Fall 2013

Nonrigid image registration using uncertain surface constraints with application to radiation therapy

Cheng Zhang
University of Iowa

Recommended Citation

Zhang, Cheng. "Nonrigid image registration using uncertain surface constraints with application to radiation therapy." PhD (Doctor of Philosophy) thesis, University of Iowa, 2013.
http://ir.uiowa.edu/etd/2029.

NONRIGID IMAGE REGISTRATION USING UNCERTAIN SURFACE CONSTRAINTS WITH APPLICATION TO RADIATION THERAPY

by

Cheng Zhang

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

December 2013

Thesis Supervisor: Professor Gary E. Christensen

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Cheng Zhang

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Electrical and Computer Engineering at the December 2013 graduation.

Thesis Committee:

Gary E. Christensen, Thesis Supervisor

Joseph M. Reinhardt

Xiaodong Wu

Hans J. Johnson

Mona K. Garvin

To Ping

ACKNOWLEDGEMENTS

First of all, I would like to thank most sincerely my Ph.D advisor, Prof. Gary E. Christensen for his exceptional leadership, insightful guidance, wholehearted encouragement and endless enthusiasm. He is a great thinker and educator. He led me to the field of image registration and paid tremendous attention and care on my research. He introduced to me the spline model for the observational data, which turned out the foundation of the whole work. This dissertation would be impossible without his help and direction.

I am grateful for the support from Prof. Jeffrey F. Williamson and Prof. Martin J. Murphy of Virginia Commonwealth University. As the PI and Co-PI of PPG project (the umbrella project of this work), they were helpful in the development of the proposed image registration method, providing me with excellent clinical expertise.

Special thanks to my thesis committee: Prof. Joseph Reinhardt, Prof. Xiaodong Wu, Prof. Hans Johnson and Prof. Mona Garvin for their valuable comments and suggestions. I also would like to thank Prof. Punam Saha and John Bayouth for serving as committee members in my proposal defense.

I would like to thank Dr. Elisabeth Weiss, Dr. Jian Wu, Yichao Wang, Ford Sleeman of Virginia Commonwealth University, for providing the patient data and help in the study of segmentation error. I would like to thank Prof. Erwei Bai for his explanation on the sampling theory, Dr. Youbing Yin for his help with the B-spline image registration, Prof. Grace Wahba of University of Wisconsin-Madison for answering my question about the proof in her spline book, Prof. Yuedong Wang of University of California, Santa Barbara for his open-source package of smoothing spline regression, Prof. Sebastian Kurtek of Ohio State University for helping me to understand the idea of shape space.

Sincere thanks go to the fellow students in Prof. Christensen's research group, Kunlin Cao, Joo-hyun Song, Ying Wei, Weichen Gao, Matineh Shaker, Jeffrey Hawley, Bowen Zhao and Teresa Stout. I would also like to acknowledge the help I received from my friends in the Department of Electrical and Compute Engineering and other departments.

Last but not least, I would like to thank deeply my wife, my son, my parents, my parents-in-law and my brother for their love, patience, hope and encouragement.

This work was supported by the NCI Grant No P01CA116602.

ABSTRACT

An important research problem in Image-Guided Adaptive Radiation Therapy (IGART) is how to accurately deform daily onboard Cone-Beam CT (CBCT) images to higher quality pretreatment magnetic resonance (MR) or Fan-Beam CT (FBCT) images, enabling cumulative dose to be evaluated and tumor response to be tracked. In the case of IGART for prostate cancer, the question becomes to accurately register the critical organs such as bladder, prostate and rectum. All are soft tissues and their boundaries can not consistently be identified using CBCT. As such it is challenging to register these soft organs precisely if the intensity difference serves as the only similarity measure.

Organ surfaces are routinely contoured as part of standard treatment planning procotol. In this work we assume that the organ surfaces are provided either by manual or automatic segmentation and can be used to improve the correspondences at structure boundaries. Unfortunately, these standard segmentations are often inaccurate so that the inclusion of the surfaces into the registration process may give little improvement.

Originating from this specific problem, this work tries to answer a more generalized question. Given two intensity images and their associated inaccurate object surfaces, can we design a non-rigid registration algorithm with improved registration accuracy? Influenced by the ideas of data assimilation (DA) and smoothing spline regression (SSR), this report provides a solution consisting of three components: statistical shape modeling, spline-based surface estimation, and surface constrained non-rigid image registration.

We surveyed several distinct surface registration algorithms and evaluated their performance on patient data. The shape models of the pelvic organs were built using training data. For image registration, the input surface is a combination of the current observed and the one predicted by our shape model. This hybrid surface was validated to be more accurate and therefore the image registration produced smaller registration error. Experiments were performed using both simulated data and real clinical data. Results show that the proposed method achieves satisfactory improvement.

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
INTRODUCTION

Image registration is routinely used to compare images between individuals, detect and measure changes between longitudinal images of the same individual, fuse information from different imaging modalities, and many other applications. Large registration error can occur around the region of interest (ROI) if intensity contrast at the boundary of the ROI is insufficient. To overcome this problem, a priori information of the ROI can be used to guide the registration process. For example, labeled landmarks, 2D contours or 3D segmentations can serve as surrogate features to augment and improve the intensity-only image registration. However, these geometric features often contain errors at the boundary too. It is well known that interobserver variability varies as a function of modality, organ, contour orientation, surface position of the object, and training.

For many registration algorithms, the uncertainty with the identification of the features is simply ignored. Segmentations are assumed accurate and the similarity between them are directly used for registration. For certain applications, doing so is legitimate if the features, even though having errors, are accurate enough in the sense that they provide correct matching clues at the regions with low intensity contrast. However, in cases where the segmentation error can not be ignored, the former strategy will bring a new source of errors into the registration process. The registration result becomes unstable.

This work makes an attempt to improve registration accuracy under such a situation. The primary application area of this dissertation is image-guided prostate cancer radiotherapy and this application requires accurate and robust non-rigid image registration of the pelvic anatomy. The boundaries of prostates are delineated by an expert but with large interobserver segmentation error. Details are given in Sec. 2.5.3.

The basic idea proposed by this work is called surface assimilation, a method of estimating the true surface from a priori observation data and the current observation. Inspired by the similar concept of Data Assimilation from the subject of weather predication [30], we obtain the general knowledge of the surface from statistical shape analysis. The a priori knowledge about the shape are integrated with the observed surface under the spline-based estimation framework for noisy observational data. Loosely speaking, we reconstruct a surface with small segmentation error by correcting an observed surface with large segmentation error using an a prori shape model. The surface estimation problem is formulated as a spline-based optimal estimation problem.

The proposed algorithm works on 1D parametric domains (curve) and 2D parametric domains (cylinder and sphere). We claim that the reconstructed surface is a better representation of the underlying anatomy. Therefore image registration involving a pair of more accurate surfaces will lead to less registration error. We choose the problem of registering prostate CT images as the working example throughout the paper. It is worth mentioning, however, that our method is general and can be applied to many body sites and different modalities.

1.1 Background

External beam radiotherapy (EBRT) is the most common treatment option for prostate cancer currently [103]. The patient lies on a movable table and an external source of radiation is pointed at a particular part of the body, delivering therapeutic radiation dose to the tumor. In order to compensate for the various uncertainties in the treatment delivery, healthy tissue surrounding the tumor is inevitably irradiated. However the tolerance to the radioactive rays of healthy tissue and the tumor are often different. Even though DNA of both cancer cells and the healthy ones will be damaged, the recovery rate of healthy ones is faster than those in the tumor. Normally the prescribed dose will be delivered in fractional amounts during multiple treatment sessions with the purpose of allowing healthy tissue to recover while still providing lethal dose to the cancer cells. In the treatment planning phase, diagnostic imaging, such as CT or MRI, is collected. Critical organs, such as bladder and rectum, together with the tumor site are contoured by experts and are fed into the planning system to generate a radiotherapy plan.

1.1.1 Uncertainty in Radiation Therapy

If dose can be delivered accurately to the tumor, a higher dose (dose escalation) can be administered so that the treatment session (typically one or two weeks in case of prostate radiotherapy) can be shortened, more healthy tissue will be spared from the unintentional exposure and the patient survival rate is expected to improve. However great attention must be paid to avoid increased toxicity to normal tissues in an attempt of dose escalation. There are many sources of uncertainty in radiotherapy. Among them the geometric uncertainty is perhaps one of major risks to dose escalation. Geometric uncertainty can be roughly classified into intrafractional or interfractional uncertainty. Intrafractional uncertainty is the error that occurs within a treatment fraction, such as the scattering uncertainty in radiation rays, bladder shape changes due to voiding. Interfractional uncertainty is the dominating error that happens from day to day, such as the table setup errors and misalignment errors due to organ motions.

Large interfractional uncertainty impacts the reproducibility of dose delivery from day to day causing tumors to be underdosed and the nearby critical organs overdosed. To reduce the these errors, image-guided radiotherapy (IGRT) [39] was invented. With the integrated on-board, low-dose, fast cone-beam CT (CBCT) imaging capability, a tumor image can be acquired before each treatment fraction. Rigid registration of treatment-day and planning-day images enables physicians to align the patient table to assist the localization of tumors. Meijer et al. [94] showed that the prostate planning margin can be reduced from 11 mm to 8 mm when aligning the patient using rigid registration to align bony anatomy. However this improvement would be compromised by the fact that the tumor and pelvic organs, as soft tissue, may undergo large shape changes in position, size and shape over the course of radiotherapy.

1.1.2 Nonrigid Image Registration in Radiation Therapy

Nonrigid image registration is capable of estimating voxel-by-voxel correspondence between two images and has been studied recently with rising interests in IGRT aiming to lower the interfractional uncertainty. By registering daily CBCT images to a higher quality pre-treatment fan-beam CT (FBCT) image, nonrigid image registration provides a means to compute cumulative treatment dose and track tumor response. In Image-guided Adaptive Radiotherapy (IGART), 3D image data is registered to construct 4D voxel trajectories (i.e., time-dependent DVFs) that describe deforming patient anatomy over a course of radiation therapy [97, 135, 6]. By doing so an initial treatment plan can be modified on a daily basis to accommodate the shape changes of delicate organs, such as bladder, prostate and rectum.

Nonrigid image registration algorithms based exclusively on intensity difference have been applied to registering FBCT and CBCT pelvic images [54, 86]. However due to the low contrast of soft tissue in CBCT imaging, the intensity difference along the boundaries of the prostate and other organs are poor. Therefore the results of the solely intensity-driven registration algorithms are often not accurate.

Various approaches have been proposed for improving the accuracy. Based on the geometric entities used, these approaches can be roughly classified into three categories, namely landmarks-, label- and surface-based. Using the SIFT feature detection technique, Chao et. al [18] extracts the automatic landmarks from the source and target bladder images to guide the registration process. Label images (binary masks), when available from either manual or automatic segmentations of organs can be combined with the intensity feature to aid the registration results at the low contrast region [64] or account for large organ deformation [21]. For surface-based approaches, [70, 102, 78] register the source and target surfaces first and then use the obtained transformation as the initial condition or extra constraint in volume registration.

Studies have been done on the automatic FBCT/CBCT registration problem. Zhou *et al.* [144] proposed registering two images after an automatic prostate segmentation based on Active Shape Models (ASM). With the focus on single organ object, this method contains two steps: the segmentation step and the registration step. Lu *et al.* [87] proposed a joint segmentation and registration algorithm in which the bladder, prostate and rectum are segmented by level-set method.

It worth mentioning that a good nonrigid image registration algorithm in IGRT needs to be fast, accurate and robust. Although nonrigid registration algorithms are generally more time consuming than rigid ones, recent progress in parallel computing makes it possible to reduce the running time of nonrigid image registration [69, 113]. The major obstacle to wide adoption of nonrigid image registration lies in the difficulty in increasing the accuracy and robustness of currently available algorithms.

1.1.3 Interobserver Segmentation Error

Landmark-based registration [108, 137] have taken into account the landmark positioning error leading to more robust algorithms [25]. However, the existing surface and intensity-driven nonrigid image registration methods assume the segmentations are accurate. This is generally not true for CBCT images where the boundaries of the prostate and the surrounding organs at risk (OARs) are often difficult to distinguish from surrounding tissue. Previous studies [104] have shown that interobserver segmentation errors (ISE) vary dramatically with location on the prostate surface and OARs. Recent studies [135, 89] suggest the inter-observer delineation variability of CBCT needs to be considered in the CBCT-based adaptive radiotherapy. If the inaccurate segmentations are used in image registration, large registration error will occur. Therefore, one of the major tasks of this study is to propose a new nonrigid image registration algorithm with full consideration of organ segmentation uncertainty aiming to find a more accurate image alignment.

For a long time, the study of estimating the interobserver segmentation error has been lacking. It is perhaps partially due to the absence of reliable tools of 3D surface analysis. The recent development in 3D surface registration make it feasible to qualify the segmentation errors, therefore facilitate the design of new image registration algorithm capable of minimizing the impact of segmentation uncertainty. Initial efforts were made in [138] to estimate the reliability of the organ boundary delineation, nevertheless, how to make use of this reliability information in image registration remains unclear.

1.2 Organization of Dissertation

Chapter 1 describes the background and motivation of this thesis. We briefly present the idea of image-guided adaptive radiation therapy (IGART) and discuss the opportunities and challenges for deformable image registration in IGART. From there, the theme of this dissertation, a general problem dealing with the uncertainty in surface segmentation and image registration, is defined.

Chapter 2 introduces the idea of statistical shape model. After reviewing various surface registration techniques, we present our work on choosing the appropriate methods of shape modeling for the pelvic organs. We also explore the possibility of applying surface registration to study the interobserver segmentation error.

Chapter 3 introduces the theory of data assimilation and spline model for observational data. We combine these theories, together with the statistical shape model, to solve the problem of reconstructing 2D curve and 3D surface from noisy observed data. Proof-of-concept experiments are done to demonstrate the ability of the algorithm.

Chapter 4 describes the application of the proposed algorithm to the real patient data from the IGART project. We design a nonrigid volumetric registration method by combining the intensity and surface similarity. Results from evaluation experiments further validate the performance of the proposed methods.

Chapter 5 concludes this dissertation and summaries the major contributions. It also discusses the outlook of the future work.

Appendix A provides more information on the patient data used in the experiments of this dissertation.

CHAPTER 2
SURFACE REGISTRATION AND STATISTICAL SHAPE MODEL

2.1 Introduction

A good understanding of organ shapes and inter-observer segmentation variability can be used to design image segmentation and registration algorithms. As mentioned in Chapter 1, organ segmentation among experts is not uniformly consistent and contours can have uncertainties that vary with the position on the organ surface. Fig. 2.1 shows a pair of FBCT / CBCT images from one patient that were repeatedly traced by five different experts (Fig. 2.2 and Sec. 2.1 show these segmentations in 3D). From the overlaid prostate contours, it is easy to see that the inter-observer segmentation errors (ISE) in CBCT are larger than those in FBCT, plus the prostate base seems to have larger ISE among experts within a CBCT image. If arbitrary expert segmentation is used in registration process, this large segmentation uncertainty will lead to large registration uncertainty and may pose a risk to patient. So what can we do to reduce the negative effect of ISE in the image registration?

