Published in final edited form as:

Med Image Anal. 2017 July ; 39: 29–43. doi:10.1016/j.media.2017.04.001.

Learning Non-rigid Deformations for Robust, Constrained Point-based Registration in Image-Guided MR-TRUS Prostate Intervention

John A. Onofrey^{a,*}, Lawrence H. Staib^{a,b,c}, Saradwata Sarkar^e, Rajesh Venkataraman^e, Cayce B. Nawaf^d, Preston C. Sprenkle^d, and Xenophon Papademetris^{a,c}

^aDepartment of Radiology & Biomedical Imaging, Yale University, New Haven, Connecticut, USA

^bDepartment of Electrical Engineering, Yale University, New Haven, Connecticut, USA

^cDepartment of Biomedical Engineering, Yale University, New Haven, Connecticut, USA

^dDepartment of Urology, Yale University, New Haven, Connecticut, USA

^eEigen, Grass Valley, California, USA

Abstract

Accurate and robust non-rigid registration of pre-procedure magnetic resonance (MR) imaging to intra-procedure trans-rectal ultrasound (TRUS) is critical for image-guided biopsies of prostate cancer. Prostate cancer is one of the most prevalent forms of cancer and the second leading cause of cancer-related death in men in the United States. TRUS-guided biopsy is the current clinical standard for prostate cancer diagnosis and assessment. State-of-the-art, clinical MR-TRUS image fusion relies upon semi-automated segmentations of the prostate in both the MR and the TRUS images to perform non-rigid surface-based registration of the gland. Segmentation of the prostate in TRUS imaging is itself a challenging task and prone to high variability. These segmentation errors can lead to poor registration and subsequently poor localization of biopsy targets, which may result in false-negative cancer detection. In this paper, we present a non-rigid surface registration approach to MR-TRUS fusion based on a statistical deformation model (SDM) of intra-procedural deformations derived from clinical training data. Synthetic validation experiments quantifying registration volume of interest overlaps of the PI-RADS parcellation standard and tests using clinical landmark data demonstrate that our use of an SDM for registration, with median target registration error of 2.98 mm, is significantly more accurate than the current clinical method. Furthermore, we show that the low-dimensional SDM registration results are robust to segmentation errors that are not uncommon in clinical TRUS data.

Graphical abstract

*Corresponding Author (Telephone number: +1-203-737-2835).

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.

Keywords

non-rigid registration; statistical deformation model; dimensionality reduction; prostate; magnetic resonance imaging; trans-rectal ultrasound; image-guided biopsy

1. Introduction

Accurate and robust non-rigid image registration is critical for image-guided interventions. Non-rigid registration is a challenging task that is further complicated if the fusion involves images of different modalities. Intensity-based registration algorithms must deal with complex intensity relationships, which makes finding the optimal solution difficult. As an alternative to maximizing an intensity-based similarity metric, such as mutual information, anatomical structures of interest may first be segmented from the images, and then be registered to each other to perform the image fusion. However, image segmentation itself remains a challenging problem. Segmentations are prone to errors and to high variability, especially when these segmentations are performed under the limited time constraints of a clinical procedure. In this paper, we present a non-rigid registration solution that is both accurate and robust to potentially inaccurate segmentations. Our approach utilizes a statistical model of non-rigid deformation due to the specific image-guided intervention learned from a large set of clinical training data to effectively constrain the registration process. We demonstrate results applying our approach to non-rigidly register prostate surfaces segmented from pre-procedure magnetic resonance (MR) imaging (MRI) and intra-procedure trans-rectal ultrasound (TRUS) imaging for image-guided prostate cancer biopsy localization.

Prostate cancer is one of the most commonly occurring forms of cancer and one of the major causes of cancer-related death in the U.S. It is estimated that over 180,000 men will be diagnosed with prostate cancer in the year 2016, and over 26,000 men will die of the disease (Siegel et al., 2016). Biopsy is the current clinical standard for diagnosing prostate cancer, and the standard procedure in the U.S. utilizes TRUS imaging for biopsy guidance. TRUS imaging itself cannot be reliably used for targeting suspicious lesions because of poor image quality and lack of contrast. The current biopsy sampling technique takes 12 tissue cores, in a systematic, non-patient-specific plan, from different regions of the prostate. This systematic survey is extremely limited in spatial coverage of the prostate gland, which results in severe under-sampling. Using this general approach to biopsy sampling has been likened to a game of chance with no more than 0.45% of an average-sized gland being biopsied (Ukimura et al., 2012). In fact, the statistics compiled from systematic ultrasound-guided biopsies are staggering: (i) as much as 30% of serious tumors may be missed on a

first-time biopsy (Taira et al., 2010); (ii) the cancer detection rate for patients with a prior negative biopsy is lower than 20% and the detection of clinically significant cancers is even lower (Roehl et al., 2002); (iii) biopsy pathology underestimates significance of disease 40% of the time leading to incorrect staging (Isariyawongse et al., 2008); and (iv) approximately 50% of detected tumors are indolent (Cooperberg et al., 2007).

Cancerous lesions are, however, more readily seen using other imaging modalities, such as MRI. Compared to TRUS, multi-parametric MRI (mpMRI) that combines T2-weighted imaging with functional sequences, *e.g.* diffusion-weighted MRI, spectroscopic MRI and dynamic contrast enhanced MRI, shows significantly better sensitivity and specificity for staging, localizing and detecting prostate cancer (Barentsz et al., 2012; Villers et al., 2006; Kirkham et al., 2006; Kim et al., 2007). However, performing biopsies under intra-procedure MRI guidance can be time consuming, expensive and impractical (Ukimura et al., 2012; Pondman et al., 2008). To avoid these problems, current practice aims to fuse pre-procedure mpMRI with intra-procedure TRUS imaging. Here, suspicious prostatic tissue is first identified via mpMRI and then TRUS imaging is used to provide targeted image-guided navigation for biopsy. A rigid image fusion method between the MRI and TRUS is inadequate for accurate biopsy guidance due to prostate gland deformation caused by: (i) variations in patient orientation (Moradi et al., 2012); (ii) changes in bladder volume or rectal filling; (iii) the presence or absence of an endorectal MR coil (Alterovitz et al., 2006); and (iv) deformation caused by hand-held TRUS probes (Krücker et al., 2007). Non-rigid registration is therefore necessary to compensate for these deformations.

1.1. MR-TRUS Image Fusion Methods

A variety of intensity-based and segmentation-based methods have been presented to non-rigidly fuse pre-procedure MRI and intra-procedure TRUS imaging. Mitra *et al.* (Mitra et al., 2012b) presented an intensity-based registration based on image textures, but validated it only on 2D midland MR-TRUS slices, while Sun *et al.* (Sun et al., 2013) presented 3D non-rigid MR-TRUS registration using local intensity descriptors. Intensity-based registration is difficult due to TRUS's poor signal-to-noise ratio and the complex intensity relationship between TRUS and T2-weighted MRI. Moreover, corresponding anatomical features may not be visible in both modalities due to differences in the imaging physics, which violates a typical requirement of intensity-based registration methods for matching structures to exist in both images.

Segmentation-based registration methods, on the other hand, fuse the MRI and TRUS images by first identifying a set of specific fiducial landmarks or the whole prostate gland, and then registering these structures. Because automatic prostate segmentation is itself a challenging task, the majority of MR-TRUS fusion methods rely on either manual or semi-automated segmentation in both the MR and TRUS images. Manually localized fiducials have been used for rigid (Xu et al., 2007; Bublely et al., 2013), affine (Kaplan et al., 2002), and non-rigid (Mitra et al., 2012b; Dickinson et al., 2013) alignment. Some fusion methods perform an initial alignment using a limited number of fiducials for global prostate shape alignment and then perform a non-rigid registration thereafter to correct for deformations (Xu et al., 2008; Cool et al., 2011; Hu et al., 2012, 2015). In contrast to the limited number

of points used by fiducial registration methods, which cannot be used to accurately account for non-rigid deformations, whole gland segmentations capture the entire prostate and provide a high number of boundary points that may be used to estimate non-rigid surface deformations. These whole gland segmentations may be represented either as image-based segmentations or as surfaces. Moradi *et al.* (Moradi et al., 2012) non-rigidly register manually segmented prostate binary image masks, while Fedorov *et al.* (Fedorov et al., 2015) non-rigidly register signed distance maps of manually segmented prostates. Surface-based non-rigid registration methods treat the segmentations as points in 2D slice contours (Reynier et al., 2004; Cool et al., 2011; Mitra et al., 2012a; Zettinig et al., 2015), as points in 3D surfaces (Narayanan et al., 2009; Karnik et al., 2010; Natarajan et al., 2011; Fedorov et al., 2015; Khallaghi et al., 2015b,a; van de Ven et al., 2015; Onofrey et al., 2015b), or as a set of basis functions, for example as spherical harmonics (Moradi et al., 2012). The state-of-the-art, clinical Artemis prostate biopsy system and its ProFuse Bx multimodal image fusion software (Eigen, Grass Valley, CA), relies on intra-procedure, semi-automated TRUS segmentation (Ladak et al., 2000) and a deformable surface registration algorithm to align the prostate surfaces segmented in both the MRI and TRUS images (Narayanan et al., 2009).

Alternative fusion approaches make use of MR prostate segmentation, which may be performed manually prior to the biopsy procedure free of clinical time constraints, but forgo manual segmentation of the TRUS image. Starting with a manual MR segmentation, these methods perform MR-TRUS fusion by (i) using the manual MR prostate segmentation to automatically segment the prostate in the TRUS image, which implicitly defines the deformation (Martin et al., 2010), (ii) performing model-to-image-based registration to the TRUS image (Hu et al., 2012, 2015), (iii) maximizing the alignment of a manual MR prostate segmentation to an automatically segmented TRUS prostate probability map image (Sparks et al., 2015), or (iv) automatically segmenting the TRUS prostate first and then performing surface-based registration (Zettinig et al., 2015). Yet another approach to MR-TRUS fusion involves acquisition of a separate TRUS volume prior to the biopsy procedure. This TRUS image serves as a reference to which the MRI is first registered; thereafter, the MR-TRUS fusion is updated with respect to this reference volume using intensity-based TRUS registrations of intra-procedure images (Xu et al., 2008; Cool et al., 2011). While Karnik *et al.* (Karnik et al., 2010) demonstrated that such intensity-based, mono-modal TRUS-TRUS registration performed better than surface-based registration of prostates segmented from these images, these methods make certain assumptions about how the prostate deforms between the TRUS reference acquisition and intra-procedure imaging. Xu *et al.* (Xu et al., 2008) assumes limited deformation between the MRI and the TRUS reference by performing an affine registration with updates using only 2D TRUS slices, whereas Cool *et al.* (Cool et al., 2011) perform non-rigid fusion between the MRI and TRUS reference (both images acquired on the same day) but register the intra-procedure TRUS to the TRUS reference using only rigid registration, which assumes no deformations between the two TRUS acquisitions.

Segmentation-based registration methods involving some level of manual interaction, such as the ones mentioned above, are time-consuming, are highly operator dependent, and are a significant source of variability (Smith et al., 2007; Karnik et al., 2010). Surface-based registration algorithms such as robust point matching (RPM) (Rangarajan et al., 1997; Chui

and Rangarajan, 2003) and coherent point drift (CPD) (Myronenko and Song, 2010) attempt to handle noisy surface data by allowing for soft point correspondences and outliers in the data. MR-TRUS fusion methods have made use of RPM (Khallaghi et al., 2015b; Onofrey et al., 2015b) and CPD (Fedorov et al., 2015; Khallaghi et al., 2015b,a; Zettinig et al., 2015). These methods have been shown to be robust to noisy segmentations with synthetically added perturbations (Onofrey et al., 2015b; Khallaghi et al., 2015a) and to partial segmentations missing large sections of the prostate (Fedorov et al., 2015; Khallaghi et al., 2015b,a).

