Fig. 10). They both have difficulties estimating a certain number of shapes. In the case of testing MR shapes with both US and MR models, interesting similarities can be found (Fig. 11). The $e_{G_{i}}^{2}$ values of these two scenarios follow each other much more closely across prostates than the previous case where the US shapes were tested as samples. Overall, while the US SSM model performs better than MR in-group, they have comparable results in cross-group modeling. Cross-group results also appear to be reasonable when compared with in-group results for both modalities (Case 1 compared to Case 4 and Case 2 compared to Case 3 respectively in Table III).

### D. In-group and Cross-group SSM Comparison Using the Specificity Measure

The specificity measure quantifies the ability of a SSM to generate only valid shapes [33]. The specificity value is evaluated by generating a test set comprised of a large number of random shapes by applying (11) while restricting each parameter $h_i$ of the vector $b$ to follow a normal distribution with zero mean and $\lambda_i$ standard deviation. The squared error $e_{S_{i}}^{2}$ between each generated shape and the closest match of the training set is then calculated according to

$$e_{S_{i}}^{2} = |\mathbf{x}_i - \mathbf{x}'_i|^2$$

And the specific value $S$ is the average of all squared errors:

$$S = \frac{1}{N} \sum_{i=1}^{N} e_{S_{i}}^{2}$$

In (16) and (17), $\mathbf{x}'_i$ is the random shape and $\mathbf{x}_i$ its closest shape in the training set. $N$, the number of random shapes generated, is set to 10,000 in this study.

As with the generalization measure, four scenarios were taken into account for both in-group and cross-group approximation: US training set and US test set; MR training set and MR test set; US training set and MR test set; and MR training set and US test set. With 15 modes, all combinations showed stabilized S measures and the corresponding results are shown in Table IV.

<table>
<thead>
<tr>
<th>Case</th>
<th>Model Imaging Modality</th>
<th>Test Imaging Modality</th>
<th>S Measure Observer 1</th>
<th>S Measure Observer 2</th>
<th>S Measure Observer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>US</td>
<td>US</td>
<td>2.5</td>
<td>3.6</td>
<td>4.2</td>
</tr>
<tr>
<td>2</td>
<td>MR</td>
<td>MR</td>
<td>3.5</td>
<td>6.4</td>
<td>5.5</td>
</tr>
<tr>
<td>3</td>
<td>US</td>
<td>MR</td>
<td>4.5</td>
<td>8.8</td>
<td>6.4</td>
</tr>
<tr>
<td>4</td>
<td>MR</td>
<td>US</td>
<td>3.0</td>
<td>7.3</td>
<td>7.0</td>
</tr>
</tbody>
</table>
As can be observed from Table IV, US SSM shows a better ability to generate valid shapes within group, as the $S$ values in the first row are consistently smaller than those in the second row. This may be due to the MR SSM’s less decoupled modes as shown in Section III.B. In comparison to the obvious superiority of US in-group $S$ measure over that of MR, two out of three observers show smaller differences in cross-group comparisons: differences between US and MR for observer 2 and 3 are $-2.8 \times 10^3$ mm$^2$ and $-1.3 \times 10^3$ mm$^2$ respectively in in-group performance, while the cross-group differences are 1.5$\times 10^4$ mm$^2$ and -0.6$\times 10^3$ mm$^2$ respectively. This shows that the cross-group performance between the two modalities is more comparable.

IV. DISCUSSION

A. Prostate Volume and Shape

The volume ratio and geometric characteristics observed in this study are comparable to the results found in the literature. Whether in [34], [35], where pre-implant US and post-implant MR contours were compared, or in [20], where post-implant US and post-implant MR contours were compared, the US/MR ratio was reported to be relatively stable, varying from 0.90 to 0.95, and is comparable to the ratio of 0.96 found in this study. In the aforementioned studies, however, degraded imaging conditions for both US and MR were reported due to implanted seeds [20], which did not exist in our study. As we only used pre-implant images for contouring, prostate edema, which could have influenced the results reported in studies involving post-implant images, did not have any influence.