This study addresses ways to reduce negative effects of ISE through two steps. First, we need a method to quantify ISE. This study uses surface registration and Principal Component Analysis (PCA) to do so. Second, we use the organ shape model and the knowledge about ISE to reconstruct a new surface closer to the true one, the main idea of Chapter 3. Therefore, we also need methods to summarize the organ shape statistics. Due to the similarity of both problems, we choose a similar surface analysis method as that for ISE quantification.

In this chapter we survey the common surface registration techniques, evaluate the candidate algorithms, summarize the organ shape statistics and estimate the segmentation errors at the sites of prostate, bladder and rectum, respectively. The output of this chapter is to provide the appropriate surface analysis tools for the following chapters.

2.2 Surface Registration: A Survey

Image registration is the enabling technique for statistical shape analysis, anatomical atlas building, longitudinal study, computer vision, pattern recognition and knowledge discovery. Image registration has been widely used in the medical image analysis. This section gives a brief review of surface registration algorithms with bias on those commonly used in shape model building[1]. For an earlier review of surface registration, see [5]. Surface registration algorithms can be grouped into three categories: pointset, mesh and volume. For each category, the methods for pair-wise surface registration will be reviewed. The group-wise surface registration is based on these pair-wise methods. The representative group-wise algorithms will be discussed in Sec. 2.3.

2.2.1 Surface Registration via Landmark Matching

A surface can be represented by a set or cloud of points. A typical landmark-based surface registration needs to know the corresponding points in template and

---

[1] Geng [62] made substantial contribution to this section

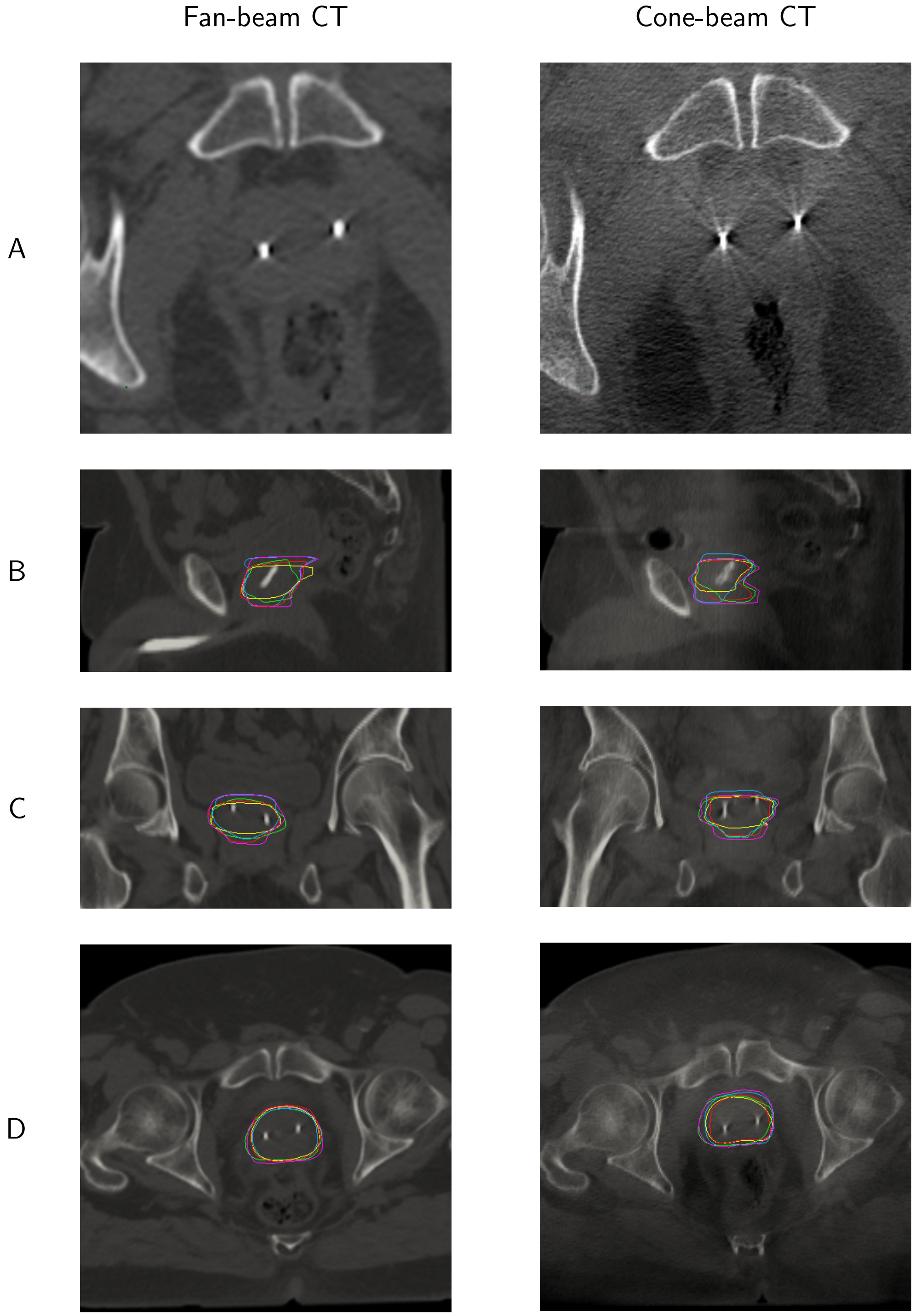


Figure 2.1: One patient's FBCT/CBCT images in the study. From left to right FBCT and CBCT grayscale image with the prostate contoured by five experts marked with different colors. From top to bottom: A. The original image, B. Sagittal view, C. Coronal view, D. Transverse view.

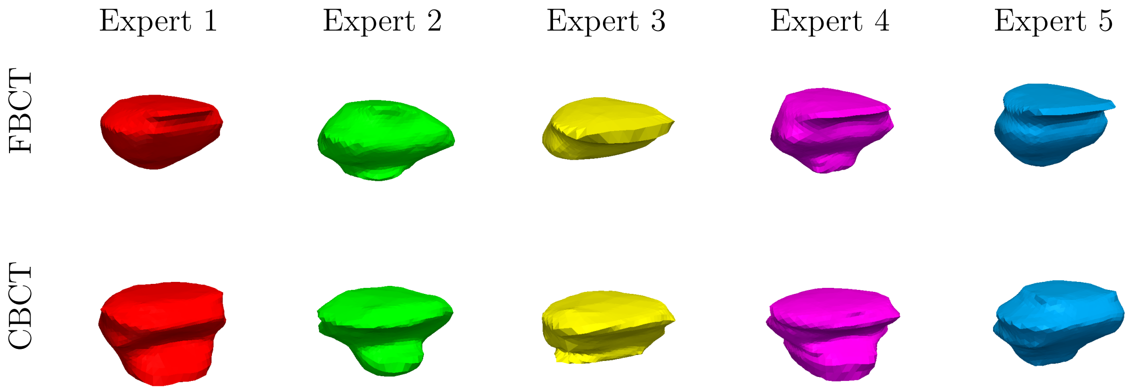
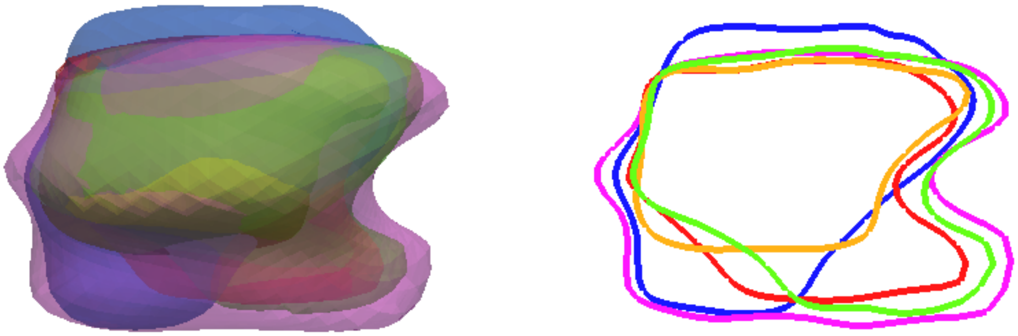


Figure 2.2: A patient's FBCT/CBCT prostate segmentations in the study. From top to bottom: FBCT and CBCT prostates segmented by five experts marked with different colors. These meshes are reconstructed from stacks of 2D manual contours. Large segmentation variation among experts are visible at the base of the prostate.



(a) 3D view of prostate segmentations (b) 2D view of prostate segmentations

Figure 2.3: 3D and 2D overlays of five prostate segmentations from the same patient contoured by five experts on the same day CBCT image.

target images before solving for the spatial transformation. Iterative Closest Point (ICP) [8] is a classic rigid landmark matching algorithm. It automatically estimates the correspondence between two point sets and updates the correspondences at every step of minimizing the distance between the two. While simple and fast, the plain ICP is sensitive to noise and initial configuration. ICP-derived methods [111, 92, 47, 112] relax the correspondence constraint and model landmark matching as a probability density estimation problem to improve its robustness.

As the seminal nonrigid landmark registration algorithm, the Thin Plate Spline (TPS) method introduced by Bookstein [10] estimates a smooth transformation between the spaces of two surfaces. It assumes the correspondence between the two point sets is known and therefore suffers from the uncertainty of landmark locations. Rohr [108] proposed a robust TPS method by considering the landmark uncertainty in registration cost function. Chui and Rangarajan [25] combined TPS and the robust point matching (RPM) to search for the unknown correspondence. Dalal [29] used TPS to measure the distance between two surfaces. Davis [38] and Worz [137] introduced Elastic Body Spline (EBS) and Gaussian Elastic Body Spline (GEBS), respectively, aiming to find more physics-based transformations. Wang [133] proposed a shape-based 3D surface correspondence approach using geodesics and local geometry.

Besides the spatial coordinates of the landmarks, other geometric features, if salient and reliable, can be used as the matching driving force. For example, in brain cortex registration problem where the sulci and gyri present salient local geometric characteristics. Curvatures and normal vectors of the crest lines (ridges and valleys) can constrain the registration under either ICP or TPS framework [119].

Partially due to the coarse discretization of a surface as a point set, it is hard to ensure that the deformed point set lies exactly on the target surface. This could be a major obstacle for the landmark-based methods to be used in complicated 3D shape analysis.

2.2.2 Surface Registration using Parameterized Surfaces

Surface registration between two parameterized surfaces is suitable for the applications requiring fine and dense correspondence on surfaces. The basic idea is to view the source surface X as a mapping from a (u, v) parameter domain into a 3D Euclidean space as $X : u \times v \to \mathcal{R}^3$ and the target surface $Y : u \times v \to R^3$. Surface matching features (e.g. curvature) are transferred to the parametric space and two parametric surfaces are aligned over the common (u, v) domain using various registration techniques. Based on their topological difference, open surface is normally flattened to a 2D (u, v) domain where $u = [0, 1]$, $v = [0, 1]$, while a closed surface is mapped to a unit sphere S^2.

Many surface parameterization methods have been proposed (for a detailed review see [53]). Some [48, 49] of them are based on minimization of the metric distortion. Some focus on surface areal distortion minimization [43, 46]. Some are based on harmonic maps [142]. The methods of conformal parameterization [85, 67, 66] maps the similar triangles on the original surface to the similar triangle on the sphere S^2. It is bijective angle-preserving, independent of triangulation of surfaces making it widely used in surface parameterization.

In the parametric space, surface registration tries to find a re-parameterization (or transformation field) $\gamma(u,v)$ optimizing the correspondence between $X(u,v)$ and $Y(\gamma(u,v))$. Since the domain of the transformation field is a unit sphere instead of Euclidean space, covariant derivatives are often used to account for the intrinsic geometry of the surfaces. Some approaches parameterize a transformation field, restricting the search space of the allowed re-parameterization (e.g., Cauchy functions [35]). Joshi *et al.* [78] presented a method of registering cortex surfaces via thin-plate spline (TPS) method on the flattened surfaces. Horkeaw [73] parameterized a transformation field with B-splines on 2D shape manifold. Some methods allow for large deformation due to the use of non-parametric re-parameterization. Tosun *et al.* [123] aligned cortical surfaces driven by optical flow. Twining *et al.* [126] extended the fluid volumetric image registration [24, 23] to shape manifold. Yeo *et al.* [140] proposed the Spherical Demons registration algorithm which is an extension of Diffeomorphic Demons [121, 130] method on sphere.

2.2.3 Surface Registration via Measure Matching

Like parametric surface registration, i.e. aligning two surfaces without reducing surface (2D shape manifold) to zero-dimensional point set, Glaunès and Vaillant [63] presented surfaces as measures of distributions, and surface similarity is defined into a Hilbert space endowed with a computable norm. The surface registration is solved as a variational problem under a framework of large deformation diffeomorphism.

Vaillant and Glaunès [128] extended the surface measure to a generalized distribution called currents from the theory of geometric measure. The currents representation preserves the geometry of the structure because both location and the first order local geometric feature appear in currents. Vaillant [129] used these methods to help the study of brain cortex structures.

2.2.4 Surface Registration via Level sets

Huot and Yahia [139, 75] proposed surface matching algorithm based on level sets. This method projects a surface into a cost hypersurface space and matches two surface by finding the surface propagation under a level set formulation. It can handle curvature singularities, large deformation and arbitrary topology. Luethi [91] represented surfaces using distance map and level set where the dense correspondence between two surfaces are estimated by a modified Demons algorithm [121] with curvature driving force.

2.2.5 Surface Registration via Image Matching

Nonrigid image matching (or registration) seeks a point-wise correspondence between two images by minimizing a cost function of the similarity between images. Choosing the similarity measure is a major part of nonrigid image registration algorithm. For the images from the same imaging modality or their intensity relationship is known, the squared intensity error, the difference of intensity gradient or the intensity correlation coefficients (CC) can be used as the similarity measure. For multi-modality image registration, the mutual information (MI) [136, 90] from information theory has been used as the similarity measure. MI usually are sensitive to the initial overlap between the targets. Studholme [115] proposed a normalized mutual information (NMI) to overcome this issue. Pluim [101] and Zitova [146] gave two comprehensive reviews on mutual information registration.

With the closed surfaces rasterized as binary label images, nonrigid image matching can establish the correspondence between two surfaces. Frangi [56] used image-based affine transform and nonrigid B-spline registration to model 3D cardiac Active Shape Models. Taron [120] used the deformation fields to model different shapes. One of the advantages of this method is that multiple surface objects can be analyzed together. The weakness is that the nonrigid matching generally require more computational resources. Besides, the deformed segmentation may not fully overlap the target creating additional errors.

2.3 Group-wise Surface Registration

A statistical shape model is built from a group of training surfaces. The quality of the correspondence heavily affects the quality of the shape model. If the correspondence is unknown, we can estimate it using the group-wise registration. Group-wise registration methods generally involve a global criterion measuring the group-wise correspondence and can be roughly grouped into two categories: 1) Explicit template-based group-wise registration and 2) Implicit template-based group-wise registration. The explicit ones require a template surface to be selected and every other surface is registered to the template pair-wisely. Bias can occur in selecting the template, i.e., the template can be unrepresentative about the population. Therefore, the shape model could be biased. Frangi [56] registered surface iteratively in a natural coordinate system to reduce the bias, but it could not be removed. On the other hand, the implicit registration registers all inputs to a common and hidden coordinate system, so is unbiased. In the following section we introduce six representative group-wise surface registration methods. Four of them were evaluated in this study and MDL was chosen to build the shape model of the prostate for future experiments.