1.2. Modeling Prostate Deformation

Prior knowledge about either (i) prostate gland tissue mechanics or (ii) statistics of prostate shape and gland deformation has been incorporated into MR-TRUS fusion methods. To address the former, including biomechanical computational models into the image registration process can constrain the non-rigid deformation to be physically realistic. A general criticism of surface-based registration methods is that they can only reliably register data that is close to the surface itself since there is no information provided to the registration algorithm inside the shape boundary far from the surface. Biomechanical models address this limitation by acting as a regularization constraint on the non-rigid transformation (Moradi et al., 2012; Hu et al., 2012, 2015; Fedorov et al., 2015; Khallaghi et al., 2015b,a; van de Ven et al., 2015). These models treat the prostate as a linear elastic material, and require tuned parameters to best approximate the elastic tissue properties of the gland. While biomechanical models may be used to realistically interpolate the deformation within a surface, the model cannot be used to drive the registration process itself since the model requires the specification of boundary constraints (deformation forces) to the model's surface. These forces are typically found using some form of surface-based registration method or alignment with image-based features, both of which are themselves sensitive to segmentations errors.

The other form of prior knowledge used by MR-TRUS fusion methods involves statistical shape models (SSMs) and statistical deformation models (SDMs), both of which attempt to handle noisy measurements by learning the distribution of either plausible prostate shape or plausible prostate deformation from a population of examples. SSMs attempt to model the variation of shape within a particular class of objects (Heimann and Meinzer, 2009). For example, active shape models use a principal component analysis (PCA) of landmark point displacements to constrain realizations of an object's shape to be from the distribution of the probabilistic model (Cootes et al., 1995). SSMs rely upon landmark points to describe the object's shape, however, manual identification of corresponding landmarks is expensive and automatic segmentation is itself a challenging task. SDMs, on the other hand, avoid the explicit segmentation of the object of interest, and instead model how an object deforms according to a deformation field. For example, an SDM may use a PCA of B-spline control point displacements in free-form deformations to learn the distribution of these non-rigid transformations (Rueckert et al., 2003). In contrast with biomechanical models, SSMs and SDMs can be used to directly drive the registration process.

SSMs and SDMs have been shown to be promising for MR-TRUS fusion (Yan et al., 2010; Hu et al., 2012, 2015; Khallaghi et al., 2015a). In particular, Hu *et al.* (Hu et al., 2012, 2015) train a patient-specific deformation model using simulated finite element mesh deformations generated from manual segmentation of the prostate and surrounding anatomy, while Khallaghi *et al.* (Khallaghi et al., 2015a) use an SSM to estimate prostate shapes given segmentations with missing data. In general, PCA-based models suffer from the curse of dimensionality where the number of training samples is limited compared to the higher dimensionality of the shape or the deformation field. For this reason, Hu *et al.* (Hu et al., 2012, 2015) simulated a large number of possible prostate deformations to generate synthetic training samples. However, from our own work (Onofrey et al., 2015a), we have seen that even low-quality training data, provided in large enough quantity, is useful for modeling non-rigid deformations in a population.

1.3. Contributions

We present a method to directly improve the MRI to TRUS fusion procedure in clinical prostate biopsy procedures. This procedure currently relies on the heavily operator-dependent, semi-automated segmentation of the prostate from the ultrasound image that is performed by the clinician during the biopsy. Our new fusion method uses a training set of over 100 MR-TRUS images to learn the space of prostate deformations so as to improve both the accuracy of the method and the robustness of the clinician's workflow. Our method, described in Sec. 2, builds a statistical deformation model (SDM) of typical prostate deformations in the context of biopsy (caused primarily by the insertion of the trans-rectal probe and the force this exerts on the prostate). We perform a principal component analysis (PCA) of intra-subject non-rigid deformations between pre-procedure MRI and intra-procedure TRUS imaging to build a statistical deformation model (SDM) of procedure-induced non-rigid deformations (Onofrey et al., 2016) from a training set of MRI-TRUS image-guided procedures derived from a population of clinical cases. While PCA has difficulty modeling high-dimensional data, we aim to capture the gross, intra-subject deformations observed between MRI and TRUS image acquisitions with our SDM. We adopt a robust point matching (RPM) framework (Rangarajan et al., 1997) to perform PCA-based SDM registration of MRI and TRUS prostate surfaces, and we provide a review of RPM in Appendix A. We emphasize the key advantages of this approach: (i) RPM does not require homologous points between surfaces like SSMs (Heimann and Meinzer, 2009); (ii) RPM is a robust method that can discard non-matching points as outliers; (iii) the SDM will constrain the deformation to be realistic based on the actual deformations encountered during the TRUS procedure as observed from a large population of training data; and (iv) the SDM is a low-dimensional parameterization that allows for highly efficient computation that fits within the time constraints of TRUS-guided biopsy procedures (typically 15 minutes). In Sec. 3, we present results that expand upon our previous work (Onofrey et al., 2015b) to include both rigorous validation testing using synthetic data with realistic segmentation errors and quantitative landmark target registration error tests with clinical data. We conclude, in Sec. 4, that our results demonstrate that our approach robustly registers surfaces in the presence of inaccurate prostate segmentations to provide effective patient-specific, non-rigid image fusion.

2. Methods

We train our statistical deformation model (SDM) using data from a clinical database of $N=104$ patients who underwent prostate biopsy at our institution. Fig. 1 outlines our method and workflow. For each patient $i = 1, \dots, N$ in the database, we have a pre-procedure T2-weighted MRI $I_{MR,i}$ and an intra-procedure TRUS image $I_{TRUS,i}$. The MR images were acquired at our institution without the use of an endorectal coil. The 3D TRUS images were created using the clinical Artemis system (Eigen, Grass Valley, CA), which reconstructs a 3D image volume from a systematic series of 2D B-mode images, which are acquired by a manual sweep but the probe is stabilized on the arm and rotates around its long axis. Clinical experts manually segment the prostate in each image according to standard clinical protocols using a semi-automated segmentation tool (Eigen, Grass Valley, CA). This process involves first identifying four or more points on the prostate boundary in 2D slices to initialize a contour that is then automatically refined according to image features (Ladak et al., 2000), and finally adjusted manually, if necessary. We denote these segmented prostate surfaces $S_{MR,i}$ and $S_{TRUS,i}$. In this work, we model the prostate surface S as a discrete set of points within a triangulated mesh. First, Sec. 2.1 details how we create a template prostate shape that will serve as a reference for our SDM. Sec. 2.2 then presents our method for creating an SDM of procedure-induced deformation. Finally, Sec. 2.3 describes how we use this deformation model to non-rigidly register MR and TRUS surfaces.

2.1. Creating a Template Prostate Shape

In order to create a model of intra-procedure prostate deformation from a population of training data, we first normalize our training data for differences in prostate volume by registering all subjects to a common reference space by creating a template prostate shape S_{SDM} . We create this template prostate shape using the segmented MR surfaces $S_{MR,i}$ instead of the TRUS segmentations because delineation of the gland boundary is much less ambiguous in MRI. First, we select a single subject to serve as the prostate template reference space. The whole gland volume distribution of the $N=104$ segmented MR prostates in this database is $55 \pm 37 \text{ cm}^3$ (mean $\pm$ SD), with median 46 cm^3 , minimum 17 cm^3 , and maximum 262 cm^3 . The chosen subject's prostate had a volume of 32 cm^3 , which was approximately equal to the mode of the population's distribution. The remaining $N-1$ subjects are then linearly registered to this single reference prostate using a robust point matching (RPM) surface registration algorithm (Rangarajan et al., 1997) with a similarity transform that allows for rigid translation, rotation, and isotropic scaling. Our use of the similarity transform mitigates prostate shape bias, as individual prostate shapes will be preserved while normalizing for prostate shape. For each of the $N-1$ registered MR surfaces, we then create a binary image mask of the prostate in the template reference space. We average the $N-1$ masks in the reference space to create a probability image of the prostate shape from which we extract the 0.5 probability surface to be used as the template prostate shape S_{SDM} for the SDM. This template prostate had a volume of 34 cm^3 . To avoid bias introduced by this reference subject, we exclude this subject from the SDM training procedure that follows (we use the remaining $N-1$ subjects). This template surface S_{SDM} serves as the reference space $\Omega_{SDM} \subset \mathbb{R}^3$ in which we construct our SDM.

2.2. Training the Non-rigid Statistical Deformation Model

To model the distribution of non-rigid deformations between the pre-procedure MR prostate surfaces and the intra-procedure TRUS prostate surfaces as a result of the biopsy procedure, we start by aligning S_{MR} with S_{TRUS} in the template prostate reference space. For each subject $i = 1, \dots, N - 1$ in the training dataset, we first perform intra-subject registration to rigidly align $S_{TRUS,i}$ to that subject's corresponding pre-procedure prostate $S_{MR,i}$ using standard rigid RPM (Rangarajan et al., 1997). We denote this transformation $T_{TRUS \rightarrow MR,i}$ where $p \rightarrow q$ indicates a rigid transformation from space p to q . At this point, the TRUS surfaces are rigidly registered to their MR surfaces, and any remaining non-rigid deformations between the two are accounted for after normalization to the prostate template in template space. To this end, we non-rigidly register $S_{MR,i}$ to S_{SDM} with standard non-rigid RPM surface-based registration (Chui and Rangarajan, 2003; Papademetris et al., 2003). These registrations, which we denote $T_{MR \rightsquigarrow SDM,i}$ where $p \rightsquigarrow q$ indicates non-rigid transformation from space p to q , use a free-form deformation (FFD) transformation model parameterized by 5.0 mm isotropic B-spline control point spacing (Rueckert et al., 1999). We then warp both $S_{MR,i}$ and $S_{TRUS,i}$ to the SDM reference space by concatenating and applying the transformations:

$$S'_{MR,i} = S_{MR,i} \circ T_{MR \rightsquigarrow SDM,i}, \quad (1)$$

$$S'_{TRUS,i} = S_{TRUS,i} \circ T_{TRUS \rightarrow MR,i} \circ T_{MR \rightsquigarrow SDM,i}, \quad (2)$$

where $\circ$ is the transformation operator. At this point, $S'_{MR,i}$ and $S'_{TRUS,i}$ are rigidly registered to each other within the SDM reference space, and non-rigid deformations remain between the two.

Making the assumption that these deformations are the result of introduction of the ultrasound probe for the biopsy procedure and changes in patient positioning with respect to the pre-procedure MRI, we now construct our SDM of procedure-induced deformation across the training population. To capture these non-rigid deformations, we non-rigidly register $S'_{MR,i}$ to $S'_{TRUS,i}$ using non-rigid RPM (Papademetris et al., 2003) with 10.0 mm isotropic FFD control point spacing. We choose a 10.0 mm control point spacing here, instead of the higher-resolution 5.0 mm used to normalize $S_{MR,i}$ to S_{SDM} , because we are more interested in capturing the gross prostate deformation characteristics than with exactly matching the surfaces, which most likely contain segmentation errors. In the fixed SDM reference space, this FFD parameterization results in a $6 \times 5 \times 5$ FFD grid with $d = 450$ degrees of freedom in 3D, and we concatenate these control point displacements into a single vector $\mathbf{d}_i \in \mathbb{R}^d$. The $N - 1$ transformations $T_{MR \rightsquigarrow TRUS,i}$ yield a training sample set of deformations $\mathbf{d}_i$, which are realizations of the distribution of deformations $\mathcal{D}$ between the pre-procedure MRI and intra-procedure TRUS imaging as a result of the procedure. We approximate the distribution of deformations $\mathcal{D}$ by performing a principal component

analysis (PCA) of the FFD transformations (Rueckert et al., 2003), where PCA estimates the eigenvectors of $\mathcal{D}$'s covariance matrix

$$\mathbf{d} = \bar{\mathbf{d}} + \Phi \mathbf{w} \quad (3)$$

where $\bar{\mathbf{d}} = \frac{1}{N-1} \sum_{i=1}^{N-1} \mathbf{d}_i$ is the mean deformation of the training registrations, $\Phi = (\phi_1 | \dots | \phi_J) \in \mathbb{R}^{d \times J}$ is the matrix of orthogonal eigenvectors, and $\mathbf{w} \in \mathbb{R}^J$ is a vector of model coefficients. The eigenvectors $\phi_1, \dots, \phi_J$ are the J principal modes of intra-patient deformation due to the procedure. The number of training samples and the transformation's dimensionality determine the number of eigenvectors such that $J = \min\{N-1, d\}$, where typically $N-1 \ll d$. The j -th eigenvalue λ_j estimates the sample variance along the eigenvector ϕ_j and we sort them in decreasing order $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_J$. An SDM using J_v parameters accounts for $0 \leq v \leq 100\%$ of the model's cumulative variance

$\sum_{j=1}^{J_v} \lambda_j \leq v \sum_{j=1}^J \lambda_j$. Thus, using the first J_v coefficients of $\mathbf{w}$ in (3) provides a low-dimensional parameterization (J_v DoFs) of a higher-dimensional FFD $\mathbf{d}$, which we denote $T_{SDM_v}(\mathbf{x}; \mathbf{w})$ for all points $\mathbf{x}$ in the reference image domain Ω_{SDM} . Fig. 2 illustrates some example deformations generated by our SDM.