In terms of the prostate shape, similar results were obtained in this study as were found in [20] by Smith et al. In their study, by comparing post-implant US contours obtained from a 3D TRUS device and post-implant MR contours, the US contours were reported to be longer and flatter than the MR contours by 1.1 mm and 2 mm, respectively. In our study, the US average model exceeded the MR average model in length and height by 2.1 mm and 1 mm, respectively. In both studies, the US contours were found to be larger than the MR contours near the base on the sides and the anterior surface of the gland. The relative flatness and increase in length of the US contours can be explained by the insertion of the TRUS probe, which presses against the rectum and the prostate. The difference in
shape can also be attributed to the different imaging requirements of US and MR, since standard MR bore could not accommodate a patient in the lithotomy position, which is required for TRUS imaging. This introduces therefore different patient positioning, and thus difference in prostate orientation and shape [19]. Although there is a three-week gap between US and MR scan, we expect the influence of rectal and bladder filling difference to be minimal, as all patients were instructed to have a comfortably full bladder and empty rectum before each scan session.

As far as the indentation is concerned, Smith et al. found very prominent deformation of the inferior half of the prostate in the US contours compared to the MR contours, which was not observed in our study. This difference is most likely associated with differences in insertion angle and y-axis positioning of the US probe in the rectum. Also, Smith et al. showed that the undeformed prostate imaged by MR was convex. This convex shape can be observed in our Fig. 5 and, more importantly, the natural drooping effect was made evident in our PCA analysis on MR shape vectors in Fig. 8, where different degrees of indentation were created with the fourth mode.

B. US and MR Statistical Shape Models

In this study, a new method was used for landmark placement (RBF interpolation combined with rigid-body registration) in order to go from 2D slices to 3D shapes. RBF were used especially to overcome several obstacles encountered in SSM building: insufficient sampling along the z-axis, different inter-slice distances between the US and MR modalities, and different number of points representing each prostate shape. Hodge et al. [16] used a 3D TRUS imaging system to directly obtain volumetric images based on which contouring was performed. In this way, contours that are consistently and equally distributed in space were obtained. When it comes to the correspondence of points located on corresponding slices, Hodge et al. tested several labeling methods. In our study, one of the tested methods (equally spaced rotational samples) was applied, which had been shown to provide accurate results and easy implementation. Instead of breaking 3D shapes or image volumes into corresponding slices and then into corresponding points for labeling as was done in [16] and the present study, other approaches include the Iterative Closest Point (ICP) algorithm introduced by Besl and McKay [36] and the Softassign Procrustes introduced by Rangarajan et al. [37], which operate directly in 3D space and do not require equal numbers of points between two models. The method used in this study takes into consideration the physical meaning behind numerical models by aligning cranial-caudal axis of each prostate to the same axis in a uniform coordinate system and adjusting the orientation of the axial plane. For the purpose of the current study, it was found sufficient to apply the same correspondence and labeling procedure to both imaging modality groups, despite the numerical errors (assumed to be identical for both imaging modalities) that may be introduced during this processing. Nonetheless, other correspondence methods can be applied in future studies [33] to see whether the results in the current study exhibit a dependence on correspondence and labeling methods.

By the application of the generalization measure and the specificity measure on both SSMs, we found that the US SSM performed better than the MR SSM in both measures, particularly in-group and slightly cross-group. However the cross-group performance is comparable between the two modalities and remain reasonable when compared to in-group G and S measures, indicating interchangeability between the two imaging modalities. The prostate SSM has traditionally been created from the US imaging modality [9], [16]; therefore there does not exist, to the authors’ knowledge, MR prostate SSM in the literature with which we could compare our results. However, MR images have been used for SSM building for brain structures [38], [39] and cardiac structures [17], [23] where they have proven effective. The inferior model quality of the MR SSM could be due to the fact that the oncologists participating in this study have more experience dealing with US images. This may result in the findings in [19], where the intra-observer variability is found to be lower for US modality in mid-gland slices as well as the prostate volume, the better consistency of which can give the US SSM a more robust performance. Also the difference in imaging resolution (0.18 mm for US images and 0.44 mm for MR images in both x and y directions) is advantageous for the US imaging modality. Finally, it could also be due to the correspondence/labeling method applied. As shown in [16], different labeling methods do influence the G and S values of one particular SSM, but since we fixed the labeling method for both imaging modalities, the next step of research can be to study whether they are equally influenced by different correspondence and labeling methods. Another improvement that can be made in future studies is to use contouring results obtained from the same oncologists involved in this study but over a certain number of repetitions, in order to take into account intra-observer effects.