2.3.1 Group-wise Registration using SPHARM First Order Ellipsoid

SPHARM [12] is a smooth, resolution-controllable shape descriptor. By defining the Fourier expansion on spherical domain, a closed surface $\mathbf{x}$ can be represented by a series of Spherical Harmonics (SPHARM) in Eq. (2.1).

$$\mathbf{x}(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \mathbf{c}_l{}^m \mathbf{Y}_l{}^m(\theta, \phi) \tag{2.1}$$

where (θ, ϕ) is the coordinate on the sphere domain, $\mathbf{Y}_l{}^m$ denotes a spherical harmonic function of degree l and order m and $\mathbf{c}_l{}^m$ is the corresponding coefficient.

By optimizing an equal area mapping and minimizing angular distortions [12], a shape object is mapped from the 3D quadrilateral voxel mesh to the unit sphere. From there the coefficients for spherical harmonics are computed to represent the original shape object. With these harmonics defined naturally on (θ, ϕ) domain, it is straightforward to obtain a spherical parameterization of a surface. With a icosahedron subdivision on the sphere, the original shape can be represented as an uniformly distributed landmark mesh [80]. Both representations are extensively used in the experiment of surface registration.

The correspondence of a group of surfaces can be estimated by aligning their SPHARM parameterization such that the ridges of the first order ellipsoid coincide. One problem of this method is that the rotational symmetry of first order ellipsoid is undistinguishable. Arbitrary matching the ridges may weaken the quality of the correspondence [117].

2.3.2 Particle-based Group-wise Registration

Following the strategy of minimizing information content across an ensemble of surfaces, the particle-based groupwise registration offers a non-parametric correspondence finding method [17]. Each surface is associate with a particle system, represented as a pointset $z = (x_1, x_2, \ldots, x_N)$, treated as a realization of a random variable Z. Suppose every point x_i with the PDF function $p(x_i)$ defined as

$$p(x_i) = \frac{1}{N(N-1)} \sum_{\substack{j=1 \\ j \neq i}}^{N} G(x_i - x_j, \sigma) \tag{2.2}$$

where $G(x_i - x_j, \sigma)$ is a isotropic Gaussian with standard deviation σ. The amount of information carried by a surface particle system z is therefore given as

$$\begin{aligned} H(z) &= -\int p(z) \log(p(z))\, \mathrm{d}z \qquad (2.3) \\ &\approx \sum_{i} \log \frac{1}{N(N-1)} \sum_{j, j \neq i} G(x_i - x_j, \sigma) \qquad (2.4) \end{aligned}$$

For a collection of M surfaces, with their particle ensemble $\mathcal{E} = z^1, z^2, \ldots, z^M$, we can place these particles row by row into a matrix and the points with the same indices (each column) can be modeled as a random variable. The information content across $\mathcal{E}$ can be computed as $H(\mathcal{E})$.

$$H(\mathcal{E}) \approx \frac{1}{2}\log|S| = \frac{1}{2}\sum_{j}\log\lambda_{j} \tag{2.5}$$

where λ_{j} is the eigenvalue of the shape covariance matrix S (see Eq. (2.15)).

For the group-wise surface registration, the cost function

$$Q = H(\mathcal{E}) - \sum_{k=1}^{M} H(z^k) \tag{2.6}$$

is minimized by finding $\hat{\mathbf{z}} = (\hat{z}^1, \hat{z}^2, \ldots, \hat{z}^M) = \arg\min_{\hat{\mathbf{z}}} Q$ with the constraint that $\hat{z}^1, \hat{z}^2, \ldots, \hat{z}^M$ are on the original input surfaces. Minimizing the first part of Eq. (2.6) leads to a compact shape model. The summation terms favor a equidistantly distributed particle coverage of the surfaces which is desirable for accurate surface representation. The particle-based method requires no parameterization on surface therefore can be used for a wide range of surface topology, such as open surface, closed surface and multi-objects [16].

2.3.3 Group Registration based on Minimum Description Length

The minimum description length (MDL) can serve as a global criterion in group-wise surface registration for statistical shape modeling. Inspired by the principle of Occam's razor that the "best" model should describe the entire training set most efficiently, Davies *et al.* [33] proposed encoding the training set with the minimum description length $\mathcal{L}_{total}$ consisting of two components:

$$\mathcal{L}_{total} = \mathcal{L}_{param} + \mathcal{L}_{data} \tag{2.7}$$

When $\mathcal{L}_{total}$ is minimized, a good surface correspondence is obtained. This method assumes the point distribution model (PDM) [27], a special multivariate Gaussian model, can describe the surface population.

To overcome the slow convergence of the original method, Thodberg [122] proposed a simplified and efficient MDL algorithm where $\mathcal{L}_{total}$ is defined as:

$$\mathcal{L}_{total} = \sum_m \mathcal{L}_m \quad \text{where} \quad \mathcal{L}_m = \begin{cases} 1 + \log(\lambda_m/\lambda_{cut}) & \text{for } \lambda_m \geq \lambda_{cut} \\ \lambda_m/\lambda_{cut} & \text{for } \lambda_m < \lambda_{cut}. \end{cases} \tag{2.8}$$

λ_{cut} is a cut-off value deciding which modes are considered as systematic variations and which ones are noise. λ_m corresponds to the m^{th} eigenvalue of the local shape features. $\lambda_{\,cut}$ and $\lambda_{\,m}$ reflect the code length of $\mathcal{L}_{param}$ and $\mathcal{L}_{data}$, respectively.

Heimann [72] improved the computation speed further by choosing Gaussian envelop function to represent the correspondence $\phi(u, v)$ and using the shape index S and the curvedness C as the local shape features, defined as

$$C = \frac{2}{\pi} \ln \sqrt{(\kappa_1^2 + \kappa_2^2)/2}, \qquad S = -\frac{2}{\pi} arctan \frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2} \tag{2.9}$$

where κ_1 and κ_2 are the two principal curvatures. C and S improve the curvature measurement by decoupling the size and shape aspects of the curvature. C indicates how curved a surface is. S describes the concaveness and convexness of the surface.

2.3.4 Groupwise registration in Metric Shape Space

Surface registration may be sensitive to the parametrization of surfaces, i.e. the different parameterization of surfaces results in the different correspondence. To overcome these limitations, Kurtek [82] proposed a Riemannian framework for surface registration. Each surface is, by a special representation, q-map, mapped into a metric shape space. The distance between two surfaces in this space is invariant to rigid motion, global scaling, and re-parametrization. Due to the property of this $\mathbb{L}^2$ metric, the deformation of the parametric mesh on the fixed surface follows the geodesic between the two surfaces. The registration result is, therefore, inverse-consistent.

For surface registration on the spherical domain, an algorithm starts by parameterizing a closed surface making it a 2D manifold represented by $u-$, $v-$parameters. Many algorithms match two surfaces in this domain and the obtained deformation field is defined in its tangent space. There is an issue of handling the parameterization variability. In case of conformal mapping, it is not fully automatic and the north pole has to be specified manually. Selecting a different north pole will result in different spherical parameterizations. Suppose $Y(u,v)$, $X(u,v)$ are two parametric surfaces, $\gamma(u,v)$ is an arbitrary re-parameterization. It is obvious that $\|Y(u,v) - X(u,v)\| \neq \|Y(\gamma(u,v))) - X(\gamma(u,v))\|$. It means even though the shapes of $Y(u,v)$ and $X(u,v)$

are fixed, the different parameterization of Y and X will result in different distances. The variability of parameterization will eventually compromise the quality of shape model. In the parameterization-invariant shape analysis technique [83] $Y(u,v)$ and $X(u,v)$ will be indirectly represented by q-maps $Q(Y)$ and $Q(X)$ with guaranteed property $\|Q(Y(u,v)) - Q(X(u,v))\| = \|Q(Y(\gamma(u,v))) - Q(X(\gamma(u,v)))\|$. It implies that the surface-to-surface correspondence remains fixed no matter what the initial surface parameterization is.

In this metric shape space we can further compute the geodesics between parametric surfaces [84]. With this ability, it is claimed the Karcher mean and Karcher variance of shapes lead to more accurate statistics computing of shape population [84].

2.3.5 Transitive inverse-consistent surface registration

Geng *et al.* [62] developed a group-wise registration method, called transitive inverse consistent manifold registration (TICMR). It jointly registers a group of three manifolds embedded in a higher dimensional image space and needs no common reference frame. TICMR works for landmarks, surfaces and images. For the surface registration, TICMR uses the iterative closest point (ICP) and linear interpolation on the triangulated mesh to approximate a dense correspondence. A minimizer of elastic energy on the surface provide extra regularization. Including the inverse consistency error (ICE) [22] and the transitivity error (TE) [20] into the cost function, TICMR can reduce certain local mis-alignment that occurs in uni-directional method.

TICMR can register a group of N images simultaneously by exhaustively pairing images up. However it is prohibitive as the computational complexity will rise exponentially. To make it computation feasible, a clustering TICMR was developed in [61]. This method clusters the population into groups of 3. If the group number is larger than 3, cluster the subgroups to group of three recursively until every group contains at most three surfaces. We can optimize the set of transformations in each group using TICMR.

2.3.6 Group-wise image registration

As an extension to the pair-wise image registration (surveyed in Sec. 2.2.5), the group-wise image registration can be used to study the shapes population. Cabezas [14] reviewed the groupwise image registration for atlas building. In the implicit group-wise registration, the global transformations are estimated simultaneously. Twining *et al.* [125] used the minimum description length (MDL) to guide the group-wise non-rigid image registration. Studholme [114], Bhatia *et al.* [9] and Zhang and Rangarajan [143], extending the concept of mutual information in pair-wise multi-modality image registration, created a measure for population alignment. These methods are computationally expensive since the computation increases exponentially when the number of the input images increase linearly. Zollei *et al.* [147] proposed an efficient registration method, *congealing* technique, where the objective function avoids the expensive computation of the joint density function when the number of inputs is large.

2.4 Statistical Shape Models

Statistical shape models (SSM) carry the shape knowledge of the anatomical objects from the observed medical image datasets. They provide a promising means to constrain, drive or predict in various medical imaging applications such as image segmentation, registration and interpretation. To build a SSM, one needs to capture the average shape and discover the shape variation modes from a given training population.

2.4.1 Methods

This section follows the concept of Point Distribution Model proposed by Cootes *et al.* [27]. Suppose the training dataset contains M shapes $\{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^M\}$, and the group-wise correspondence is known (e.g., via the group-wise surface registration). Each shape consists of N pairs of evenly distributed sample points. The i-th shape is represented as:

$$\mathbf{x}^i = (x_1^i, y_1^i, z_1^i, x_2^i, y_2^i, z_2^i, \ldots, x_N^i, y_N^i, z_N^i)^T \tag{2.10}$$

and the sample matrix X is defined as:

$$X = \begin{pmatrix} x_1^1 & x_1^2 & \ldots & x_1^M \\ y_1^1 & y_1^2 & \ldots & y_1^M \\ z_1^1 & z_1^2 & \ldots & z_1^M \\ \vdots & \vdots & & \vdots \\ x_N^1 & x_N^2 & \ldots & x_N^M \\ y_N^1 & y_N^2 & \ldots & y_N^M \\ z_N^1 & z_N^2 & \ldots & z_N^M \end{pmatrix} \tag{2.11}$$

The mean shape is obtained by averaging X row by row:

$$\bar{\mathbf{x}} = (\bar{x}_1, \bar{y}_1, \bar{y}_1, \bar{x}_2, \bar{y}_2, \bar{y}_2, \ldots, \bar{x}_N, \bar{y}_N, \bar{z}_N) \tag{2.12}$$

where

$$\bar{x}_j = \frac{1}{M}\sum_{i=1}^{M} x_j^i, \qquad \bar{y}_j = \frac{1}{M}\sum_{i=1}^{M} y_j^i \quad \text{and} \quad \bar{z}_j = \frac{1}{M}\sum_{i=1}^{M} z_j^i \tag{2.13}$$

To capture the shape variation, we subtract the mean shape from each shape:

$$\delta\mathbf{x}^i = \mathbf{x}^i - \bar{\mathbf{x}} \tag{2.14}$$

And we obtain the covariance matrix of size $3N \times 3N$:

$$S = \frac{1}{M}\sum_{i=1}^{M} \delta\mathbf{x}^i (\delta\mathbf{x}^i)^T \tag{2.15}$$

Solving the eigen-decomposition equation

$$S\mathbf{p}^i = \lambda_i \mathbf{p}^i \tag{2.16}$$

gives the eigenvectors $\{\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^{3N}\}$ and the corresponding eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_{3N}\}$ of S. S is then decomposed into $3N$ shape variation modes. Each eigenvector $\mathbf{p}^i$ captures one of these (directions), and λ_i tells its significance. The eigenvectors of S serve as a orthogonal basis of the space of shape. Given a particular shape $\mathbf{x}$, it can be represented by:

$$\mathbf{x} = \bar{\mathbf{x}} + P\mathbf{b} \tag{2.17}$$

where $\mathbf{b} = (b^1, b^2, \ldots, b^{3N})^T$ is a vector of weights, defining a coordinate in the parameter shape space, $P = (\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^{3N})$ maps the shape parameter to the original shape space and the mean shape translates the mapped one to the origin.

The eigenvalue λ_i of S indicates how much variation appears on the particular direction, the axis represented by $\mathbf{p}^i$. Thus, we can sort the eigenvalues from large to small, and choose the first $t(< 3N)$ ones to approximate a shape:

$$\mathbf{x} \approx \bar{\mathbf{x}} + P_t\mathbf{b}_t \tag{2.18}$$

where $P_t = (\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^t)$ and $\mathbf{b}_t = (b^1, b^2, \ldots, b^t)$. One approach to choosing t is to compare the sum of variance up to t and the total variance $\lambda_{total} = \sum_{i=1}^{3N} \lambda_i$ such that:

$$\sum_{i=1}^{t} \lambda_i \geq \alpha\lambda_{total} \qquad 0 < \alpha < 1 \tag{2.19}$$

where the choice of α depends on the requirement of the application. Usually, a small number of modes can describe most of the shape variation, leading to a compact shape model.

Carrying the knowledge of shape it models, the Eq. (2.18) can generate new examples of shapes by varying the random variable $\mathbf{b}_t$. When $\mathbf{b}_t$ is assumed as a multi-variate Gaussian random variable [27], a proper range of b_i can be:

$$-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i} \tag{2.20}$$

The PDM can be generalized to continuous domains. For a smooth parametric curve (1D manifold), the shape probabilistic model Eq. (2.18) can be defined as

$$\hat{\mathbf{x}}(t) = \bar{\mathbf{x}}(t) + \sum_{i=1}^{M} b_i \mathbf{p}_i(t) = \bar{\mathbf{x}}(t) + P(t)\mathbf{b} \tag{2.21}$$

where t is a continuous variable. For a smooth parametric surface (2D manifold), $\mathbf{x}(u, v)$ like the bladder and prostate surfaces in this work, we choose the shape model given by

$$\hat{\mathbf{x}}(u, v) = \bar{\mathbf{x}}(u, v) + \sum_{i=1}^{M} b_i \mathbf{p}_i(u, v) = \bar{\mathbf{x}}(u, v) + P(u, v)\mathbf{b} \tag{2.22}$$

An illustration of how the shape changes over the first shape mode can be seen in Fig. 2.4.