2.3. Surface-based Statistical Deformation Model Registration

Given a new pre-procedure MR and intra-procedure TRUS image pair for a patient not in our training set, we use the learned SDM to drive the non-rigid registration process of the segmented prostate S_{MR} to the segmented intra-procedure prostate S_{TRUS} . Following the method described in Sec. 2.2 to register and then warp the training data to the prostate template space, we (i) rigidly register S_{TRUS} to S_{MR} using linear RPM, (ii) non-rigidly register S_{MR} to the reference template S_{SDM} using standard non-rigid RPM, and then (iii) transform the two surfaces into SDM space using (1) and (2). With the surfaces transformed to the SDM reference space, our proposed method non-rigidly registers S'_{MR} to S'_{TRUS} by constraining the non-rigid RPM registration framework to deform according to the SDM transformation. Rather than estimate an FFD deformation by optimizing the control point locations individually as is usually done, the SDM's weights $\mathbf{w}$ in (3) offer a low-dimensional parameterization of a dense FFD non-rigid transformation $T_{SDM}(\mathbf{x}, \mathbf{w})$, $\forall \mathbf{x} \in \Omega_{SDM}$. For each iteration k of the annealing procedure, our SDM RPM method estimates the eigenvector weights $\hat{\mathbf{w}}$ in (3) to minimize the cost function in (A.1):

$$T_{SDM}^k(\cdot, \hat{\mathbf{w}}) = \arg \min_{\mathbf{w}} \sum_i c_i \|T_{SDM}(\mathbf{x}_i, \mathbf{w}) - \mathbf{x}'_i\|^2 + f(t) R(T_{SDM}), \quad (4)$$

where $\mathbf{x}_i \in S'_{MR}$ and $\mathbf{x}'_i$ are the estimated RPM point correspondences from points $y_i \in S'_{TRUS}$, $c_i \in \mathbb{R}$ is the confidence of each correspondence determined by the RPM estimate of point i 's "outlierness", and $R(T)$ is a regularization function weighted by a scalar

function $f(t) \in \mathbb{R}$. While the SDM's eigenvectors in (3) are orthogonal, their linear combination does not necessarily guarantee parameterization of a smooth FFD transformation. We therefore include the regularization term to enforce smoothness constraints onto the transformation estimate by penalizing the parameterized FFD's total bending energy (Rueckert et al., 1999). Additionally, we bound the SDM transformation to have weights within 3 standard deviations of the mean value, *i.e.* $|w_j| \leq 3\sqrt{\lambda_j}$ where λ_j is the eigenvalue corresponding to the j -th eigenvector. We solve this optimization problem using conjugate gradient descent at each annealing step.

Thus far, we have estimated and applied the transformation T_{SDM} within the prostate template space Ω_{SDM} . To fuse the MR data to the TRUS data within the native TRUS image space Ω_{TRUS} in which the urologist is viewing the biopsy procedure, we concatenate the following transformations to align S_{MR} to S_{TRUS} :

$$\hat{S}_{\text{MR}} = S_{\text{MR}} \circ T_{\text{MR} \rightsquigarrow \text{SDM}} \circ T_{\text{SDM}} \circ T_{\text{SDM} \rightsquigarrow \text{MR}} \circ T_{\text{MR} \rightarrow \text{TRUS}} \quad (5)$$

where $T_{\text{SDM} \rightsquigarrow \text{MR}} \approx T_{\text{MR} \rightsquigarrow \text{SDM}}^{-1}$ and $T_{\text{MR} \rightarrow \text{TRUS}} = T_{\text{TRUS} \rightarrow \text{MR}}^{-1}$ invert the transformations in (2). In the case of $T_{\text{SDM} \rightsquigarrow \text{MR}}$, the inverse of the FFD transformation does not have a closed-form solution and we instead approximate its solution. For the remainder of this paper, we refer to this entire chain of transformations interchangeably with T_{SDM} .

3. Results

We tested the performance of our surface registration algorithm using both synthetic deformation tests (Sec. 3.2) and tests using real clinical data with manually identified landmarks (Sec. 3.3). These landmarks consisted of cysts and other “cyst-like” anatomical structures appearing in both the MR and TRUS images of 22 subjects. Sec. 3.1 details our methods for evaluating these experiments. We compared our non-rigid statistical deformation model (SDM) registration approach against surface-based robust point matching (RPM) registrations (Rangarajan et al., 1997; Chui and Rangarajan, 2003; Papademetris et al., 2003) with the following transformation models: (i) rigid transformation, (ii) affine transformation, and (iii) standard, unconstrained, non-rigid free-form deformation (FFD) transformation. Parameters for the RPM annealing process were held constant across all registration methods, with initial and final annealing temperatures $t_{\text{init}} = 2.0$ and $t_{\text{final}} = 0.5$, respectively. During the registration's annealing process, we gradually reduced the regularization parameter from its initial value $f(t_{\text{init}}) = 1.0$ down to a final value $f(t_{\text{final}}) = 0.1$ in order to slowly relax the transformation's smoothness constraint as the registration progressed. For the non-rigid RPM registrations, we used the same 10.0 mm isotropic FFD control point spacing used to parameterize the SDM in Sec. 2.2. Additionally, for the clinical landmark tests in Sec. 3.3, we compared our results to two other methods: (i) the adaptive focus deformable model (AFDM) non-rigid surface-based registration algorithm (Shen et al., 2001) currently implemented in the clinical Artemis system (Narayanan et al., 2009); and (ii) a distance map-based registration similar to the one presented by Fedorov *et al.* (Fedorov et al., 2015). For the distance map registration

experiments, we tested rigid, affine and non-rigid FFD transformations, where the FFD control point spacing was set to be the same as the RPM and SDM methods. We computed the distance maps over isotropically resampled versions of the MR and TRUS image volumes with spacing 1.0 mm^3 and 0.5 mm^3 , respectively. With the exception of the AFDM method, we implemented all registrations algorithms in C++ as part of BioImage Suite (Joshi et al., 2011), and performed all tests on a 4 GHz Intel Core i7 iMac workstation.

3.1. Evaluation Methods

We evaluated registration performance using two different methods: (i) volume of interest (VOI) overlap in our synthetic tests (Sec. 3.2), and (ii) landmark target registration error (TRE) in our clinical data tests (Sec. 3.3). We detail each of these methods in the following sections.

3.1.1. Evaluating Synthetic Data—The most reliable way to evaluate registration performance is to calculate the overlap of small volumes of interest (VOIs) (Rohlfing, 2012). For the prostate, we make use of the clinical Prostate Imaging Reporting and Data System (PI-RADS) (Weinreb et al., 2016) parcellation standard to delineate 37 clinically-relevant anatomical VOIs within the prostate (Fig. 3). An expert rater manually segmented all 37 PI-RADS VOIs in a single subject's MR image and we denote this atlas $A_{\text{PI-RADS}}$. We then propagated this atlas to a test subject's MR image space $I_{\text{MR},i}$ by non-rigidly registering the whole prostate gland contour in $A_{\text{PI-RADS}}$ to $S_{\text{MR},i}$ using FFD RPM to generate the subject-specific PI-RADS parcellation $A_{\text{PI-RADS},i}$. While the propagated VOIs do not correspond exactly to each subject's unique prostate anatomy because they were registered with a surface-based registration algorithm that ignores internal prostate anatomical information, they visually appeared qualitatively close and serve their purpose as a parcellation with anatomical significance for synthetic data generation in our evaluation. To evaluate registration accuracy, we calculated the Dice overlap (DO) for each of the $v = 1, \dots, 37$ PI-RADS VOIs:

$$\text{DO}_v(A, B) = 2 \frac{|A_v \cap B_v|}{|A_v| + |B_v|}, \quad (6)$$

where X_v denotes the v -th VOI from image X . We also calculated the mean Dice overlap (MDO) for each registration result by averaging over all VOIs, $\sum_v \text{DO}_v$, which has been used as a summary statistic to quantify overall registration performance rather than analyze each individual VOI Dice overlap separately (Klein et al., 2009). We scale DO values by 100 to aid in interpretation.

3.1.2. Evaluating Clinical Data—From our database of $N = 104$ TRUS biopsy subjects, an expert manually identified a total of 29 corresponding landmarks in 22 subject MR-TRUS image pairs (maximum number of landmarks per subject was 2). These landmarks, which consisted of cysts and “cyst-like” anatomical features, were distributed throughout the prostate gland across the 22 subjects (Fig. 4 and Table 1), with 20 landmarks localized in the midgland and 9 in the base, the majority of which were in the transition zone. Rather than

localize the landmarks using a single point, the expert segmented the cysts by manually painting a 3D volume of interest (VOI) mask using BioImage Suite (Joshi et al., 2011). For each landmark VOI_v , $v \in \{1, 2\}$ in each modality $m \in \{\text{MR}, \text{TRUS}\}$ in these $s = 1, \dots, 22$ subjects, we then computed the VOI center of gravity $\mathbf{c}_{s,v,m} \in \Omega_m \subset \mathbb{R}^3$. To compute the fiducial localization error (FLE) of these landmarks, a single rater repeated this cyst painting localization procedure twice (due to the time-consuming nature of this protocol, only 10 of the 22 subjects were segmented twice) and we then averaged the two landmark center of gravity estimates. We calculated the distribution of FLE (mean $\pm$ SD) to be 0.07 ± 0.03 mm and 0.40 ± 0.72 mm for the MR and TRUS landmarks, respectively. We then calculated the landmark target registration error (TRE) (West et al., 1997)

$$\text{TRE}_{s,v} = \|\mathbf{c}_{s,v,\text{TRUS}} - T(\mathbf{c}_{s,v,\text{MR}})\| \quad (7)$$

for the given transformation T .

3.2. Validation Using Synthetic Data

We leveraged our large database to generate a set of realistic, synthetic TRUS prostate shapes with known non-rigid deformations and known segmentation errors on which we can test different registration approaches. For these validation tests, we partitioned our database into two sets: (i) a set of 50 subjects as training data, and (ii) a set of 50 subjects for testing. Using these two sets, we followed the method in Sec. 2.2 to create two independent SDMs, one for training and one for testing, which we denote $T_{\text{SDM}}^{\text{train}}$ and $T_{\text{SDM}}^{\text{test}}$, respectively. We selected the same single subject to be used as a prostate template reference shape S_{SDM} for both SDMs using the procedure from Sec. 2.1.