C. SSM Methodology and Independent Component Analysis (ICA)

The SSM was chosen to represent prostate shapes and as a tool for shape comparison partly because of its popularity and success to be combined with image search algorithms for contour segmentations. Since the objective of our study is to provide insight into the feasibility of coupled US-MR segmentation methods based on SSM, it makes sense to compare the SSM shape models constructed from the data pertaining to these two modalities. Also, as can be seen from the previous analysis, the SSM can not only be used to compare prostate shapes across different imaging modalities, but also to visualize how the shapes vary within the same imaging modality through PCA eigenvectors. The RBF interpolation preceding the SSM modeling is a necessary step. More than a simple visualization tool, it brings all prostate shapes to a unified 1800 data point representation and increases the data sample density. However, this approach is not without its limitations. The PCA associated with
conventional SSM building is designed to address linear (or slightly non-linear) data and assumes a Gaussian distribution of the data points. Therefore in this study, we also assume the data points obtained from both imaging modalities are normally distributed. Another limitation is that PCA returns eigenvectors that represent global deformations [40]. Since more localized deformations would be likely to be included in eigenshapes describing less variance, they are likely to be ignored by PCA. However, PCA is able to rank the order of dominant eigenshape components, which is relevant for our analysis as they show the most dominant deformations that exist in both US and MR prostate shapes.

On the other hand, ICA, originally developed to separate mixed audio signals into independent sources, is getting more and more popular in areas such as feature extraction, signal processing and image processing [41]. It has the potential to be combined with SSM [40] and provide an alternative method to obtain deformation components. It does not need the assumption of Gaussian distribution of data, and it is able to exploit higher-order moments of the statistical distributions. If used in shape analysis, ICA is able to capture better local deformations compared with PCA, as it maximizes the independence between the non-Gaussian component vectors [40]. However, the standard ICA algorithm does not sort the components in the order of variance. An appropriate ordering scheme has to be carefully investigated before ICA can be applied to shape models for dimension reduction, which is not in the scope of this current study. However, ICA and PCA can be potentially combined to provide both global and local deformation information. It is an option we will try to explore to improve our MR SSM results in future studies. Since compared with US PCA results, the MR PCA results show less decoupled principal modes, we have a reason to suspect that the MR shape in this study might not follow a Gaussian distribution. Therefore, our next step will include applying ICA to the MR data set to see if better decoupled components can be obtained.

V. CONCLUSIONS

In this study, TRUS and MR pre-implant prostate scans for 40 patients were segmented by three experienced oncologists. The 2D contours from both imaging modalities were interpolated into 3D shapes by combining radial basis function interpolation and principal axis rigid registration. The placement of corresponding landmarks was accomplished by geometric registration. By comparing the mean shapes of the resultant US and MR statistical shape models, geometric characteristics consistent with those reported in prostate implant literature were found. The effect of the ultrasound probe on prostate drooping was not readily apparent from inspecting the mean shape. While we expected the prostate drooping to be solely due to the ultrasound probe insertion, the principal component analysis on the MR data suggests that prostates naturally possess varying degrees of drooping. Therefore the drooping demonstrated in US shapes is a combined effect of ultrasound probe insertion and prostate’s natural drooping. Statistical models based on MR and US contours were tested using the generalization and specificity measures. It was found that, in this group of contours, US SSM models perform in general better than MR SSM models. However, we believe that there is relatively good compatibility between US and MR imaging modalities in terms of statistical shape modeling, due to the comparable generalization measure (G) and specificity measure (S) during cross-group analysis. Another reason is since drooping naturally exists in prostates, it can be taken into account by the MR SSM. In future studies, different correspondence / labeling methods can be used and oncologists with more MR image segmentation experience can be recruited in order to further investigate why the US model and MR model had different performances regarding the used metrics in this study. The possibility of incorporating ICA into the shape analysis together with PCA for better SSM construction will also be explored.

REFERENCES