2.4.2 Evaluation measures

Evaluating the quality of the statistical shape model is an important task for the whole dissertation. Under the assumption that shape sample is governed by a Gaussian distribution, the quality of model depends on that of correspondence estimated by the group-wise registration methods surveyed in Sec. 2.3. If the manual landmarks are available, the first choice is to use the distance to these landmarks as the gold standard to evaluate the correspondence. Unfortunately, we do not have

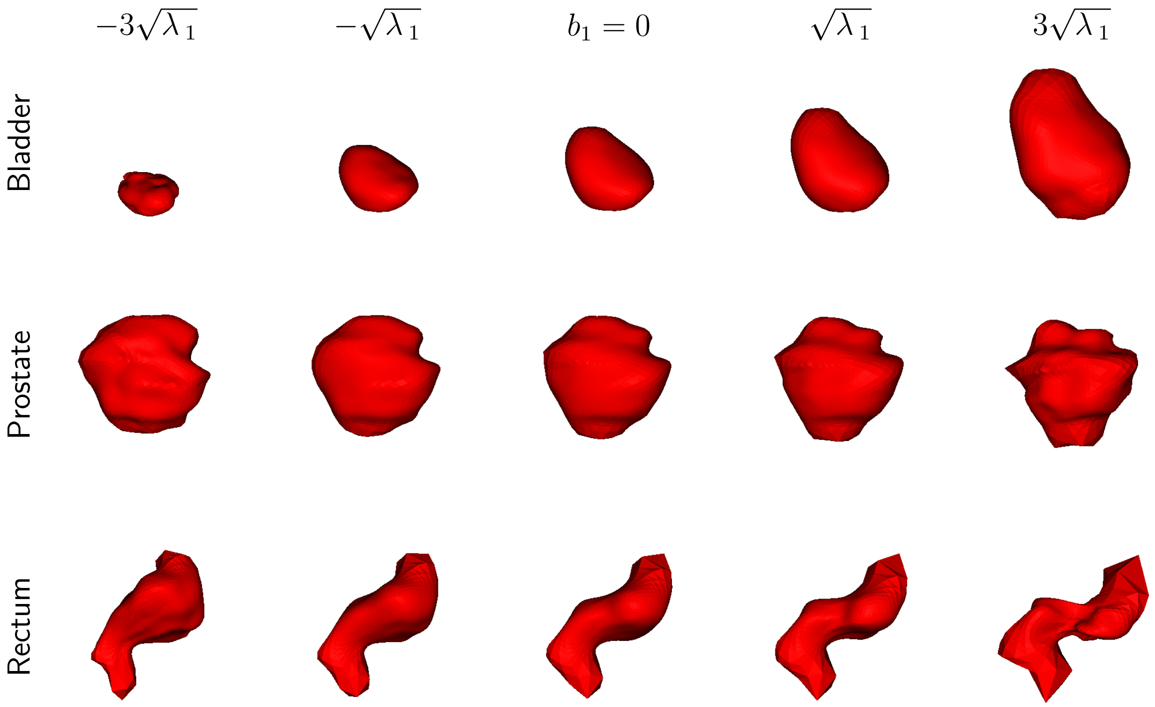


Figure 2.4: Illustration of how the shapes of the bladder, prostate and rectum of Patient A from NKI dataset changes. Only the first shape mode b_1 is used. b_1 varies from $-3\sqrt{\lambda_1}$ to $3\sqrt{\lambda_1}$. The mean shapes of each organs are shown in the middle column with $b_1 = 0$. The groupwise surface correspondence was estimated by MDL method.

them in our dataset. To evaluate in absence of ground truth, we use two published and widely accepted measures [36, 117]: *specificity* and *generalization ability*, covering different properties that a good model is expected to have.

Model generalization: a overfitted model is unable to generalize to unseen shapes. Given limited training samples, a model with strong generalization ability can perform well the *unseen* shape instances. The generalization ability is quantified using leave-one-out trials. Suppose the training set is given by $A = \{\mathbf{x}^i : i = 1, \ldots, M\}$ and the number of variation modes is n_m. We build a model from all training samples except the j-th. We can fit this model to the j-th example $\mathbf{x}^j$ to obtain an approximation $\tilde{\mathbf{x}}^j$. The average distance between the fitted model $\tilde{\mathbf{x}}^j$ and the excluded $\mathbf{x}^j$ over the whole trials [36] is a function $G(t)$.

$$G(n_m) = \frac{1}{M} \sum_{j=1}^{M} \|\mathbf{x}^j - \tilde{\mathbf{x}}^j\| \tag{2.23}$$

Model specificity: a model is specific if it only generates valid shape instances that are similar to those in the training set [36]. One way of measuring the specificity is to compare the probability density function (p.d.f) of a model and that of the generated instances. Suppose we have a set of randomly generated examples $B = \{\mathbf{y}^j : j = 1, \ldots, n_s\}$ from a model with n_m shape parameters, the average distance between each of them to training set suggests how specific these generated sample to the model. Therefore the specificity $S(k)$ is defined as:

$$S(n_m) = \frac{1}{n_s} \sum_{j=1}^{n_s} \min_i \|\mathbf{y}^j - \mathbf{x}^i\| \tag{2.24}$$

where n_s is the number of samples drawn from the shape model. In the experiment n_s was chosen 10,000 times.

2.4.3 Experiment

Our NKI database contains the FBCT pelvic images from 19 patients acquired daily through the treatment session (see [41] for detailed description). Table 2.1 provides a simple overview of the availability of the patient data. The bladder, prostate and rectum on each image were contoured slice-by-slice by a single expert. These organ segmentations comprise the training population of our experiment.

Two types of shape model can be constructed: one is for the whole patient population (population level), the other is for the specific patient (individual level). Due to the limited size of the shape instances, it is unrealistic to build an accurate model for the entire population (including the unseen shapes). Besides, our immediate use shape model is to help the image registration for a particular patient. Therefore, our experiment only evaluates the shape model at the individual level.

Based on our requirement on algorithms and the availability of the implementations, we choose five algorithms as the candidates. 1) SPHARM First order ellipsoid matching [116] (SPHARM); 2) Entropy-based particle system with SPHARM initialization [17] (Particle); 3) SPHARM registration followed by Particle registration [99]

Table 2.1: Numbers of CT scans (days) for each patient in NKI dataset

Patient	Total #	Scan													
	of scans	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	13	X	X	X	X	X	X	X	X	X	X	X	X	X	
B	10	X	X	X	X	X	X	X	X	X	X				
C	12	X	X	X	X	X	X	X	X	X	X	X	X		
D	11	X	X	X	X	X	X	X	X	X	X	X			
E	13	X	X	X	X	X	X	X	X	X	X	X	X	X	
F	9	X	X	X	X	X	X	X	X	X					
G	12	X	X	X	X	X	X	X	X	X	X	X	X		
H	13	X	X	X	X	X	X	X	X	X	X	X	X	X	
I	12	X	X	X	X	X	X	X	X	X	X	X	X		
J	11	X	X	X	X	X	X	X	X	X	X	X			
K	12	X	X	X	X	X	X	X	X	X	X	X	X		
L	14	X	X	X	X	X	X	X	X	X	X	X	X	X	X
M	11	X	X	X	X	X	X	X	X	X	X	X			
N	11	X	X	X	X	X	X	X	X	X	X	X			
O	6	X	X	X	X	X	X								
P	13	X	X	X	X	X	X	X	X	X	X	X	X	X	
Q	13	X	X	X	X	X	X	X	X	X	X	X	X	X	
R	12	X	X	X	X	X	X	X	X	X	X	X	X		
S	11	X	X	X	X	X	X	X	X	X	X	X			

(SPHARM+Particle); 4) Minimum Description Length [72] (MDL); 5) SPHARM registration followed by MDL (SPHARM+MDL). The building steps are illustrated in Fig. 2.5.

Parameterizing organ surfaces

We choose the parametric surface as the shape representation. With the shape decomposition ability of SPHARM (Sec. 2.3.1), we converted each binary organ segmentation to the block representation and then the icosahedron or the (u, v) parametric representations, respectively. This process is illustrated in Fig. 2.6. Icosahedron

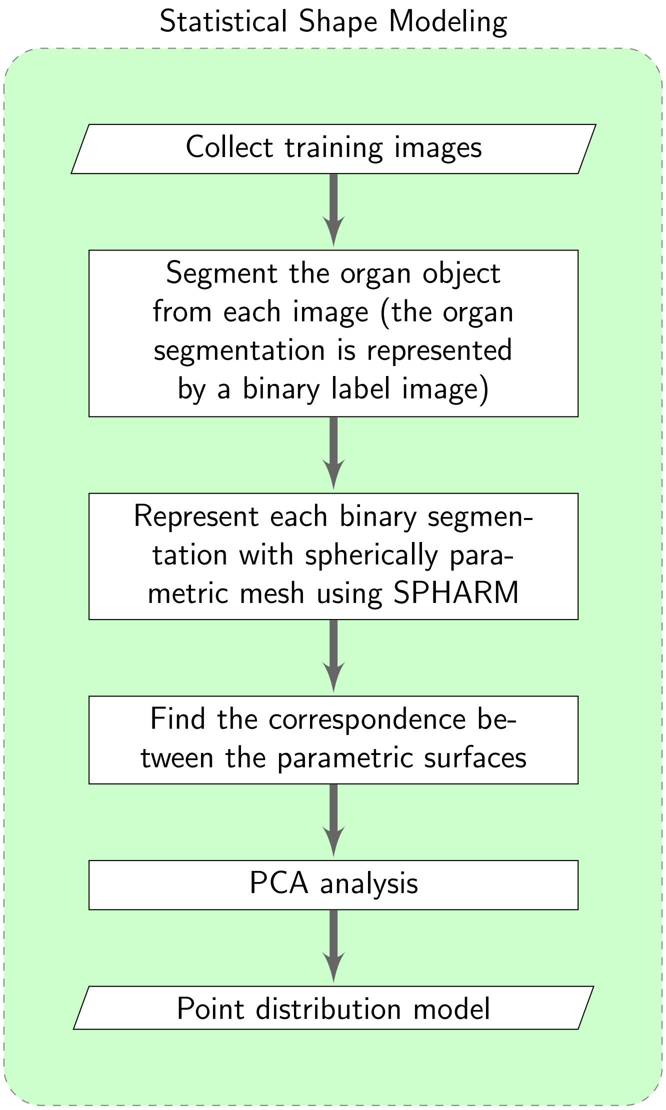


Figure 2.5: Steps for building the organ shape model.

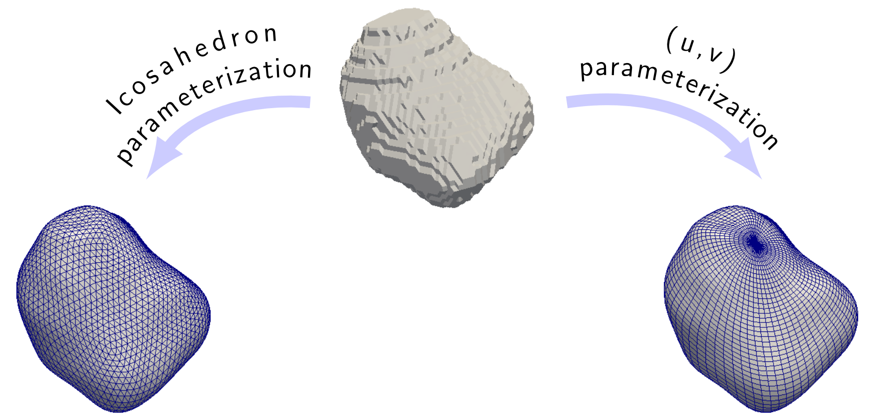
representation is preferred over the (u, v) one for shape model building since the uniformly distributed point coverage reduces the bias. However, for the computation of surface assimilation in the next chapters, a (u, v) representation is needed since it provides a natural coordinate system on the sphere domain.

2.4.4 Results

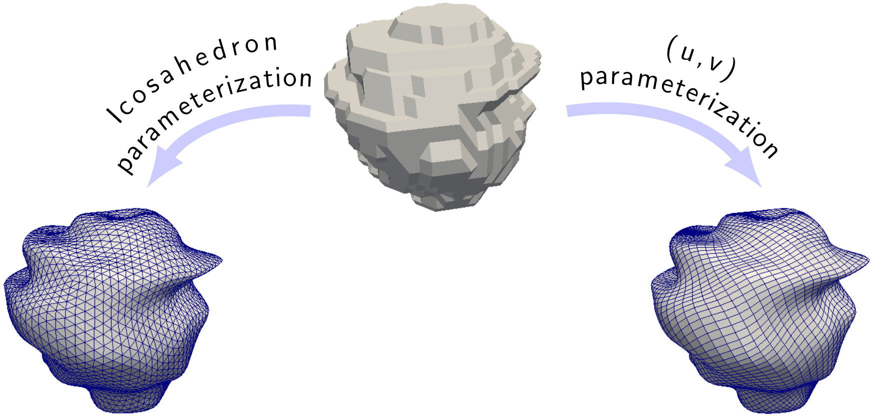
The evaluation results of the five surface registration algorithms on the bladder, prostate and rectum dataset are reported in Fig. 2.7, Fig. 2.8 and Fig. 2.9, respectively. In each figures, the x-axis is the number of shape variation modes n_m and the y-axes associated with $G(n_m)$ and $S(n_m)$ are the errors of the generality ability and specificity, respectively. For a specific algorithm and a given n_m, the errors are averaged cross the patient population. An algorithm with smaller errors is better.

The evaluation results show the pattern that all algorithms' generality errors go down and specificity errors slightly go up as more shape modes are involved. The error ranges on the bladder and prostate are similar and both are much smaller than that on the rectum. We attribute this pattern to the topological difference between the spherical-like bladder, prostate and the tubular-like rectum. This may suggest: 1) the candidate algorithms are not good at tubular structure modeling 2) the shape consistency among rectum surfaces is bad. Due to these reasons, we did not use the rectum data in our experiments in Chapter 3 and Chapter 4.