We selected 10 subjects from the training set as example prostates, and we index these subjects $s = 1, \dots, 10$. Because prostate gland segmentation is so much more reliable within the MR images, we use the MR prostate segmentations $S_{\text{MR},s}$ for these examples. For each subject s , we generated $t = 1, \dots, 100$ synthetic deformations using the training SDM by randomly sampling weights $\mathbf{w}$ for the model in (3). We sampled SDM deformation weight values $w_j \sim \mathcal{T}(0, \lambda_j)$, where λ_j is the j -th eigenvalue, using the first $j = 1, \dots, 5$ eigenvectors in (3) to create the deformation $T_{\text{SDM},s,t}^{\text{train}}$. Using (1) and (5), where in (5) $T_{\text{MR} \rightarrow \text{TRUS}}$ is set to be the identity transform since we do not have to transform to TRUS image space with these synthetic examples, we create a total of 1,000 synthetically deformed MR prostate surfaces $\hat{S}_{\text{MR},s,t}$ that realistically mimic the intra-procedure deformations observed in our training data. These synthetic surfaces and their respective transformations function as our synthetic ground-truth since they are not subject to potential errors introduced by the segmentation of the prostate in TRUS imaging.

To simulate the variability of the TRUS segmentation process, we first calculated the spatial distribution of TRUS segmentation variability shown in Fig. 5 (Appendix B), which we then used to create a spatial model of TRUS segmentation variance (Appendix C). As described in the Appendices, this model required a TRUS template prostate shape S_{SDM} in order to

calculate segmentation variability across different subjects. Thus, we first warped this surface to each $S_{MR,s}$ using a non-rigid FFD RPM transformation (Chui and Rangarajan, 2003; Papademetris et al., 2003), and used this warped version of the template $S_{TRUS,s}$ as our TRUS examples. We then applied the same sample SDM deformations $T_{SDM,s,t}^{train}$ to these surfaces as was done for the MR surfaces using (5) to create $\hat{S}_{TRUS,s,t}$. Finally, we added synthetic segmentation noise to each $\hat{S}_{TRUS,s,t}$ based on the segmentation variability model in Appendix C. To do this, we randomly sampled values $u_j \sim \mathcal{T}(0, \sigma v_j)$, where v_j is the j -th eigenvalue, using the first $j = 1, \dots, 10$ modes in (C.1) with a noise level scalar $\sigma = \{0.25, 0.5, 1.0, 2.0, 4.0\}$. A noise level $\sigma = 1.0$ represents the standard amount of segmentation variance observed in the population (assuming a Gaussian normal distribution), while noise levels $\sigma < 1.0$ and $\sigma > 1.0$ represent lower than and higher than observed levels of segmentation variance, respectively. As $\sigma \rightarrow 0$, the synthetic examples approach the variance-free segmentation case, *i.e.* minimal artificial segmentation errors. Thus, for each noise level σ , we created 1,000 synthetic TRUS shapes $\hat{S}_{TRUS,s,t,\sigma}$ which resulted in 5,000 synthetic samples in total. Fig. 6 shows some example prostate shapes with both synthetic deformations and synthetic segmentation errors.

Our registration tests compared the performance fusing $S_{MR,s}$ to $\hat{S}_{TRUS,s,t,\sigma}$ using a variety of surface-based registration methods: (i) rigid RPM (Rangarajan et al., 1997), (ii) affine RPM, (iii) non-rigid FFD RPM (Papademetris et al., 2003), and (iv) our SDM RPM method. Here, the FFD method serves as a surrogate for the AFDM method since we were unable to test the synthetic examples within the existing clinical device itself. For our SDM method, we used the SDM learned from the testing data T_{SDM}^{test} , and we compared results using $J = \{1, 3, 6, 11, 16, 50\}$ eigenvectors, which accounted for 25, 50, 75, 90, 95, and 100% of the model's cumulative variance, respectively. We denote transformations from these test methods T_r with $r = \{\text{Rigid RPM, Affine RPM, Non-rigid FFD RPM, SDM}\}$.

We evaluated registration performance by calculating the Dice overlap of PI-RADS parcellation VOIs. Given the PI-RADS parcellation for each test subject $A_{PIRADS,s}$ (Sec.

3.1.1), we used the known synthetic deformations $T_{SDM,s,t}^{train}$ generated above to we warped this atlas to create gold-standard target atlas images $\hat{A}_{PIRADS,s,t}$. Since we added segmentation error noise to simulate TRUS segmentation error after the synthetic deformation was applied, we only required a total of 1,000 $\hat{A}_{PIRADS,s,t}$ atlases to serve as ground-truths, irrespective of the segmentation error noise level added (these atlases represent ground-truth parcellations in the absence of segmentation error). We then calculated the Dice overlap for each VOI v as $DO_v(T_r \circ \hat{A}_{PIRADS,s,t}, A_{PIRADS,s})$ in (6) as well as the mean Dice overlap (MDO), $\Sigma_v DO_v$.

Fig. 7 shows the distribution of MDO results for each level of synthetic noise tested. We used a Wilcoxon rank-sum test with significance level 0.05 to test for significant differences in MDO between each pairwise comparison of the registration methods. For the standard noise level $\sigma = 1.0$, all registration methods were found to have statistically significant differences in MDO ($p < 0.05$), with the exception of pairwise comparisons between the SDM $J = 11, 16, 50$ methods, which showed no significant differences. The SDMs with $J > 1$ all had significantly higher median MDO than the other RPM registration methods. Rigid

RPM registration performed significantly worst (median MDO = 61.1), while affine RPM (median MDO = 74.4) significantly outperformed non-rigid FFD RPM (median MDO = 70.2) and SDM 1 (median MDO = 73.1). The SDM methods showed much better worst-case performance with minimum MDO = 46.7 for all SDMs J , compared to minimum MDO values 25.9, 37.2, and 23.6 for Rigid, Affine, and FFD, respectively. At the lowest level of noise ($\sigma = 0.25$), affine registration significantly outperformed both rigid and non-rigid FFD registration, as well as the SDMs with $J = 1, 3$. The SDMs with $J = 6$, on the other hand, had significantly higher MDO than affine registration at low noise. All methods showed significant decrease in MDO as noise increased from one level to the next (with the exception of the rigid and SDM 1 methods from $\sigma = 0.25$ to $\sigma = 0.5$), performance of the SDM degraded gracefully with increasing levels of noise. As the noise level increased, affine registration performance approached that of both rigid and non-rigid FFD registration, while the SDM performance decreased at a slower rate. Affine registration median MDO performance decreased by 27.1% from 79.1 at $\sigma = 0.25$ to 57.7 at $\sigma = 4.0$, while SDM 11 performance decreased by 16.7% from 82.2 at $\sigma = 0.25$ to 68.5 at $\sigma = 4.0$. Furthermore, the small interquartile range values demonstrate that the SDM methods show more consistent and robust registration performance across all noise levels.

We can further break the MDO results down into their constituent VOIs to examine registration performance spatially. Fig. 8 shows individual VOI DO results averaged across the population for each level of noise tested and for each of the tested methods (we only show SDM 11 since this method had the highest MDO results and used the fewest number of eigenvectors while showing no statistically significant differences between SDM 16, 50 at standard levels of noise $\sigma = 1.0$ in Fig. 7) (see Supplementary material Fig. S1 for individual VOI DO distribution values). Qualitative inspection of the VOI results show the SDM method with the highest DO performance throughout the prostate regions at all noise levels. Affine registration had slightly lower DO performance than the SDM at low noise levels, most notably in the peripheral zone (PZ) at the base. Unlike the affine transformation, the SDM maintained excellent registration performance in the transition zone (TZ) across all noise levels. Only at the highest levels of noise did the SDM performance start to degrade in the PZ and anterior fibromuscular stroma (AFS). Rigid registration, on the other hand, shows poor registration performance across the base, midgland, and apex, with particularly poor results in the base PZ and apex AFS. While the non-rigid FFD registration performed better than the rigid registration method at lower noise levels (higher DO values throughout the prostate for $\sigma \leq 1.0$), the unconstrained FFD method's performance approached that of rigid registration at high levels of noise. Furthermore, at low noise levels, the non-rigid FFD showed particularly poor performance in the base PZ and AFS, areas of the prostate which are most difficult to segment in the TRUS images and subject to the most variability (Fig. 5).

Table 2 summarizes these individual VOI overlap results by prostate regions. For all registration methods, the midgland exhibited the highest mean DO values. While affine registration had the highest mean DO in the apex, the SDM 11 had the highest mean DO values in both the midgland and base as well as the smallest standard deviation values. Rigid registration performed worst across all prostate regions.

Running times for the SDM 11 were 3.3 ± 0.4 seconds compared to 12.0 ± 1.7 seconds (mean $\pm$ SD) for the FFD across the 1000 synthetic examples with $\sigma = 1.0$, a 73% reduction in processing time on average. This SDM running time excludes the time for registering the MR surface to the SDM template shape because this task may be performed offline prior to any biopsy procedure and the time for the final warping procedure in (5) is negligible.

3.3. MR-TRUS Landmark Registration Error with Clinical Data

3.3.1. SDM Registration—To test our SDM method, we performed leave-one-out testing by holding test subject s out from the training dataset, and then computed the SDM as in Sec. 2.2 with the remaining subjects. Instead of generating a new SDM template prostate shape S_{SDM} for each leave-one-out test, we used the same one created in Sec. 2.2 since the removal of a single subject from the training data had negligible effect on its creation. Using the 29 landmarks identified in the $s = 1, \dots, 22$ subjects (Sec. 3.1.2), we calculated the distribution of landmark target registration error (TRE) using (7).

We tested SDM registration using the first $J = \{1, 2, 3, 4, 5, 17, 35, 102\}$ eigenvectors, which accounted for 28, 49, 61, 71, 75, 95, 99, and 100% of the model's cumulative variance, respectively. Fig. 9 summarizes the TRE distribution results. Using a Wilcoxon rank-sum test with a significance level of 0.05 to test for significant differences in TRE between each of the eight SDM methods, we found no significant differences for all values of J . The SDM using $J = 17$ eigenvectors had the lowest median TRE of 2.98 mm of the test values. All methods had approximately equal standard deviations of TRE (~ 2 mm). Maximum TRE decreased with increasing number of eigenvectors used.

3.3.2. Comparison to Other Registration Methods—We compared our SDM results (using $J = 17$) to the seven following registration methods: (i) the non-rigid adaptive focus deformation model (AFDM) (Shen et al., 2001); RPM surface-based registration using (ii) rigid (Rangarajan et al., 1997), (iii) affine, and (iv) non-rigid FFD (Papademetris et al., 2003) transformations; and distance map registration using (v) rigid, (vi) affine, and (vii) non-rigid FFD (Fedorov et al., 2015) transformations. We summarize the TRE results for each method in Fig. 9 and Table 3.

The AFDM algorithm had TRE 6.11 ± 4.75 mm (mean $\pm$ SD) with median, minimum, and maximum TRE 4.24 mm, 0.68 mm, and 19.75 mm, respectively. In terms of median TRE performance, the AFDM method had the highest median error while the SDM when using $J = 17$ eigenvectors had the lowest median TRE of 2.98 mm, a reduction of 30% in median TRE across the population. The results also showed that AFDM is highly sensitive to incorrectly segmented prostate glands in the TRUS images. In this case, the registration algorithm aligns the MR and TRUS prostate surfaces to a fault (maximum TRE was 19.75 mm), whereas all the RPM and distance map registration methods had much lower maximum TRE performance, which indicates lower sensitivity to errors in the segmentation process. The standard, non-rigid FFD RPM method showed the lowest maximum TRE performance (6.61 mm) of the tested non-rigid registration methods, while the non-rigid distance map method had a maximum TRE of 9.77 mm, compared to SDM 17's maximum TRE of 8.69 mm. Rigid RPM achieved the lowest minimum TRE value of 0.18 mm. While

the surface-based RPM methods outperformed their distance map registration counterparts in terms of median TRE, the distance map methods exhibited lower standard deviation values. We used a Wilcoxon rank-sum test with a significance level of 0.05 to test for significant differences in TRE between each of the registration methods. Both the non-rigid FFD RPM method and our SDM using $J = 17$ eigenvectors had significantly lower TRE ($p < 0.05$) than AFDM. The RPM FFD and rigid distance map method results were also significantly different. There were no statistically significant differences between all other pairwise registration method comparisons. The running time for the SDM using 17 eigenvectors was 3.7 ± 0.5 seconds across all subjects, whereas the running time for the non-rigid RPM method was 13.1 ± 2.0 seconds. In this case, the SDM registration method resulted in a computational improvement of 72%. Running time for the distance map FFD was 243.5 ± 80.2 seconds.