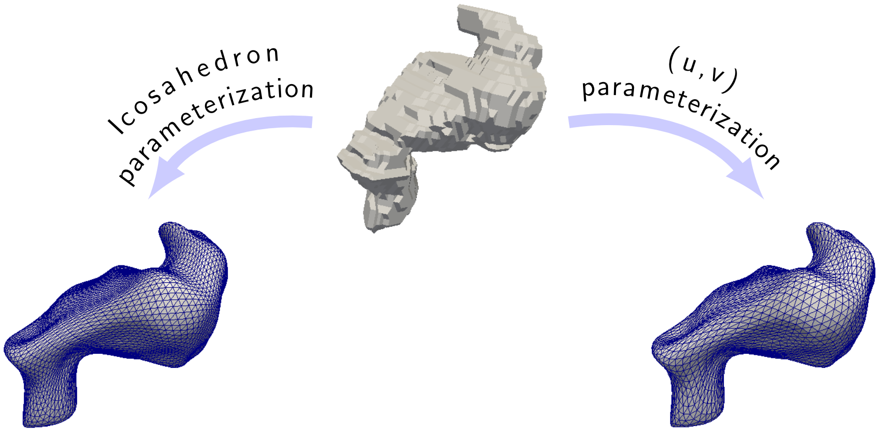
Some patients have more than 12 organ segmentations. Thus n_m along the x-axis could be larger than 10 (we need 11 segmentations to complete the computation



(a) Bladder parameterization

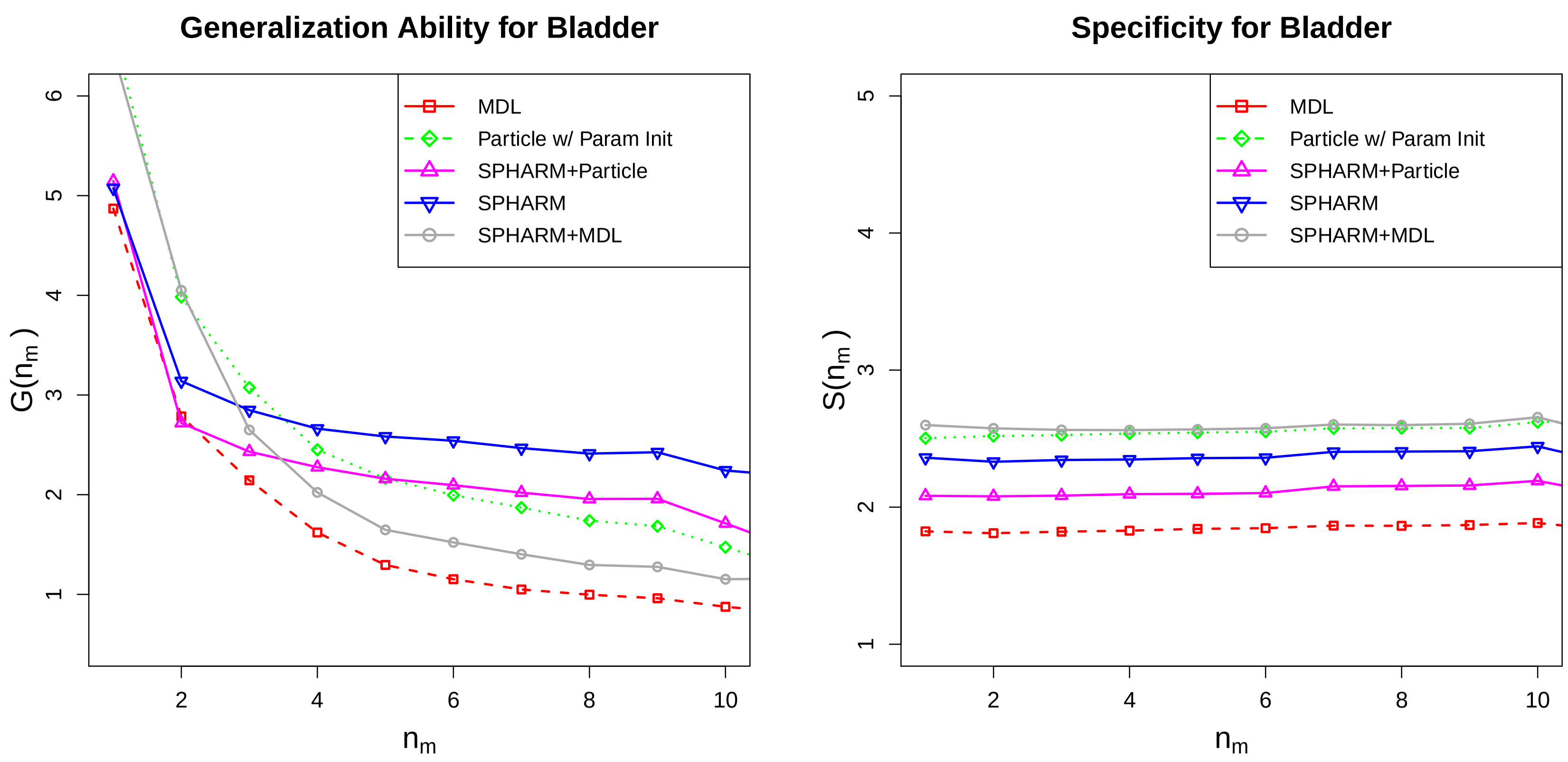


(b) Prostate parameterization



(c) Rectum parameterization

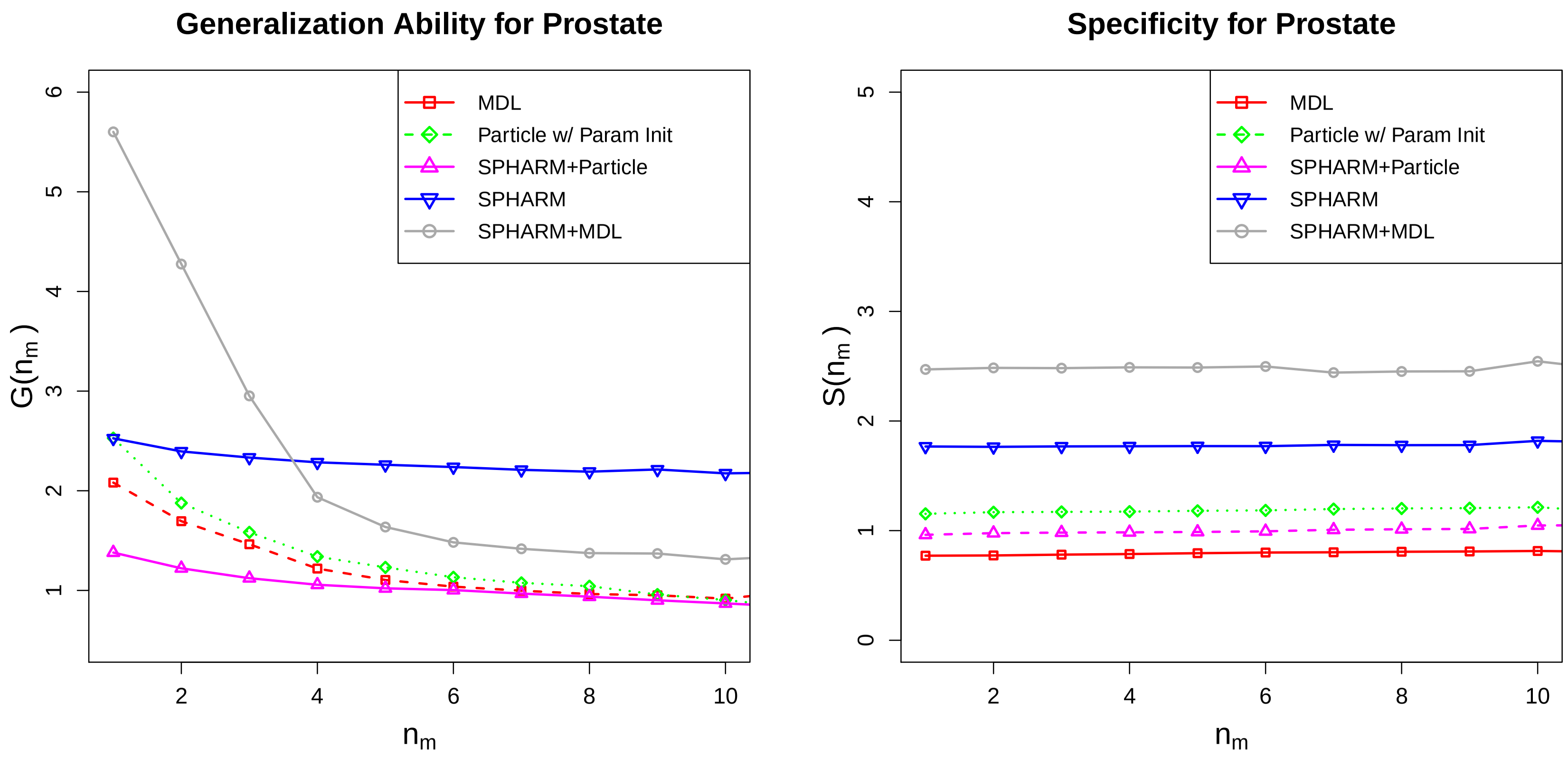
Figure 2.6: Illustration of surface parameterization using SPHARM.



(a) Generalization ability error

(b) Specificity error

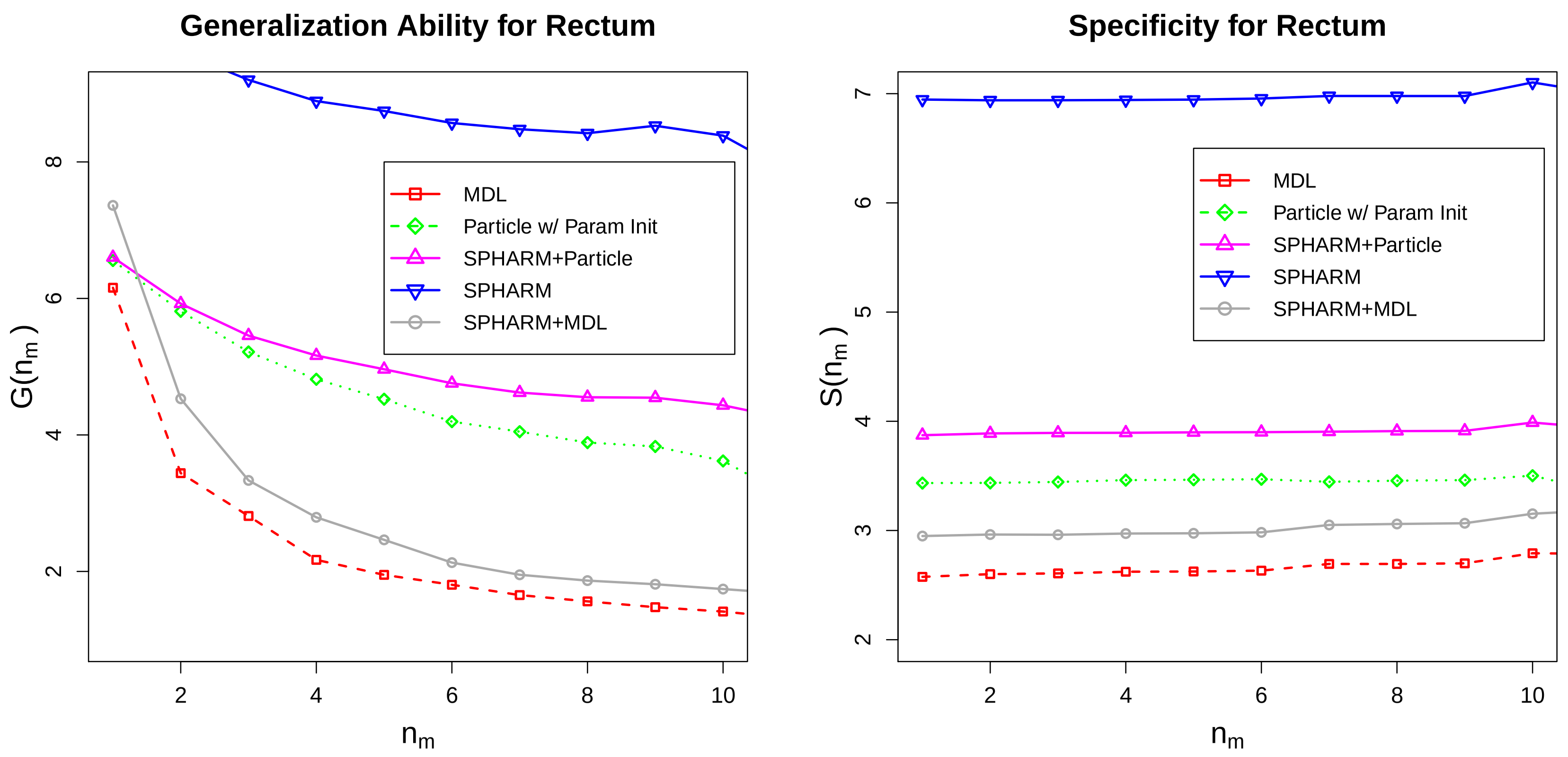
Figure 2.7: Comparison of five surface registration algorithms on the bladder data. n_m is the number of shape modes. The value of each point is averaged over the population and lower is better.



(a) Generalization ability error

(b) Specificity error

Figure 2.8: Comparison of five surface registration algorithms on the prostate data. n_m is the number of shape modes. The value of each point is averaged over the population and lower is better.



(a) Generalization ability error

(b) Specificity error

Figure 2.9: Comparison of five surface registration algorithms on the rectum data. n_m is the number of shape modes. The value of each point is averaged over the population and lower is better.

if $n_m = 10$). However, the average over the population is unreliable when n_m is beyond 11 (we do not have enough samples for the average).

In the case of shape analysis, the generalization ability is more important than the model specificity because we do not generate new surfaces. In the case of shape simulation and prediction, model specificity weighs over the generalization ability.

When the number of shape variation modes is larger than 6, our experiment shows MDL is consistently better than other algorithms. The results agree with [117, 34] where MDL is the winner over SPHARM and DevCov on the site of lateral ventricles and femoral head. We also see that the concatenation of algorithms (one algorithm followed by another), like SPHARM+MDL and SPHARM+Particle, helps little.

We conclude by choosing MDL as the shape modeling method for the remaining chapters. However the implementation of MDL [72] we have is much slower (at least 10 times) than other non-MDL methods. The Particle method offers a better trade-off between the accuracy and running time. In the case of prostate, the performance of Particle is closer to MDL.

To leave us enough data for validation experiments done in Chapter 3 and Chapter 4, we choose the first 8 organ shapes for every patient to build the shape model.

2.5 Inter-observer Segmentation Error

Given multiple segmentations on the same object, we propose to quantify the organ boundary segmentation error denoted by $\omega(u, v)$ using the Principal Component Analysis (PCA) with the assumption that the inter-observer segmentations follow Gaussian model. The correspondence among these manual surfaces can be estimated using the group-wise surface registration methods in the previous section.

The quantification of the inter-observer segmentation error (ISE) serves two purposes: 1) it helps to generate the simulated manual segmentations used in the experiments of Chapter 3 and Chapter 4, 2) it predicts the segmentation errors on a new manual segmentation, playing an important role in the algorithms in Chapter 3 and Chapter 4. We demonstrate the idea in the following study which quantifies the ISE for FBCT/CBCT.

2.5.1 Description of Patient Data

A multiple observer-contouring study was performed on a set of FBCT and CBCT that were obtained as part of an IRB-approved protocol [141]. All patients underwent primary external beam radiotherapy and had tumors that were confined to the prostate (2 x T1cN0, 1 x T2aN0, 1 x T3aN0). Three patients had markers or dosimeters implanted prior to treatment: Patient 1 had 2 DVS devices (Sicel Technologies, Morrisville, NC) implanted in the prostate, patient 3 3 gold markers and patient 4 3 Calypso markers (Calypso Medical Technologies, Seattle, WA). Fan beam CTs of the pelvis were acquired with continuous 1.5 mm slices on a 16 slice scanner with a 60 cm field-of-view (140kV, 350 mAs, Brilliance Big Bore, Philips Medical Systems, The Netherlands). Kilovoltage CBCTs were acquired in half-fan mode with bowtie filter and anti scatter grid (125kVp, 80mA, 25 ms, Varian Medical Systems, Palo Alto, CA). 630 projections were acquired over a 360 degree rotation in approximately one minute. The field-of-view was 48x48 cm^2 in axial plane and 14.25 cm longitudinal length. A 512 matrix was used for both imaging modalities. Pixel size for CBCTs was either 0.059 x 0.059 x 0.15/0.1 cm^3 or 0.49 x 0.49 cm x 0.15 cm^3. For FBCTs pixel size was either 0.097 x 0.097 x 0.15 cm^3 or 0.117 x 0.117 x 0.15 cm^3. All imaging was performed without intravenous contrast. Structure delineation On both imaging modalities prostate, bladder and rectum were delineated independently by 5 medically experienced personnel using commercially available treatment planning software (Pinnacle version 8.1, Philips Medical Systems, Milpitas, CA). Contouring of FBCT and CBCT images was performed independently on each modality according to a detailed contouring protocol that among others provided instructions on the choice of window level and a description of anatomical details to be considered during the contouring process.

2.5.2 Method

Let $X_i(u, v)$ and $Y_i(u, v)$, $i = 1 \dots n$, denote the manually segmented prostate surfaces traced on FBCT and CBCT images, respectively, by the i-th expert. Each randomly sampled FBCT prostate surface, $X(u,v)$, can be represented as

$$X(u,v) = \bar{X}(u,v) + \sum_{i=1}^{n} b_i \beta_i(u,v) \tag{2.25}$$

where $\bar{X}(u,v)$ is the mean prostate surface, b_i is a random variable with normal distribution $b_i \sim N(0,\sigma_i^2)$, σ_i^2 is the i-th eigenvalue of the covariance matrix, and $\beta_i(u,v)$ is the corresponding vector-valued eigenfunction, representing the i-th variation mode. Similarly, $Y(u,v)$ can be represented as

$$Y(u,v) = \bar{Y}(u,v) + \sum_{i=1}^{n} c_i \psi_i(u,v) \tag{2.26}$$

where $\bar{Y}(u,v)$ is the mean shape of CBCT prostate, $c_i \sim \mathcal{N}(0,\epsilon_i^2)$, and $\psi_i(u,v)$ is the corresponding variation mode.

The segmentation uncertainty $\omega(u,v)$ is designed to have the following property. $\omega(u,v)$ is low at the regions where the experts' manual contours closely agree with each other. Experts are confident about the boundary location either because of the boundary is easy to distinguish or is supported by experts' knowledge. When we register these regions, more weight should be given to these regions. The segmentation uncertainty is high at regions of large inter-observer discrepancies, indicating the lack of expert consensus on boundary location. We thus should give less weight when matching these regions. To simplify the computation we define

$$\omega(u,v) = \mathrm{tr}(\mathrm{Cov}\,(X(u,v))) + \mathrm{tr}(\mathrm{Cov}\,(Y(u,v))) \tag{2.27}$$

At each point (u_0, v_0), the trace of the covariance $\mathrm{Cov}\left(X(u_0, v_0)\right)$ is derived as

$$\begin{aligned}
\mathrm{tr}(\mathrm{Cov}\left(X(u_0, v_0)\right)) &= \mathrm{tr}(\mathrm{E}\left\{[X(u_0, v_0) - \bar{X}(u_0, v_0)][X(u_0, v_0) - \bar{X}(u_0, v_0)]^{\mathsf{T}}\right\}) \\
&= \mathrm{tr}(\mathrm{E}\left\{[\sum_{i=1}^{n} b_i \beta_i(u_0, v_0)][\sum_{i=1}^{n} b_i \beta_i(u_0, v_0)]^{\mathsf{T}}\right\}) \\
&= \sum_{k=1}^{3} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{E}\left\{b_i b_j\right\} \beta_{ik}(u_0, v_0) \beta_{jk}(u_0, v_0) \\
&= \sum_{k=1}^{3} \sum_{i=1}^{n} \sigma_i^2 \beta_{ik}^2(u_0, v_0)
\end{aligned}$$

where i has the same meaning in Eq. (2.25) and k represents the k-th direction among x, y and z directions. Similarly, $\mathrm{tr}(\mathrm{Cov}\left(Y(u_0, v_0)\right)) = \sum_{k=1}^{3} \sum_{i=1}^{n} \epsilon_i^2 \psi_{ik}^2(u_0, v_0)$.

2.5.3 Preliminary Results for Individual Patient

A multiple observer-contouring study was performed on a set of FBCT and CBCT from 4 patients. On both imaging modalities prostate, bladder and rectum were delineated independently by 5 medical experts using commercially available treatment planning software (see **??** for description). This is a different dataset with that used in Sec. 2.4.

We computed the FBCT/CBCT segmentation errors for each of four patients following the steps in Fig. 2.10. The correspondence among the training surfaces were computed using MDL (see Sec. 2.3.3). The boundary variance among the manual surfaces was computed using the method in Sec. 2.5.2 and are illustrated in Fig. 2.13.

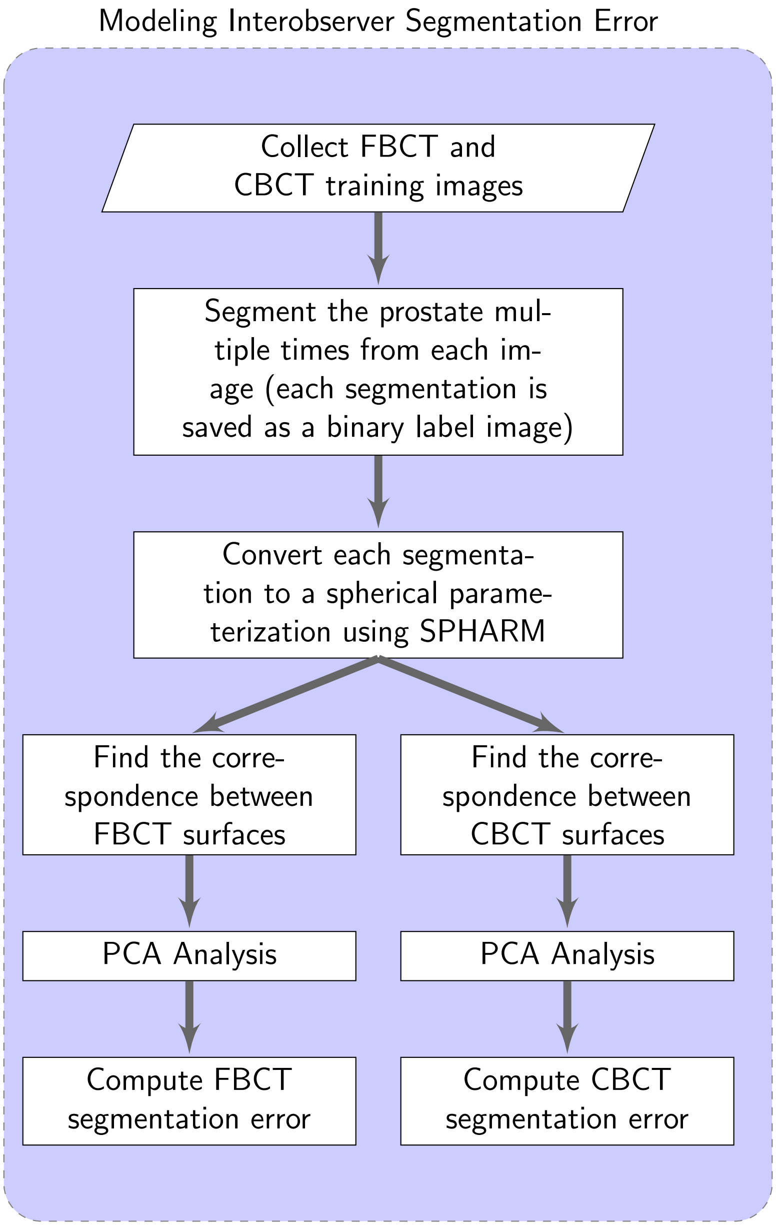


Figure 2.10: Steps for modeling the inter-observer segmentation error.

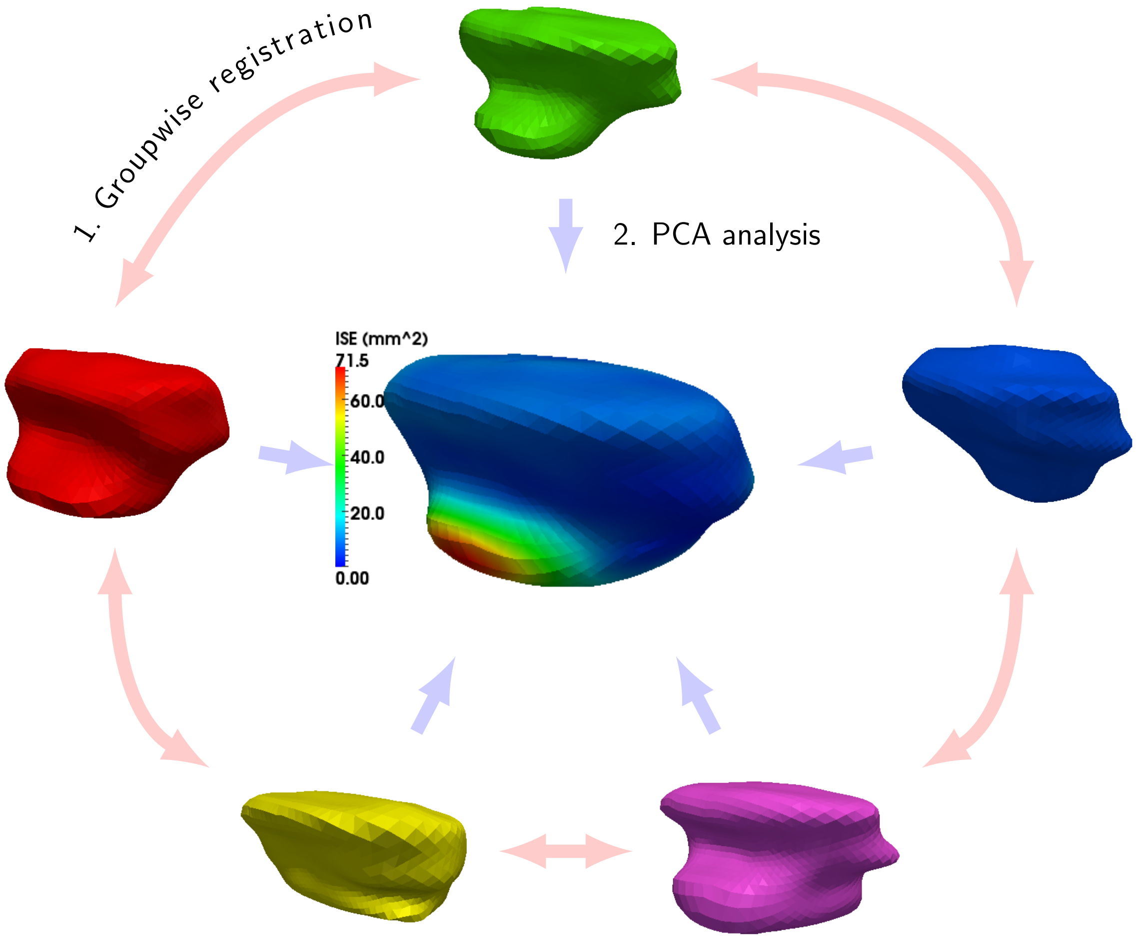


Figure 2.11: Procedures for quantifying the inter-observer segmentation error (ISE). Given five instances (in solid colors) of segmentations of the same prostate, we first run group-wise surface registration to estimate the correspondence (red arrows) and then use PCA to compute the distribution of segmentation errors on the average surface (in rainbow color).

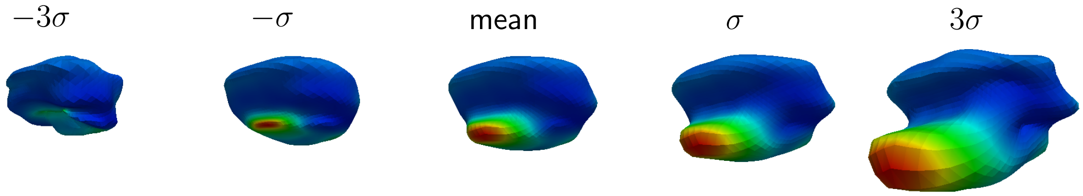
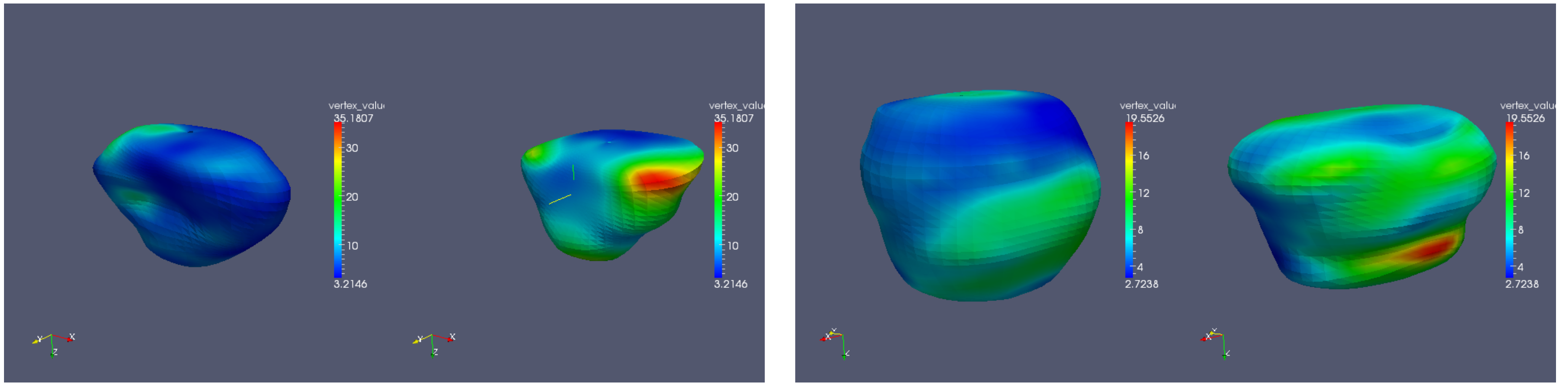
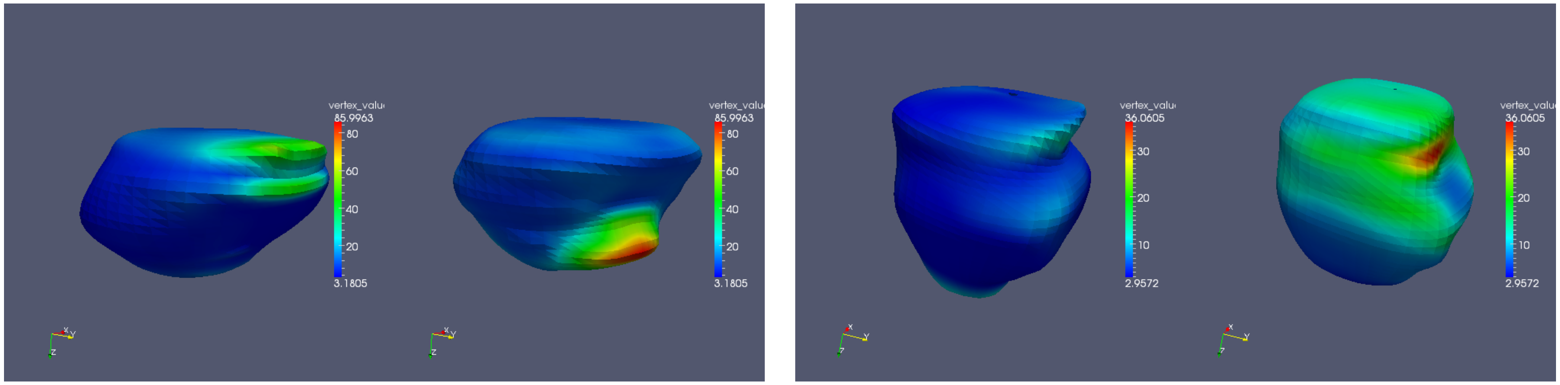


Figure 2.12: Illustration of the first variation mode of the ISE model. The mean shape ($\sigma = 0$) is shown at the center. This example shows that a large segmentation error happened at the base of the prostate.