3.3.3. Registration Performance with Segmentation Error—To test the stability of the surface-based registration approaches, we added synthetic noise to mimic prostate segmentation variability in the TRUS images. Similar to as done in Sec. 3.2, we used our large database to calculate the distribution of TRUS segmentation variability (Appendix B), and created a spatial model of TRUS segmentation variance (Appendix C). Our model of segmentation variability across subjects required a TRUS template prostate shape $\mathcal{S}_{\text{SDM}}$, which we warped to each $\mathcal{S}_{\text{TRUS},s}$ in our testing set using a non-rigid FFD RPM transformation (Chui and Rangarajan, 2003; Papademetris et al., 2003). This warped version of the template $\hat{\mathcal{S}}_{\text{TRUS},s}$ replaces the original surface $\mathcal{S}_{\text{TRUS},s}$ in these tests. We then added realistic synthetic segmentation noise to each $\hat{\mathcal{S}}_{\text{TRUS},s}$ using our model of segmentation variability in Appendix C by randomly sampling values $u_j \sim \mathcal{T}(0, \sigma v_j)$, where v_j is the j -th eigenvalue, using the first $j = 1, \dots, 10$ modes in (C.1) with a noise level scalar $\sigma = 1.0$. A noise level $\sigma = 1.0$ represents the standard amount of segmentation variance observed in the population (assuming a Gaussian normal distribution). Finally, for each subject $s = 1, \dots, 22$, we created 100 synthetic TRUS shape $\hat{\mathcal{S}}_{\text{TRUS},s,\sigma}$, which resulted in 2,200 synthetic samples in total. We then registered each test subject's MR prostate surface $\mathcal{S}_{\text{MR},s}$ to the 100 TRUS samples $\hat{\mathcal{S}}_{\text{TRUS},s,\sigma}$ using the following RPM surface-based registration methods: (i) rigid (Rangarajan et al., 1997), (ii) affine, (iii) non-rigid FFD, and (iv) our proposed non-rigid SDM using $J = 17$ eigenvectors. Table 4 shows the registration results. Using a Wilcoxon rank-sum test with a significance level of 0.05 to test for significant differences in TRE between each of the four methods, we found both our SDM and the RPM FFD methods had significantly lower TRE than the rigid and affine RPM methods. There were no significant differences between the RPM FFD and our SDM 17 methods, and no significant differences between the RPM rigid and RPM affine methods.

4. Discussion and Conclusion

We have demonstrated the learning and the application of statistical deformation models (SDMs) to perform constrained, non-rigid surface-based registration for fusing MR and TRUS images in image-guided prostate biopsy. Our method works by learning a model of the deformations typically encountered during the biopsy procedure given a large database of prostate biopsy interventions. We then apply that model to guide the registration of novel

MR-TRUS prostate image pairs. We evaluated the performance of our SDM registration method by performing tests using both synthetic and real clinical data (Sec. 3), and we demonstrate the effectiveness of our approach to significantly improve registration performance compared to the method currently used by an existing clinical device. In addition to the SDM's significantly improved registration performance, the synthetic tests highlight the SDMs more robust and more consistent registration results, which is critical for clinical image-based procedures, where time is limited, data is often times imperfect, and minimizing the maximum registration error is important. In other words, the registration algorithm should be robust to noisy data and performance should degrade gracefully. Compared to the other registration methods tested, our results (Fig. 7) demonstrate the SDM's characteristics of: (i) *reliability* in that it does not fail badly (higher minimum Dice overlaps); (ii) *robustness* to increasing levels of segmentation errors (higher mean Dice overlaps); and (iii) *consistency* in that performance holds over the population (smaller interquartile range).

Our landmark registration tests using clinical data, with target registration error (TRE) 3.50 ± 2.02 (mean $\pm$ SD), were within the clinically significant threshold of 5 mm (Karnik et al., 2010). Nevertheless, localizing corresponding landmarks within clinical MR and TRUS images is extremely hard, and from our experience, subject to high variability. Landmark TRE is the most widely reported metric for MR-TRUS registration performance by the methods in Sec. 1.1, but results between different studies are difficult to compare due to different landmarks being used by each study as well as the challenges in localizing these landmarks. While prostate landmarks common to all patients have been used to assess MR-TRUS registration performance, for example the entry and exit points of the urethra, verumontanum, and the midpoint of the seminal vesicles entry (Hu et al., 2012; Fedorov et al., 2015), we found that both intra- and inter-rater identification accuracy was too low for us to reliably use these landmarks to quantify registration error. Localizing these landmarks was challenging due to the ambiguous delineation of the prostate gland surface in parts of the TRUS imaging with little tissue contrast. We found that patient-specific cysts and other "cyst-like" anatomical structures appearing in both the MR and TRUS images were the most reliable to localize, and these types of landmarks have also been used to calculate TRE in other studies (Hu et al., 2012, 2015; Fedorov et al., 2015; Khallaghi et al., 2015b,a). The 29 identified landmarks used in Sec. 3.3 were distributed throughout the prostate volume of 22 subjects (Fig. 4), however, our use of a limited number of subjects with identifiable cysts does have the disadvantage that they could potentially bias our registration results.

To address these challenges, the detailed synthetic fusion results in Sec. 3.2 are invaluable for a more comprehensive evaluation. Compared to the landmark results, which at most report registration error results at 2 point locations in the prostate for a given subject, the synthetic results show a more holistic assessment of registration performance using multiple, clinically-relevant VOIs spanning the entire prostate gland volume. Our calculated distribution of TRUS segmentation variability (Fig. 5) shows the greatest segmentation variability in the prostate apex and base regions and lowest variability in the midgland. These results agree with those presented by Smith *et al.* (Smith et al., 2007), even though Smith *et al.*'s measure of segmentation variability relates to the distribution of inter-rater segmentation variance for 10 images whereas our measure derives from a distribution of

segmentation variance across a population of more than 100 images for a single rater. That these synthetic examples are realistic to clinically-observed segmentation errors contrasts with prior studies that introduced synthetic segmentation errors at random locations on the prostate surface as first done by Onofrey *et al.* (Onofrey et al., 2015b) and later by Khallaghi *et al.* (Khallaghi et al., 2015a). The synthetic tests, furthermore, also eliminated the potential for linear registration errors that could be present in the clinical data tests. With the prostate being roughly ellipsoidal in shape, the linear registration step has rotational ambiguities that could potentially result in large registration error prior to the application of any non-rigid registration algorithms. To avoid such problems, some other methods utilize an initial manual identification of a small number of fiducials prior to registration (Hu et al., 2012, 2015). Nevertheless, our results using clinical data (Sec. 3.3) suggest that misregistration to the template shape is not an issue, as the SDM registration results performed at the level of standard non-rigid RPM and distance map image-based registration.

The synthetic test results demonstrate the utility and effectiveness of the SDM registration in comparison to the unconstrained non-rigid free-form deformation (FFD) registration method, which also serves as a surrogate registration method in these tests for the AFDM method used by the existing state-of-the-art clinical system. Here, the relatively poor performance of the standard, unconstrained FFD-based non-rigid registration was notable (Figs. 7 and 8). In cases of segmentations with even low levels of noise, which is not uncommon in clinical use (Appendix B), the linear affine registration method performed better than the unconstrained non-rigid FFD transformation. The non-rigid FFD transformation with high degrees of freedom (DoFs) offered too much flexibility and the algorithm incorrectly registered the prostate surface to segmentation noise. Similarly, other studies have reported increased TRE for unconstrained non-rigid registration methods over rigid registration (Cool et al., 2011; Khallaghi et al., 2015b) and that unconstrained, deformable registration did not show significant improvements over affine registration (Delongchamps et al., 2013; Fedorov et al., 2015). We have seen analogous results when performing intensity-based multi-modal brain registration with an SDM (Onofrey et al., 2016), where rigid registration can outperform unconstrained non-rigid registration due to noisy data. The low DoF affine transformation model limits the prostate surfaces from over-registering (fitting the noise too well), but we also see that a linear transformation is not enough to account for the non-rigid deformations of the biopsy procedure at clinically-observed levels of segmentation noise (Fig. 7), as well as in real data (Fig. 10).

The non-rigid SDM registration method is well-suited for registering surfaces under noisy conditions as it offers both a non-rigid deformation that can accommodate complex changes to prostate shape and a low-dimensional parameterization of that transformation to avoid over-fitting. We observed robust registration performance behavior in both our synthetic (3.2) and clinical (3.3.3) experiments where we tested registration using TRUS prostate segmentations with added errors that realistically mimic segmentation variability. Khallaghi *et al.* (Khallaghi et al., 2015a) had similar success using their statistical model of prostate shape to effectively constrain alignment of the prostate to segmentations with missing data. The best performing SDMs had non-rigid SDM deformations with 11 DoFs for the synthetic tests while the clinical tests had 17 DoFs. From the results in this paper, the ideal number of

eigenvectors for the prostate SDM registration methods appears to be J_v accounting for 75%–90% of the cumulative variance (Sec. 2.2).

Our use of the robust point matching (RPM) framework to handle outliers in surface data was itself a major contributor to improved registration performance. As demonstrated by non-rigid FFD RPM landmark registration results (Fig. 10), which was an unconstrained deformable transformation similar to the one used by AFDM, the RPM framework allowed for outlier points on the surface segmentations and so it was robust to noisy segmentations. The AFDM registration method, which does not model point outliers in the dataset, aligned the two surfaces, and any segmentations errors, almost perfectly. The RPM registrations with linear rigid and affine transformations models, though not significantly different from the AFDM results in terms of landmark TRE, had much better worst-case performance (lower maximum TRE over the testing samples). Alternative surface-based registration methods that accommodate missing data, such as coherent point drift (Myronenko and Song, 2010), have found similar success for MR-TRUS fusion (Fedorov et al., 2015; Khallaghi et al., 2015b,a). Distance map-based registration methods (Fedorov et al., 2015) have the potential to align the interior of the prostate gland, which is otherwise ignored in surface-based registrations. However, our experiments showed that the surface-based registrations performed on par with distance map-based registration (Fig. 10), with no significant difference being found between the surface-based RPM and distance map methods. While the distance map method registers two dense distance map images, these distance maps are still based on the initial surface segmentations to provide the distance map calculation, and are thus susceptible to segmentation errors.

The AFDM registration algorithm, in contrast, relies on accurate prostate segmentation, which is a challenge due to the poor contrast of ultrasound imaging in certain areas of the gland (Fig. 5). The high landmark TRE and large variance of the AFDM method (Fig. 10) demonstrates its sensitivity to segmentation error. While the AFDM method used by the clinical system had high landmark TRE in our one-time registration tests, the urologist, in clinical practice, has the opportunity to intervene and repeat the segmentation and fusion process if the initial results are not acceptable. Thus, we expect the registration errors for the clinical device to be lower in practice than observed for these experiments. Nevertheless, in the existing clinical system, the human operator is responsible for detecting and correcting fusion errors, whereas our proposed method provides reliable, robust, and consistent fusion results automatically.