(a) Patient 1. FBCT v.s. CBCT

(b) Patient 2. FBCT v.s. CBCT



(c) Patient 3. FBCT v.s. CBCT

(d) Patient 4. FBCT v.s. CBCT

Figure 2.13: FBCT and CBCT prostate interobserver segmentation errors for four patients. For Patient 3, The range of FBCT surface location variance is [1.9, 74.5] mm^2. The range of CBCT surface location variance is [2.5, 73.5] mm^2. Where the expert surfaces are close to each other, the variance is small indicating the low uncertainty (blue) about the boundary position between the human observers. Likewise, the large variance (red) suggests high uncertainty between human observers.

2.6 Discussion

Different shape analysis algorithms have the different assumptions therefore the different methodologies towards the shape object. Each algorithm has its own strengths and limitations. Therefore, it is necessary to evaluate the candidate methods in the context of physiologically meaningful results. This chapter evaluated different 3D shape analysis methods to seek the appropriate surface alignment method that can capture the shape knowledge of the underlying anatomical changes of pelvic organs. With generalization and specificity as the benchmark measures, the groupwise surface registration method based on the minimum description length (MDL) is chosen over other methods.

A wider range evaluation experiment can be done in the future by including the methods in the metric shape space Sec. 2.3.4 and those based on grayscale image registration. Other evaluation methods, e.g., ground truth PDM (similar to know transformation) proposed by Munsell [96], can be included. We used a fixed surface resolution (the number of points covering the surface), i.e., 2562 points (Icosahedron, d4)) in the experiment. A sensitivity analysis on the resolution can also be an interesting study since it affects the speed of the algorithms, a major concern for clinical use.

In Sec. 2.4.3, we map the pelvic organ surfaces, i.e. the bladder, prostate and rectum surfaces, onto the unit sphere, from there we establish the parametric mesh using the SPHARM method. For bladder and prostate, since their shapes are similar to a sphere, this spherical mapping does not cause significant distortion. For a rectum surface, however, is topological equivalent to a tube. In the pre-processing step we capped the two ends to make the mapping working but large mapping distortion could occur. Huysmans [76] proposed a parameterization of tubular surfaces and a group-wise surface registration based on MDL method. Future work can be done to evaluate the severity of the distortion of spherical parameterization of rectum surface. If the distortion is large, the cylindrical mapping can be an alternative option.

Comparing with the implicit representations, such as level set, curvature flow, medial representation [100], the parameterized surface is more straightforward for PCA-based shape analysis. The advantages include 1) it is a natural representation of 3D objects and the geometric features (e.g. curvature) are preserved naturally 2) its parametric form makes it easy to switch among arbitrary resolutions. 3) it is less demanding on the disk storage and computational resources than the volume-based representation. Mapping the original shapes to the spherical domain provides a common coordinate system in which the one-by-one correspondence between surfaces can be estimated. These methods, like spherical-MDL and registration in shape space, can only model one shape object. The extension to multiple shape registration has not been reported. In such applications, a volume-based or particle-based method may be preferable.

One limitation of this study is that we explain the shape changes purely based on the observed shape samples without considering the domain-specific knowledge. For instance, the shape changes of a prostate are also subject to the interaction with the seminal vesicle. As such the large deformation on a pair of prostate surfaces suggested by the pure shape analysis is implausible in clinical sense. The fullness and voidness of the bladder and the rectum also make it hard for shape analysis and inference. Therefore the clinical inputs are necessary for the accurate modeling of shape objects and the interaction among them.

Although we only built statistical shape models for individual patients, the shape model for the population, if accurate, can be used in a wider range of applications [93, 95]. Modeling and quantifying the segmentation error is a non-trivial and interesting question in radiotherapy. And it is receiving increasing attention [89].

We analyzed the shape changes and the segmentation error using Principle Component Analysis (PCA) with the assumption that they can be explained by the Gaussian model. From that we developed a method for quantification. However, due to the limited resources, this assumption was not verified. With more data available, kernel-based method [26], non-linear PCA [11], or Principal Geodesic Analysis (PGA) [52], can be applied.

The modeling of segmentation error cross the population is even harder since it requires large scale training inputs and sophisticated statistical methods. Likely, a consistent segmentation error model independent of patient-to-patient does not exist. The distribution of interobserver segmentation error (ISE) varies case-by-case. As the study of vessel contouring [19] suggests the segmentation error could be related with the surrounding intensity.

2.7 Summary

The surface assimilation method introduced in the next chapter requires a good understanding of the organ shapes and the interobserver segmentation uncertainty. We propose that both knowledge can be obtained though the statistical shape analysis (SSA). The quality of SSA, in turn, depends on that of group-wise correspondence. Therefore, we surveyed the common surface registration algorithms, selected the candidate group-wise registration methods, evaluated their performance with our organ shape database. Following the Principal Component Analysis method, we built two types of statistical models for organ segmentations using the suitable algorithms. One is the statistical shape model (SSM) summarizing the average shape and the major variation modes. SSM helps to predict where the plausible boundary is in a noisy segmentation. The other model is the one for interobserver segmentation error (ISE) characterizing the distribution of manual segmentation error along a organ surface. Both models carry important knowledge that are essential to the surface assimilation (Chapter 3) and the shape model-constrained image registration (Chapter 4).

CHAPTER 3
SURFACE ASSIMILATION

3.1 Introduction

Suppose the signal $f(t)$ is corrupted by the additive noise $\epsilon(t)$ and the observed measurement y_i of $f(t)$ at i-th location of total N observation spots can be represented as

$$y_i = f(t_i) + \epsilon_i, \quad i = 1, 2, \ldots, N \tag{3.1}$$

Assume f is smooth and has continuous derivative at least m order, and ϵ_i is a i.i.d Gaussian random variable with zero mean and variance σ^2. An estimate of f can be obtained by minimizing

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - f(t_i))^2 + \lambda \int_0^1 (f^{(m)}(u))^2 du \tag{3.2}$$

The first part, represented by the residual sum of squares (RSS), measures the fidelity to the observed signal. The second part measures the smoothness of the estimated signal. The smoothing parameter λ balances these two penalties, which is also known as the bandwidth in some statistics literature. When $\lambda = 0$, Eq. (3.2) becomes an interpolation problem. When $\lambda \to \infty$, it is close to a linear regression model. As a classic non-parametric regression problem, it has received considerable amount of attention [28]. Among the methods solving this regression model, the theory of smoothing splines proposed by Wahba *et al.* [132] is widely accepted. It shows that the minimizer of Eq. (3.2) is a natural polynomial spline and provides the automatic algorithm (Generalized Cross Validation) to estimate the optimal smoothing parameter λ.

The assumption that ϵ_i is both homoscedastic and uncorrelated is too restrictive for many practical applications. By weakening the assumption, Eq. (3.1) can be generalized in several ways. Supposing the variances of ϵ_i are unequal (heteroscedasticity), Andrews [4] showed GCV was still a feasible solution to achieve asymptotically the optimality. Altman [3], Opsomer *et al.* [98] and Brabanter *et al.* [40] considered the case that ϵ_i is correlated. Other extensions include the multi-variates regression, such as spline model on plain [55] and sphere [132], the error-in-variable, i.e. the measurements of t contain errors [7]. The limitations of the smoothing spline regression may include: 1) It requires large observational data to reveal its internal structure and properties. 2) When the errors are highly correlated, the estimation of f is likely not good enough.

Consider our original problem: given a stack of 2D manual contours of an 3D object with these contours having the segmentation (observational) errors, our goal is to estimate a new surface closer to the true one. It can be formulated as a nonparametric regression problem. The challenge is that the segmentation errors are correlated and unequal instead of i.i.d homoscedastic Gaussian random noise along the boundary. Without the prior knowledge about the signal, the nonparametric regression in this case is hard. To simplify our problem we assume the errors are uncorrelated based on the following reasons. The object boundary is contoured slice by slice and a human observer segmenting a slice generally does not refer to the previous ones. Thus the error between the slices can be treated as uncorrelated. The in-slice error is correlated, which makes it hard for the nonparametric regression to distinguish the signal from the noise. To deal with it we will use the history knowledge of the signal, i.e., the statistical shape model (SSM) from the previous chapter. As for the heteroscedasticity of the error variances, estimated by the analysis of the interobserver segmentation error (ISE), we use them as a clue for fusing the observed boundary and the predicted one from the SSM.

3.2 Data Assimilation with Shape Prior

Interestingly, combining the observed surface with the prior knowledge is similar to a Data Assimilation (DA) [79] problem. Therefore we call our surface fusion method *Surface Assimilation*. As a typical DA problem, Numerical Weather Prediction (NWP) tries to give a better estimate of current weather condition by considering the current inaccurate observation and the prediction from the historical weather data.

The data assimilation can be formulated as an optimization problem. In a simple form, given N pointwise measurements, the cost function (Eq. (3.3)) consists of the sum of the squared difference between the observed measurements y_i and the estimate $f(t_i)$ weighted by the accuracy of observation σ_i^2, plus the sum of the squared difference between the model prediction $\hat{y}_i$ and the estimate $f(t_i)$ weighted by the accuracy of prediction δ_i^2.

$$C(f) = \sum_{i=1}^{N} \frac{(y_i - f(t_i))^2}{\sigma_i^2} + \sum_{i=1}^{N} \frac{(\hat{y}_i - f(t_i))^2}{\delta_i^2} \tag{3.3}$$

The pointwise correspondence between y_i, $f(t_i)$ and $\hat{y}_i$ is established by the curve or surface registration (see Chapter 2). Under the spline-based function approximation framework, the smoothness of f in Eq. (3.3) is regularized by the constraint:

$$C(f) = \sum_{i=1}^{N} \frac{(y_i - f(t_i))^2}{\sigma_i^2} + \sum_{i=1}^{N} \frac{(\hat{y}_i - f(t_i))^2}{\delta_i^2} + \lambda \int_0^1 (f^{(m)}(u))^2 \mathrm{d}u \tag{3.4}$$

In our case, the statistical shape model (SSM) built from a set of high quality object segmentations provides the model prediction $\hat{y}$. A set of multi-observer segmented objects are collected to determine the segmentation error σ_i. As in Chapter 2, the shape model and segmentation error are both assumed as Gaussian random distribution summarized by Point Distribution Model (PDM). Suppose the shape is represented by N points and the SSM contains M variation modes described as

$$x = \bar{x} + Pb \tag{3.5}$$

where $\bar{x}$ is the mean shape, $b = (b^1, b^2, \ldots, b^M)^T$ is a $M \times 1$ vector of Gaussian random variables, $P = (\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^M)$ is a $N \times M$ matrix defining the shape variation directions. Given an observed surface y, we can obtain a predicted one by minimizing

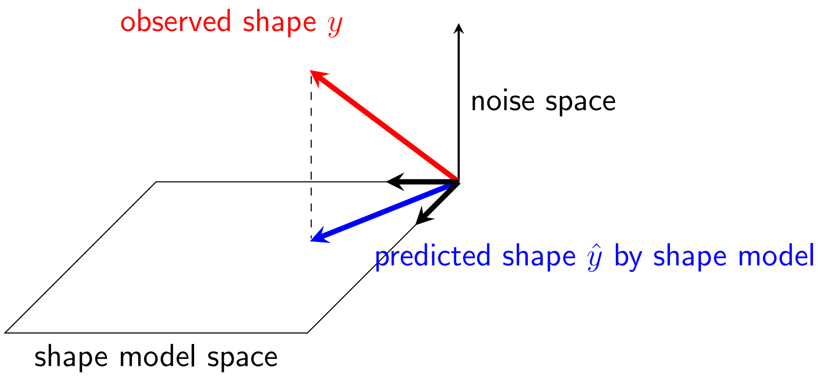


Figure 3.1: Observed shape (red) is projected onto the shape model space, forming a predicted surface (blue).

the cost function:

$$\hat{b} = \arg\min_{b}[\bar{x} + Pb - R(y)]^2 \tag{3.6}$$

where $R(\bullet)$ is the Procrustes transformation. This linear regression gives a set of shape coefficients $\hat{b}$. From it a new shape $\hat{y} = \bar{x} + P\hat{b}$, which is the predicted surface based on our prior shape knowledge Fig. 3.1. We will combine it with the observed surface to estimate the truth.

Eq. (3.6) can be viewed as a signal-noise filtering process. Suppose the statistical shape model (SSM) can capture all instances of a signal (true shape). Given a noisy signal (the true surface corrupted by the inter-observer segmentation error, ISE), the recoverability of SSM against ISE depends on the orthogonality of the space of SSM and ISE. Ideally, if the ISE noise is in the null space of SSM, by solving Eq. (3.6), ISE will be filtered out and $\hat{y}$ will be close to the true one. The worse case happens when the ISE falls into the same space of the SSM where the noise and signal are undistinguishable. Also, the recoverability is compromised by the Procrustes transform since the differences in scale, rotation and translation are removed. A signal corrupted by these transformations will be hard to identify.