Low-dimensional SDM registration is also computationally efficient. The SDM learning procedure occurs only once and the model can be stored as part of the clinical system for future use, while the warping of the pre-procedure MR surface to the SDM template may be done offline prior to the biopsy. In Sec. 3.3, we reduced registration time by 72% from non-rigid FFD (13.1 ± 2.0 seconds) to SDM 17 (3.7 ± 0.5 seconds). The surface-based registration methods provide a large speedup in computation time compared to the image-based distance map registration methods (243.5 seconds on average for non-rigid FFD). A GPU-based implementation of our algorithm could potentially reduce running times to the sub-second level. Such fast performance would have the potential for the registration to be run before each needle insertion. However, due to the relatively long time required to

reconstruct the 3D TRUS volume (approximately 30 seconds), per-biopsy fusion updates would most likely require registration of 2D TRUS slices to the original 3D TRUS volume (Xu et al., 2008; Cool et al., 2011). A good portion of the intra-procedure image fusion processing time (approximately 2–5 minutes, or higher if refinement is needed) is dedicated to the semiautomated segmentation of the TRUS prostate and verification of the initial alignment. However, our results show that the segmentation does not need to be perfect as our method is robust to segmentation errors and variability.

Our method has certain limitations. Surface-based registration algorithms are highly dependent upon the semiautomated segmentation of the prostate in both the MR and the TRUS images. The surface-based registrations used throughout this paper are inherently less accurate farther away from the surface boundary, i.e. the prostate gland center, due to the lack of internal information. Without information to guide registration inside the gland, our model training data may contain potential deformation errors. Using a distance map-based registration formulation (Fedorov et al., 2015) or incorporating a biomechanical model into the prostate deformation training and registration process (Hu et al., 2012) could address this issue, and other studies have demonstrated that FEM can be advantageous for transformation regularization (Khallaghi et al., 2015b; van de Ven et al., 2015). The SDM training and registration is also dependent upon the SDM template shape. While we created a prostate template shape with a volume close to the distribution's mode (34 cm^3), our choice of a single template might bias our registration towards prostates of similar size. An alternative template construction approach might utilize group-wise registration (Rasoulian et al., 2012) to create a less-biased prostate template. Furthermore, even though we normalized for differences in prostate volume using a similarity transform, the deformations learned by the SDM may not accurately reflect those seen in prostates with extremely large volumes, *e.g.* 9 subjects had volumes in excess of 100 cm^3 with one subject having maximum volume 262 cm^3 . To handle such cases, we plan to investigate training subject-specific SDMs where the SDM training happens in the native MR space of the given subject. We also aim to incorporate our model of segmentation variability into the registration procedure as a form of prior knowledge about where the segmentation is likely to be incorrect. Finally, our method is currently limited in its application to the image-guided biopsy configuration of the Artemis system, which creates 3D TRUS image volume reconstructions using a rotation of a 2D US probe. Nevertheless, learned models of deformation such as the one presented here could be applied to any system employing surface-based registration techniques. We further limited our training to use MR images (and their respective surface segmentations) acquired without an endorectal coil. Using two types of MR images, either acquired with or without an endorectal coil, would potentially represent two distinct classes of deformations, and constructing a model capable of handling these heterogeneous deformations remains for future work.

To summarize, we have applied a statistical model of intra-procedure prostate biopsy deformation to non-rigidly register pre-procedure MR images to intra-procedure TRUS imaging. We validated our method using both synthetic and real clinical data to demonstrate significantly improved registration quality, consistency and robustness. By constraining the non-rigid surface registration procedure to the space of previously observed intra-procedure prostate deformation, the SDM is not only effective at registering the MR and TRUS

surfaces, but also prevents the registration process from failing in the presence of prostate gland segmentation errors. Our method offers improved fusion performance over a current clinical state-of-the-art system, and may easily be incorporated into the existing clinical workflow.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health (NIH) National Cancer Institute (NCI) R41 CA186414.

References

- Alterovitz R, Goldberg K, Pouliot J, Hsu ICJ, Kim Y, Noworolski SM, Kurhanewicz J. Registration of MR prostate images with biomechanical modeling and nonlinear parameter estimation. *Med Phys*. 2006; 33(2):446. [PubMed: 16532952]
- Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, Rouviere O, Logager V, Fütterer JJ. ESUR prostate MR guidelines 2012. *European Radiology*. 2012; 22(4):746–757. [PubMed: 22322308]
- Bubley GJ, Bloch BN, Vazquez C, Genega E, Holupka E, Rofsky N, Kaplan I. Accuracy of Endorectal Magnetic Resonance/Transrectal Ultrasound Fusion for Detection of Prostate Cancer During Brachytherapy. *Urology*. 2013; 81(6):1284–1290. [PubMed: 23490527]
- Chui H, Rangarajan A. A new point matching algorithm for non-rigid registration. *Comput Vis Image Underst*. 2003; 89(2–3):114–141.
- Cool DW, Bax J, Romagnoli C, Ward AD, Gardi L, Karnik V, Izawa J, Chin J, Fenster A. Fusion of MRI to 3D TRUS for mechanically-assisted targeted prostate biopsy: System design and initial clinical experience. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2011; 6963:121–133. LNCS.
- Cooperberg MR, Broering JM, Kantoff PW, Carroll PR. Contemporary Trends in Low Risk Prostate Cancer: Risk Assessment and Treatment. *J Urol*. 2007; 178(3, Supplement):S14–S19. [PubMed: 17644125]
- Cootes T, Taylor C, Cooper D, Graham J. Active shape models-their training and application. *Computer Vision and Image Understanding*. 1995; 61(1):38–59.
- Delongchamps NB, Peyromaure M, Schull A, Beuvon F, Bouazza N, Flam T, Zerbib M, Muradyan N, Legman P, Cornud F. Prebiopsy magnetic resonance imaging and prostate cancer detection: comparison of random and targeted biopsies. *J Urol*. 2013; 189(2):493–9. [PubMed: 22982424]
- Dickinson L, Hu Y, Ahmed HU, Allen C, Kirkham AP, Emberton M, Barratt D. Image-directed, tissue-preserving focal therapy of prostate cancer: a feasibility study of a novel deformable magnetic resonance-ultrasound (MR-US) registration system. *BJU International*. 2013; 112(5):594–601. [PubMed: 23819525]
- Fedorov A, Khalaghi S, Sánchez CA, Lasso A, Fels S, Tuncali K, Sugar EN, Kapur T, Zhang C, Wells W, Nguyen PL, Abolmaesumi P, Tempany C. Open-source image registration for MRI-TRUS fusion-guided prostate interventions. *Int J Comput Assist Radiol Surg*. 2015; 10(6):925–934. [PubMed: 25847666]
- Heimann T, Meinzer HP. Statistical shape models for 3D medical image segmentation: A review. *Med Image Anal*. 2009; 13(4):543–563. [PubMed: 19525140]
- Hu Y, Ahmed HU, Taylor Z, Allen C, Emberton M, Hawkes D, Barratt D. MR to ultrasound registration for image-guided prostate interventions. *Medical Image Analysis*. 2012; 16(3):687–703. [PubMed: 21216180]

- Hu Y, Gibson E, Ahmed HU, Moore CM, Emberton M, Barratt DC. Population-based prediction of subject-specific prostate deformation for MR-to-ultrasound image registration. *Med Image Anal.* 2015; 26(1):332–44. [PubMed: 26606458]
- Isariyawongse BK, Sun L, Bañez LL, Robertson C, Polascik TJ, Maloney K, Donatucci C, Albala D, Mouraviev V, Madden JF, Moul JW. Significant Discrepancies Between Diagnostic and Pathologic Gleason Sums in Prostate Cancer: The Predictive Role of Age and Prostate-Specific Antigen. *Urology.* 2008; 72(4):882–886. [PubMed: 18384857]
- Joshi A, Scheinost D, Okuda H, Belhachemi D, Murphy I, Staib L, Papademetris X. Unified Framework for Development, Deployment and Robust Testing of Neuroimaging Algorithms. *Neuroinformatics.* 2011; 9(1):69–84. [PubMed: 21249532]
- Kaplan I, Oldenburg NE, Meskell P, Blake M, Church P, Holupka EJ. Real time MRI-ultrasound image guided stereotactic prostate biopsy. *Magnetic Resonance Imaging.* 2002; 20(3):295–299. [PubMed: 12117612]
- Karnik VV, Fenster A, Bax J, Cool DW, Gardi L, Gyacskov I, Romagnoli C, Ward AD. Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy. *Med Phys.* 2010; 37(2):802–813. [PubMed: 20229890]
- Khallaghi S, Sanchez CA, Rasoulia A, Nouranian S, Romagnoli C, Abdi H, Chang SD, Black PC, Goldenberg L, Morris WJ, Spadinger I, Fenster A, Ward A, Fels S, Abolmaesumi P. Statistical Biomechanical Surface Registration: Application to MR-TRUS Fusion for Prostate Interventions. *IEEE Trans Med Imaging.* 2015a; 34(12):2535–2549. [PubMed: 26080380]
- Khallaghi S, Sanchez CA, Rasoulia A, Sun Y, Imani F, Khojaste A, Goksel O, Romagnoli C, Abdi H, Chang S, Mousavi P, Fenster A, Ward A, Fels S, Abolmaesumi P. Biomechanically Constrained Surface Registration: Application to MR-TRUS Fusion for Prostate Interventions. *IEEE Trans Med Imaging.* 2015b; 34(11):2404–2414. [PubMed: 26054062]
- Kim CK, Park BK, Lee HM, Kwon GY. Value of Diffusion-Weighted Imaging for the Prediction of Prostate Cancer Location at 3T Using a Phased-Array Coil. *Invest Radiol.* 2007; 42(12):842–847. [PubMed: 18007156]
- Kirkham APS, Emberton M, Allen C. How Good is MRI at Detecting and Characterising Cancer within the Prostate? *Eur Urol.* 2006; 50(6):1163–1175. [PubMed: 16842903]
- Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, Christensen GE, Collins DL, Gee J, Hellier P, Song JH, Jenkinson M, Lepage C, Rueckert D, Thompson P, Vercauteren T, Woods RP, Mann JJ, Parsey RV. Evaluation of 14 nonlinear deformation algorithms applied to human brain {MRI} registration. *NeuroImage.* 2009; 46(3):786–802. [PubMed: 19195496]
- Krucker, J., Xu, S., Glossop, N., Guion, P., Choyke, P., Ocak, I., Singh, AK., Wood, BJ. Fusion of real-time transrectal ultrasound with preacquired MRI for multimodality prostate imaging. In: Cleary, KR., Miga, MI., editors. *Proc SPIE.* Vol. 6509. 2007. p. 650912
- Ladak HM, Mao F, Wang Y, Downey DB, Steinman DA, Fenster A. Prostate boundary segmentation from 2D ultrasound images. *Med Phys.* 2000; 27(8):1777–1788. [PubMed: 10984224]
- Martin, S., Baumann, M., Daanen, V., Troccaz, J. MR prior based automatic segmentation of the prostate in TRUS images for MR/TRUS data fusion. 2010 IEEE Int. Symp. Biomed. Imaging From Nano to Macro; IEEE; 2010. p. 640-643.
- Mitra J, Kato Z, Martí R, Oliver A, Lladó X, Sidibé D, Ghose S, Vilanova JC, Comet J, Meriaudeau F. A spline-based non-linear diffeomorphism for multimodal prostate registration. *Med Image Anal.* 2012a; 16(6):1259–1279. [PubMed: 22705289]
- Mitra J, Martí R, Oliver A, Lladó X, Ghose S, Vilanova J, Meriaudeau F. Prostate multimodality image registration based on B-splines and quadrature local energy. *Int J of Comp Assisted Radiology and Surgery.* 2012b; 7(3):445–454.
- Moradi M, Janoos F, Fedorov A, Risholm P, Kapur T, Wolfsberger L, Nguyen P, Tempany C, Wells W. Two solutions for registration of ultrasound to MRI for image-guided prostate interventions. *IEEE EMBC.* 2012:1129–1132.
- Myronenko A, Song X. Point Set Registration: Coherent Point Drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2010; 32(12):2262–2275. [PubMed: 20975122]
- Narayanan R, Kurhanewicz J, Shinohara K, Crawford ED, Simoneau A, Suri J. MRI-ultrasound registration for targeted prostate biopsy. *IEEE ISBI.* 2009:991–994.

- Natarajan S, Marks LS, Margolis DJA, Huang J, Macairan ML, Lieu P, Fenster A. Clinical application of a 3D ultrasound-guided prostate biopsy system. *Urologic Oncology: Seminars and Original Investigations*. 2011; 29(3):334–342. [PubMed: 21555104]
- Onofrey JA, Papademetris X, Staib LH. Low-Dimensional Non-Rigid Image Registration Using Statistical Deformation Models From Semi-Supervised Training Data. *IEEE Trans Med Imaging*. 2015a; 34(7):1522–1532.
- Onofrey JA, Staib LH, Papademetris X. Learning intervention-induced deformations for non-rigid MR-CT registration and electrode localization in epilepsy patients. *NeuroImage: Clinical*. 2016; 10:291–301. [PubMed: 26900569]
- Onofrey, JA., Staib, LH., Sarkar, S., Venkataraman, R., Papademetris, X. Learning nonrigid deformations for constrained point-based registration for image-guided MR-TRUS prostate intervention. *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*; 2015b. p. 1592-1595.
- Papademetris, X., Jackowski, AP., Schultz, RT., Staib, LH., Duncan, JS. Computing 3D Non-rigid Brain Registration Using Extended Robust Point Matching for Composite Multisubject fMRI Analysis. In: Ellis, RE., Peters, TM., editors. *Med Image Comput Comput Interv - MICCAI 2003*. Vol. 2879. Springer; Berlin Heidelberg: 2003. p. 788-795.
- Pondman KM, Fütterer JJ, ten Haken B, Kool LJS, Witjes JA, Hambroek T, Macura KJ, Barentsz JO. MR-Guided Biopsy of the Prostate: An Overview of Techniques and a Systematic Review. *Eur Urol*. 2008; 54(3):517–527. [PubMed: 18571309]
- Rangarajan A, Chui H, Mjolsness E, Pappu S, Davachi L, Goldman-Rakic P, Duncan J. A robust point-matching algorithm for autoradiograph alignment. *Medical Image Analysis*. 1997; 1(4):379–398. [PubMed: 9873917]
- Rasoulian A, Rohling R, Abolmaesumi P. Group-Wise Registration of Point Sets for Statistical Shape Models. *IEEE Transactions on Medical Imaging*. 2012; 31(11):2025–2034. [PubMed: 22692899]
- Reynier C, Troccaz J, Fournieret P, Dusserre A, Gay-Jeune C, Descotes JL, Bolla M, Giraud JY. MRI/TRUS data fusion for prostate brachytherapy. Preliminary results *Med Phys*. 2004; 31(6):1568. [PubMed: 15259661]
- Roehl KA, Antenor JAV, Catalona WJ. Serial Biopsy Results in Prostate Cancer Screening Study. *J Urol*. 2002; 167(6):2435–2439. [PubMed: 11992052]
- Rohlfing T. Image Similarity and Tissue Overlaps as Surrogates for Image Registration Accuracy: Widely Used but Unreliable. *Medical Imaging, IEEE Transactions on*. 2012; 31(2):153–163.
- Rueckert D, Frangi A, Schnabel J. Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Trans Med Imaging*. 2003; 22(8):1014–1025. [PubMed: 12906255]
- Rueckert D, Sonoda L, Hayes C, Hill D, Leach M, Hawkes D. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging*. 1999; 18(8):712–721. [PubMed: 10534053]
- Shen D, Herskovits EH, Davatzikos C. An adaptive-focus statistical shape model for segmentation and shape modeling of 3-D brain structures. *Medical Imaging, IEEE Transactions on*. 2001; 20(4):257–270.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin*. 2016; 66(1):7–30. [PubMed: 26742998]
- Smith WL, Lewis C, Bauman G, Rodrigues G, D'Souza D, Ash R, Ho D, Venkatesan V, Downey D, Fenster A. Prostate volume contouring: A 3D analysis of segmentation using 3DTRUS, CT, and MR. *Int J Radiat Oncol*. 2007; 67(4):1238–1247.
- Sparks R, Nicolas Bloch B, Feleppa E, Barratt D, Moses D, Ponsky L, Madabhushi A. Multiattribute probabilistic prostate elastic registration (MAPPER): Application to fusion of ultrasound and magnetic resonance imaging. *Med Phys*. 2015; 42(3):1153–1163. [PubMed: 25735270]
- Sun, Y., Yuan, J., Rajchl, M., Qiu, W., Romagnoli, C., Fenster, A. Efficient convex optimization approach to 3D non-rigid MR-TRUS registration. *MICCAI*; Springer; 2013. p. 195-202.
- Taira AV, Merrick GS, Galbreath RW, Andreini H, Taubenslag W, Curtis R, Butler WM, Adamovich E, Wallner KE. Performance of transperineal template-guided mapping biopsy in detecting prostate

- cancer in the initial and repeat biopsy setting. *Prostate Cancer Prostatic Dis.* 2010; 13(1):71–77. [PubMed: 19786982]
- Ukimura O, Faber K, Gill IS. Intraprostatic targeting. *Current opinion in urology.* 2012; 22(2):97–103. [PubMed: 22249373]
- van de Ven WJM, Hu Y, Barentsz JO, Karsssemeijer N, Barratt D, Huisman HJ. Biomechanical modeling constrained surface-based image registration for prostate MR guided TRUS biopsy. 2015
- Villers A, Puech P, Mouton D, Leroy X, Ballereau C, Lemaitre L. Dynamic Contrast Enhanced, Pelvic Phased Array Magnetic Resonance Imaging of Localized Prostate Cancer for Predicting Tumor Volume: Correlation With Radical Prostatectomy Findings. *J Urol.* 2006; 176(6):2432–2437. [PubMed: 17085122]
- Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, Margolis D, Schnall MD, Shtern F, Tempany CM, Thoeny HC, Verma S. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *European Urology.* 2016; 69(1):16–40. [PubMed: 26427566]
- West J, Fitzpatrick JM, Wang MY, Dawant BM, Maurer CR, Kessler RM, Maciunas RJ, Barillot C, Lemoine D, Collignon A, Maes F, Sumanaweera TS, Harkness B, Hemler PF, Hill DLG, Hawkes DJ, Studholme C, Maintz JBA, Viergever MA, Mal G, Pennec X, Noz ME, Maguire GQ, Pollack M, Pelizzari CA, Robb RA, Hanson D, Woods RP. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography.* 1997; 21:554–566. [PubMed: 9216759]
- Xu S., Kruecker J., Guion, P., Glossop, N., Neeman, Z., Choyke, P., Singh, AK., Wood, BJ. Closed-Loop Control in Fused MR-TRUS Image-Guided Prostate Biopsy. In: Ayache, N., Ourselin, S., Maeder, A., editors. *Med Image Comput Comput Interv – MICCAI 2007.* Vol. 4791. Springer; Berlin Heidelberg: 2007. p. 128-135.
- Xu S, Kruecker J, Turkbey B, Glossop N, Singh AK, Choyke P, Pinto P, Wood BJ. Real-time MRI-TRUS fusion for guidance of targeted prostate biopsies. *Comput Aided Surg.* 2008; 13(5):255–264. [PubMed: 18821344]
- Yan P, Xu S, Turkbey B, Kruecker J. Discrete Deformable Model Guided by Partial Active Shape Model for TRUS Image Segmentation. *IEEE Trans Biomed Eng.* 2010; 57(5):1158–1166. [PubMed: 20142158]
- Zettinig O, Shah A, Hennemersperger C, Eiber M, Kroll C, Kübler H, Maurer T, Milletari F, Rackerseder J, Schulte Zu Berge C, Storz E, Frisch B, Navab N. Multimodal image-guided prostate fusion biopsy based on automatic deformable registration. *Int J Comput Assist Radiol Surg.* 2015; 10(12):1997–2007. [PubMed: 26054983]

Appendix A. Robust Point Matching

The robust point matching (RPM) methodology for surface registration (Rangarajan et al., 1997) functions by alternately estimating surface point correspondences and estimating the transformation. Denoting the reference point set as X and the transform point set as Y , the goal of RPM registration is to estimate the transformation $T: X \mapsto Y$, with T^k the estimate of T at iteration k . The match matrix M estimates the fuzzy correspondence (ranging from 0 to 1) between points $T(X_i)$ and Y_j

$$M_{ij} = \frac{1}{\sqrt{2\pi t^2}} \exp \left\{ \frac{-\|T^k(X_i) - Y_j\|^2}{2t^2} \right\}$$

according to the temperature term t in an annealing formulation that controls the fuzziness of the correspondence. The standard RPM framework allows for outlier point estimation in both X and Y . To accommodate these outliers into the RPM formulation, an outlier column

C and an outlier row R measure the “outlierness” (also ranging from 0 to 1) for each point X_i and Y_j , respectively. The values of M , C and R are normalized such that:

$$\sum_j M_{ij} + C_i = 1, \forall i, \text{ and } \sum_i M_{ij} + R_j = 1, \forall j.$$

Upon normalization, RPM computes each X_i 's corresponding point X'_i as a normalized weighted sum of points Y_j such that

$$X'_i = (\sum_j M_{ij} Y_j) / (\sum_j M_{ij}), \text{ and } c_i = \sum_j M_{ij} = 1 - C_i$$

where c_i is the total confidence of the correspondence.

Given the correspondence estimate, RPM then estimates the transformation at step k as a weighted least squares fit between X_i and X'_i

$$T^k = \arg \min_T \sum_i c_i \|T(X_i) - X'_i\|^2 + f(t)R(T) \quad (\text{A.1})$$

where $R(T)$ is a regularization functional (*e.g.* bending energy) weighted by a function of the annealing temperature $f(t)$. A deterministic annealing framework is used to alternately solve for M and T . The framework begins with a high value of t , which corresponds to a rough estimate of the maximum misalignment distance, and then gradually decreases t by an annealing factor $0 < a < 1$ until t becomes sufficiently small. The RPM formulation in (A.1) allows for the transformation T to be of any form. T can be either a linear (Rangarajan et al., 1997) or a non-rigid transformation, with example non-rigid deformations parameterized by thin-plate spline (Chui and Rangarajan, 2003) and B-spline free-form deformation (Papademetris et al., 2003) transformation models.

Appendix B. Distribution of TRUS Prostate Segmentation Variability

To find the distribution of TRUS prostate segmentation variability in our database of N subjects, the same urologist who originally segmented the TRUS images resegmented each of the subjects a second time using the same semi-automated clinical protocol. We denote these two segmentations $S_{\text{TRUS},i}^t$ at time $t = 1, 2$ for subject $i = 1, \dots, N$. In order to estimate the distribution of segmentation variability both spatially and across the population, we created a TRUS template prostate surface S_t using the same method in Sec. 2.2, but used the TRUS surfaces instead of the MR surfaces to create the template surface shape. We then non-rigidly warped S_t to all instances of $S_{\text{TRUS},i}^t$ using robust point matching (RPM) surface registration (Chui and Rangarajan, 2003; Papademetris et al., 2003) with 5.0 mm isotropic control point spacing, and denote these warped surfaces $S_{\text{TRUS},i}^{t'}$. For each of the $j = 1, \dots, J$

surface points $\mathbf{p}_{i,j} \in S_{\text{TRUS},i}^{1'}$, we found the closest point $\mathbf{q}_{i,j} \in S_{\text{TRUS},i}^{2'}$, where $J=3,468$ for this template. Rather than using the Euclidean distance $\delta_{i,j} = \|\mathbf{p}_{i,j} - \mathbf{q}_{i,j}\|$ at each point $\mathbf{p}_{i,j}$ as our measure of segmentation variability, we project $\delta_{i,j}$ onto the outward-facing direction

vector $\mathbf{v}_{i,j} = (\mathbf{v}_{i,j} - \mathbf{c}_i) / \|\mathbf{c}_{i,j} - \mathbf{c}_i\|$, where $\mathbf{c}_i \in \mathbb{R}^3$ is $S_{\text{TRUS},i}^{1'}$'s center of gravity. By projecting the Euclidean distance onto this direction vector $\delta_{i,j} = \delta_{i,j} \mathbf{v}_{i,j}$, we can encode the direction of segmentation error in a way that generalizes across the population. Thus, we can calculate the pointwise distribution of segmentation variability (mean $\pm$ SD) over the population at each point $\mathbf{p}_j \in S_t$. Fig. 5 visualizes this distribution on the template prostate shape. This distribution shows the statistical variability of a single expert performing whole gland prostate segmentation in the TRUS images.

Appendix C. Modeling TRUS Segmentation Variability with PCA

Given samples from the distribution of TRUS segmentation variability, we perform a principal component analysis (PCA) of the sample covariance matrix to construct a linear model of segmentation variability. A PCA of the N training samples allows us to model correlated segmentation variance across the population. From Appendix B, for each subject $i = 1, \dots, N$, we have J pointwise segmentation distances projected along the surface vector pointing outwards from the surface center of gravity, $\delta_{i,j} \in \mathbb{R}^3$, which we concatenate into a single vector $\delta_i \in \mathbb{R}^{3J}$. These vectors are N samples from the distribution of segmentation variance $\mathcal{S}$ across the population. PCA of these samples yields a model

$$\delta = \bar{\delta} - \Psi \mathbf{u}, \quad (\text{C.1})$$

where $\bar{\delta}$ is the population mean segmentation variance, and $\Psi = [\boldsymbol{\psi}_1 | \dots | \boldsymbol{\psi}_J]$ is the matrix of J orthogonal eigenvectors $\boldsymbol{\psi}_j$ describing the principal modes of segmentation variability. These eigenvectors have corresponding eigenvalues v_j and the eigenvectors are sorted in decreasing order such that $v_1 \geq v_2 \geq \dots \geq v_J$. Fig. 6 shows some synthetic prostate shape examples created using this model.

Highlights

- Model non-rigid deformations typically encountered when fusing pre-procedure MR and intra-procedure TRUS images for image-guided prostate biopsy
- A large database of clinical prostate biopsy interventions is used to train a statistical deformation model (SDM)
- The SDM prevents the registration process from failing in the presence of prostate gland segmentation errors
- Rigorous validation using synthetic data and clinical landmarks demonstrates accurate, reliable, robust, and consistent registration results

Figure 1.

We construct a prostate statistical deformation model (SDM) that captures the principal directions of prostate deformation between its pre-procedural and intra-procedural shapes in MR and TRUS images, respectively. Using a large set of clinical training data, our approach first normalizes all the segmented prostate surfaces to a template prostate shape. Non-rigid, intra-subject, surface-based registration within template space captures the intra-procedure prostate deformation, and we use these transformations to learn our SDM. Given a novel MR-TRUS image pair, and after normalizing these prostate surfaces to template space, we perform non-rigid surface registration using the learned SDM to constrain the deformation process and robustly fuse the MR-TRUS imaging.

Intra-procedure Prostate Deformation

Figure 2.

Visualizing example non-rigid deformations of the prostate using our statistical model of intra-procedure deformation. The central prostate shape shows the template prostate with the model's mean deformation applied. The three axes show example deformations according to the first 3 eigenvectors of deformations, where $\pm\phi_i$ indicates deformation along the i -th eigenvector. The first eigenvector captures expansion of the gland, while the second and third eigenvectors capture deformations induced by insertion of the TRUS probe. The axes indicate (A) anterior, (P) posterior, (R) right, (L) left, (S) superior, and (I) inferior orientations with respect to the subject.

PI-RADS Parcellation

Figure 3.

The clinical PI-RADS reporting standard parcellates the prostate into 37 volumes of interest (VOIs). These VOIs include the following anatomical zones: peripheral zone (PZ), transition zone (TZ), central zone (CZ), urethra (U), and anterior fibromuscular stroma (AFS). These zones are subdivided into their left/right anterior (a), posterior (p), posterior medial (pm), and posterior lateral (pl) components within the prostate base, midgland, and apex.

Spatial Distribution of Landmarks

Figure 4.

The spatial distribution of 29 manually identified landmarks in 22 different subjects visualized within a template prostate shape. Landmarks are color-coded to their corresponding subject identification number in the testing set. The maximum number of landmarks per subject is 2. The labels indicate (A) anterior, (P) posterior, (R) right, and (L) left orientations with respect to the prostate. Below, we show example landmarks localized in both the MR and the TRUS images for 3 subjects.

Figure 5.

Visualizing the distribution of TRUS prostate segmentation variability on an average prostate template shape. We show the mean and standard deviation for $N=103$ subjects, with two 3D views to highlight the prostate gland base (right column) and apex (left column). Additionally, we show the prostate surface contour overlaid onto the mean intensity TRUS image (middle column). Across the population, the greatest amount of segmentation variability occurs at the prostate base, where there is much segmentation ambiguity of the prostate gland boundary as a result of poor TRUS image contrast in this region. The axes indicate (A) anterior, (P) posterior, (R) right, (L) left, (S) superior, and (I) inferior orientations with respect to the subject.

Figure 6.

Example synthetic TRUS prostate surfaces used for validation testing. We show a reference prostate shape (gray) first after synthetic non-rigid deformations (blue) and then with synthetic segmentation errors applied to the deformed prostate shapes (orange). For each synthetic deformation, we added synthetic segmentation errors at various levels of noise according to our statistical model of segmentation variability, where noise level 1.00 corresponds to the observed level of noise in clinical practice. Surfaces show the base of the prostate and the labels indicate (A) anterior, (P) posterior, (R) right, and (L) left orientations with respect to the subject.

Figure 7.

The distribution of Dice overlap results first averaged over all volumes of interest for each synthetic example and then averaged across the 1000 synthetically deformed prostate samples. We tested registration using Rigid RPM, Affine RPM, Non-rigid RPM FFD, and our proposed SDM methods, where SDM J indicates that our model used the first J deformation eigenvectors. We show registration results for prostates with different levels of segmentation noise. The boxplots shows the median, 25th and 75th percentiles (interquartile range), extremes, and outlier values.

Figure 8.

Dice overlap results for each volume of interest averaged across 1000 synthetically deformed prostate samples using Rigid RPM, Affine RPM, Non-rigid RPM FFD, and our proposed SDM methods. SDM 11 indicates that our model used the first 11 deformation eigenvectors (accounting for 90% of the model cumulative variance). For each method, we show registration results for prostates with different levels of segmentation noise.

Figure 9.

The distribution of landmark target registration (TRE) error for 29 landmarks in 22 different subjects using our proposed SDM method, where SDM J indicates that our model used the first J deformation eigenvectors. The boxplots shows the median, 25th and 75th percentiles (interquartile range), extremes, and outlier values.

Figure 10.

The distribution of landmark target registration error for 29 landmarks in 22 different subjects for different prostate registration methods. We compare our proposed SDM 17 method (SDM using 17 eigenvectors) to the adaptive focus deformable model (AFDM) method used by an existing clinical biopsy system, to distance map (Dmap) registration using rigid, affine and non-rigid FFD, and to RPM surface-based registration using rigid, affine, and non-rigid FFD transformations. The boxplots shows the median, 25th and 75th percentiles (interquartile range), extremes, and outlier values.

Figure 11.

Example MR-TRUS fusion results showing manually identified landmarks used to calculate target registration error. Here, the TRUS images are warped to the MR image space by a reslicing operation; however, in the clinical setting, the MR surfaces are actually transformed to the TRUS imaging space for guidance. We note that the transformations map points from MRI space to TRUS imaging space, which results in either the image transformation procedure reslicing the TRUS image back into MRI space or the surface/point warping procedure transforming the MR prostate surface/point into the TRUS image's space. The resliced TRUS images are shown for visualization purposes here.

Table 1

Localization counts of landmarks (29 landmarks in 22 subjects) with respect to different prostate zones (AFS, CZ, (PZ), (TZ)) and sections (apex, midgland, and base).

Section	Zones				Total
	AFS	CZ	PZ	TZ	
Apex	-	-	-	-	-
Midgland	1	-	6	13	20
Base	-	2	2	5	9
Whole Gland	1	2	8	18	29

Table 2

The distribution of Dice overlap values (mean $\pm$ SD) for the prostate apex, midgland, and base regions averaged over 1000 synthetically deformed prostate samples. We tested RPM surface-based registration using rigid, affine, and non-rigid FFD transformations, and our proposed SDM method using the first 11 deformation eigenvectors. We show registration results using the standard level of synthetic segmentation noise ($\sigma = 1.0$).

Method	Apex	Midgland	Base
Rigid	79.3 $\pm$ 6.7	82.3 $\pm$ 5.3	81.1 $\pm$ 6.3
Affine	88.0 $\pm$ 4.5	89.1 $\pm$ 3.8	86.7 $\pm$ 4.5
Non-rigid FFD	86.5 $\pm$ 5.0	86.8 $\pm$ 4.8	85.5 $\pm$ 5.0
SDM 11	87.9 $\pm$ 5.0	90.9 $\pm$ 2.7	89.5 $\pm$ 3.0

Table 3

Landmark target registration error (TRE) of the 29 landmarks in 22 subjects. We compared registration using our statistical deformation model (SDM) to robust point matching (RPM) methods using rigid, affine, standard non-rigid FFD transformation models, as well as to an adaptive focus deformable model (AFDM) used in the existing clinical system and distance map-based registration (Dmap) of the surfaces.

Method	Transformation	Mean±SD	Median	Min	Max
AFDM	Non-rigid ^{†,‡}	6.11 ± 4.75	4.24	0.68	19.75
Dmap	Rigid *	4.17 ± 1.50	4.46	0.78	6.92
Dmap	Affine	3.89 ± 1.65	3.86	0.69	7.37
Dmap	Non-rigid FFD	3.90 ± 1.92	3.67	0.67	9.77
RPM	Rigid	3.82 ± 1.81	3.89	0.18	8.84
RPM	Affine	3.90 ± 1.88	3.74	0.56	7.79
RPM	Non-rigid FFD ^{*,‡}	3.39 ± 1.51	3.56	0.61	6.61
RPM	SDM ^{17,‡}	3.50 ± 2.02	2.98	0.82	8.69

^{*}, [†], [‡] indicates statistically significant results with respect to pairwise comparisons at the level of 0.05.

Table 4

Landmark target registration error (TRE) of the 29 landmarks in 22 subjects subject to synthetic TRUS segmentation errors. We compared registration using our statistical deformation model (SDM) to other surface-based robust point matching (RPM) methods using rigid, affine, standard non-rigid FFD transformation models.

Method	Transformation	Mean±SD	Median	Min	Max
RPM	Rigid ^{*,†}	4.07 ± 1.83	4.13	0.23	9.90
RPM	Affine ^{‡,§}	4.24 ± 1.99	3.97	0.21	9.91
RPM	Non-rigid FFD ^{*,§}	3.67 ± 1.61	3.52	0.21	9.43
RPM	SDM 17 ^{†,‡}	3.69 ± 1.63	3.54	0.16	9.32

^{*}, [†], [‡], [§] indicates statistically significant results with respect to pairwise comparisons at the level of 0.05.