In many real applications, the shape model is not perfect, meaning that the signal often contains the components in the higher dimensions. In such case, even if the noise is perpendicular to the signal, the shape model prediction is not accurate enough. Therefore we use the current observed signal y to compensate. In this study we choose the following equation to assimilate these two kinds of surfaces:

$$C(f) = \sum_{i=1}^{N}(1-\alpha_i)(y_i - f(t_i))^2 + \sum_{i=1}^{N}\alpha_i(\hat{y}_i - f(t_i))^2 + \lambda \int_0^1 (f^{(m)}(u))^2 \mathrm{d}u \tag{3.7}$$

where $\alpha_i \in [0,1]$ is the pointwise normalized of Eq. (2.27) balancing the contribution from the model prediction and the current observation at the i-th point.

As we see in the later discussion, under the smoothing spline regression (SSR) framework, the first two terms are in a Hilbert space, orthogonal to the one that the third term is in. Therefore the minimizer of $C(f)$ in Eq. (3.7) can be obtained by minimizing the first two terms (Data Assimilation) and substituting y_i in Eq. (3.2) with this minimizer, and then solving Eq. (3.2) by SSR.

The first terms of Eq. (3.7) is a quadratic function. It is easy to verify that its minimizer is given by

$$\hat{f}_i = (1-\alpha_i)y_i + \alpha_i\hat{y}_i \tag{3.8}$$

Therefore the minimizer of Eq. (3.7) is obtained by minimizing the following equation

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{f}_i - f(t_i))^2 + \lambda \int_0^1 (f^{(m)}(u))^2 du \tag{3.9}$$

In the upcoming sections we introduce the theory of smoothing spline regression.

3.3 Smoothing Spline Regression

Suppose the observation of f is indirect by satisfying the relationship $y(i) = L_i f(x_i) + \epsilon_i$ and the penalty on f is some more general quadratic functionals, we can extend Eq. (3.2)

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - f(t_i))^2 + \lambda \int_0^1 (f^{(m)}(u))^2 du \tag{3.10}$$

to a more general form

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - L_i f)^2 + \lambda \int_0^1 (L_m f(u))^2 du \tag{3.11}$$

where L_i is a linear evaluation functional and L_m is an m-th order differential operator. Different applications can be formulated under Eq. (3.11) with different domains and specific splines. For example, the growth curves are defined on the continuous interval and the weather data are defined on the sphere. By considering the model space of f in a reproducing kernel Hilbert space (RKHS), we can study this generalized Penalized Least Square (PLS) problem in a unified fashion [132].

3.3.1 Reproducing Kernel Hilbert Space

This section explains briefly the mathematical background of RKHS. The following definitions and theorems were adapted from [31] and [134].

Definition 3.1 (Vector space[31]). *A vector space* $\mathbb{V}$ *is a mathematical structure over a field* $\mathbb{F}$ *with a vector addition operation* $\oplus$*, a scalar multiplication operation* $\otimes$*, a unary additive inverse operation* $\ominus$ *satisfying the following axioms: For any* $u, v, w \in \mathbb{V}$ *and* $a, b \in \mathbb{F}$*,*

- *Associative Law:* $(u \oplus v) \oplus w = u \oplus (v \oplus w)$
- *Commutative Law:* $u \oplus v = v \oplus u$
- *Inverse Law:* $u \oplus (\ominus u) = 0_{\mathbb{V}}$
- *Identity Laws:* $0_{\mathbb{V}} \oplus u = u$ *and* $1 \otimes u = u$
- *Distributive Laws:* $a \otimes (b \otimes u) = (ab) \otimes u$ *and* $(a + b) \otimes u = a \otimes u \oplus b \otimes u$*.*

A function over a vector space $\mathbb{V}$ to R defined a norm, denoted $\|\cdot\|$, if it satisfies the following properties: For any $f, g \in \mathbb{V}$ and a scalar α, (1) $\|f\| \geq 0$, (2) $\|\alpha f\| = |\alpha|\|f\|$, (3) $\|f + g\| \leq \|f\| + \|g\|$, (4) $\|f\| = 0$ if and only if $f = 0$. A norm measures the size of a vector or the distance between two vectors. A vector space with a norm is call normed vector space.

Definition 3.2 (Banach space[31]). *A Banach space is a normed vector space* X *which is complete w.r.t the norm. That is, every Cauchy sequence* $\{f_n\}_{n=1}^{\infty}$ *converges to an element* f *in* X*:* $\lim_{n\to\infty}\|f_n - f\| = 0$

Definition 3.3 (Hilbert space[31]). *A Hilbert space* $\mathcal{H}$ *is a Banach space endowed with an inner product mapping which, for any* $f, g, h \in \mathcal{H}$ *and a scalar* α*, satisfies the following properties: (1)* $< \alpha f, g >= \alpha < f, g >$*, (2)* $< f, g >=< g, f >$*, (3)* $< f, g + w >=< f, g > + < f, w >$

Theorem 3.1 (Reize representation theorem[134]). *Let* L *be a bounded linear functional on a Hilbert space* $\mathcal{H}$*. There exists a unique* $u \in \mathcal{H}$ *such that* $Lf =< f, u >_{\mathcal{H}}$ *for all* $f \in \mathcal{H}$*. The element* u *is called the representer of* L*.*

Assuming $\mathcal{H}$ is a Hilbert space of real-valued functions from χ to $\mathbb{R}$ for some measurable domain χ. For a fixed element $x \in \chi$, a evaluational functional $L_x \in \mathcal{H}$ is defined as

$$L_x f = f(x)$$

Definition 3.4 (RKHS[134]). *A Reproducing Kernel Hilbert Space (RKHS) is a Hilbert space of real-valued functions* $\mathcal{H}$ *that every evaluational functional in* $\mathcal{H}$ *is bounded.*

Due to the Reize representation theorem, for every evaluation functional L, there is a unique vector (denoted by R_x) in $\mathcal{H}$ such that

$$L_x f = f(x) =< f, R_x >_{\mathcal{H}} \tag{3.12}$$

A function $R(x,z) = R_x(z)$ is called the Reproducing Kernel (RK) since $R(x,z) = L_z R_x =< R_x, R_z >$ (the second equality due to Reize representation theorem). It is easy to check $R(x,z)$ is symmetric and nonnegative definite.

Theorem 3.2 (MooreAronszajn theorem[134]). *For every nonnegative definite function* R *on* $\chi \times \chi$ *, there exists a unique RKHS on* χ *with* R *as its RK.*

See [132] for the proof. The following properties of RKHS will be used throughout this chapter. If $\mathcal{H}_0$ and $\mathcal{H}_1$ are two closed subspaces of an RKHS $\mathcal{H}$, then $\mathcal{H}_0$ and $\mathcal{H}_1$ are also RKHS. If $\mathcal{H}_0$ and $\mathcal{H}_1$ are orthogonal complement to each other, then we say $\mathcal{H}$ can be decomposed into $\mathcal{H}_0$ and $\mathcal{H}_1$, denoted by $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$, and their reproducing kernels, R, R_0 and R_1, satisfies $R = R_0 + R_1$, vice versa [134].

3.3.2 Penalized Least Square Estimation

This section summarizes materials from [134] and [132] describing the procedures to obtain the estimate of f, denoted as $\hat{f}$, by minimizing Eq. (3.11) with the assumption that the λ is fixed. The methods of selecting λ will be discussed in Sec. 3.5.

With the assumption that $f \in W_m$ and the decomposition $W_m = \mathcal{H}_0 \oplus \mathcal{H}_1$, the PLS in Eq. (3.11) can be expressed as

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - L_i f)^2 + \lambda \|P_1 f\|^2 \tag{3.13}$$

with the smoothness penalty functionals $J_m(f) = \|P_1 f\|^2$, where P_1 is the orthogonal projection of f onto $\mathcal{H}_1$ in W_m. When f is confined in RKHS, the minimizer of

Eq. (3.13) was proved as a spline function [81].

By Reize Representation Theorem, for every bounded linear functional L_i there exists a representer η_i such as

$$< \eta_i, f >= L_i f, \quad f \in \mathcal{H}_R$$

By the property of RKHS,

$$\eta_i(x) =< \eta_i, R_x >= L_i R_x = L_{i(\cdot)} R(x, \cdot) \tag{3.14}$$

where the first equality is derived from Eq. (3.12) and $L_{i(\cdot)}$ means L_i is applied to what follows as a function of $(\cdot)$ [132]. Eq. (3.14) shows that we can compute the the representer η_i by applying the operator L_i to the reproducing kernel R. Since $R(x, z) = R^0(x, z) + R^1(x, z)$ and $\mathcal{H}_R = \mathcal{H}_0 \oplus \mathcal{H}_1$, applying P_1 to η_i projects η_i onto $\mathcal{H}_1$.

Let $\xi_i = P_1\eta_i$, by the property of RKHS and the self-adjoint P_1 such that $< P_1 g, h >=< g, P_1 h >$ for any $g, h \in \mathcal{H}$, we have

$$\xi_i(x) =< \xi_i, R_x >=< P_1\eta_i, R_x >=< \eta_i, P_1 R_x >=< \eta_i, R_x^1 >= L_i R_x^1 \tag{3.15}$$

Furthermore, since $< \eta_i - \xi_i, \xi_j >= 0$

$$< \xi_i, \xi_j >=< \eta_i, \xi_j >= L_i \xi_j = L_{i(x)} L_{j(z)} R^1(x, z) \tag{3.16}$$

Denote

$$T = \{L_i \phi_\nu\}_{i=1}^{n}{}_{\nu=1}^{p} \tag{3.17}$$

$$\Sigma = \{L_{i(x)} L_{j(z)} R^1(x,z)\}_{i,j=1}^{n} \tag{3.18}$$

where T is an $n \times p$ matrix, and Σ is an $n \times n$ matrix. In our basic PLS problem (Eq. (3.2)), L_i is the linear evaluation functional, i.e., $L_i f = f(x_i)$, thus $\xi_i(x) = R^1(x, x_i)$, $T = \{\phi_\nu(x_i)\}_{i=1}^{n}{}_{\nu=1}^{p}$, and $\Sigma = \{R^1(x_i, x_j)\}_{i,j=1}^{n}$.

The following theorem shows that the smoothing spline estimate f_λ is a linear combination of the basis of $\mathcal{H}_0$ and the representer in $\mathcal{H}_1$ [134].

Theorem 3.3 (Kimeldorf-Wahba representer theorem[132]). *Suppose T is of full column rank and let $\phi_1, \ldots, \phi_p$ span the null space ($\mathcal{H}_0$) of P_1, then f_λ, the minimizer of PLS Eq.* (3.13)*, is given by*

$$f_\lambda = \sum_{\nu=1}^{p} d_\nu \phi_\nu + \sum_{i=1}^{n} c_i \xi_i \tag{3.19}$$

where

$$\xi_i = P_1\eta_i, \tag{3.20}$$

$$d = (T'M^{-1}T)^{-1}T'M^{-1}y, \tag{3.21}$$

$$c = M^{-1}(I - T)(T'M^{-1}T)^{-1}T'M^{-1})y, \tag{3.22}$$

$$M = \Sigma + n\lambda I, \tag{3.23}$$

$$\Sigma = < \xi_i, \xi_j >. \tag{3.24}$$

Proof [134]: Suppose the minimizer f_λ is in the form

$$f_\lambda = \sum_{\nu=1}^{p} d_\nu\phi_\nu + \sum_{i=1}^{n} c_i\xi_i + \rho \tag{3.25}$$

where ρ is some element in H_R perpendicular to $\phi_1, \ldots, \phi_p, \xi_1, \ldots, \xi_n$, therefore we have

$$L_i\rho = < \eta_i, \rho >= 0 \tag{3.26}$$

and

$$\|P_1 f\|^2 = \|\sum_{i=1}^{n} c_i\xi_i + \rho\|^2 = c^\top \Sigma c + \|\rho\|^2 \tag{3.27}$$

Then Eq. (3.13) becomes

$$\frac{1}{n}\|y - Td - \Sigma c\|^2 + \lambda\, c^\top \Sigma c + \|\rho\|^2 \tag{3.28}$$

It is obvious that Eq. (3.28) is minimized when $\|\rho\|^2 = 0$. Rewriting the PLS eq. (3.13) as

$$\frac{1}{n}\|y - Td - \Sigma c\|^2 + \lambda\, c^\top \Sigma c \tag{3.29}$$

and making the first derivatives w.r.t c and d equal to zero gives

$$\begin{aligned} (\Sigma + n\lambda\, I)\Sigma c + \Sigma T d = \Sigma y, \\ T^\top \Sigma c + T^\top T d = T^\top y \end{aligned} \tag{3.30}$$

It is easy to verify that the solution [134]

$$\begin{aligned} d =&(T' M^{-1} T)^{-1} T' M^{-1} y, \\ c =&M^{-1}(I - T)(T' M^{-1} T)^{-1} T' M^{-1})y, \end{aligned} \tag{3.31}$$

where $M = \Sigma + n\lambda\, I$ satisfies. ■

Utreras [127] showed that the following equation

$$\begin{aligned} (\Sigma + n\lambda\, I)c + Td = y, \\ T^\top c = 0 \end{aligned} \tag{3.32}$$

is equivalent to Eq. (3.30). Solving c and d for Eq. (3.32) provides a numerically friendly solution. Let the QR decomposition of T be

$$T = [Q_1\, Q_2] \begin{bmatrix} R \\ 0 \end{bmatrix}$$

where $Q = [Q1\, Q2]$ is an orthogonal matrix, Q_1 is n by p, Q_2 is n by $n - p$ and R is p

by p upper triangular. Skipping the steps in between (see [132]) on the manipulation, we give the expressions of c and d directly

$$\begin{aligned} c &= Q_2(Q_2^\top M Q_2)^{-1} Q_2^\top y, \\ d &= R^{-1} Q_1^\top (y - Mc) \end{aligned} \tag{3.33}$$

We conclude this section by the following definition that plays an important role in Sec. 3.5.

Definition 3.5 (Influence matrix[132]). *$H(\lambda\,)$ is called the influence matrix if it satisfies*

$$\begin{pmatrix} L_1 f_\lambda \\ \vdots \\ L_n f_\lambda \end{pmatrix} = H(\lambda\,) y \tag{3.34}$$

With $\rho = 0$ in Eq. (3.25), we have

$$\begin{pmatrix} L_1 f_\lambda \\ \vdots \\ L_n f_\lambda \end{pmatrix} = Td + \Sigma c \tag{3.35}$$

Therefore $H(\lambda\,)y = Td + \Sigma c$. Subtracting this from the first equation in Eq. (3.32) gives $H(\lambda\,)y = y - n\lambda\, c$. Further, substituting c with Eq. (3.33)

$$H(\lambda\,) = I - n\lambda\, Q_2(Q_2^\top M Q_2)^{-1} Q_2^\top \tag{3.36}$$

3.4 Splines for Surface Assimilation

A comprehensive set of smoothing splines have been developed for different applications such as polynomial splines, periodic splines, spherical splines, vector splines. To our interests, we focus on the periodic splines for closed 2D curve assimilation and spherical splines for closed 3D surface assimilation. This section was adapted from [134].

3.4.1 Periodic splines

When the object is in form of a closed smooth curve, the polynomial splines will cause error at the starting and end points (actually overlapped). Thus we probably need splines defined on the circle satisfying the periodic boundary condition. Without loss of generality, we assume f is a periodic function defined on the unit circle with the domain $\chi = [0, 1]$.

Let $k_r(x) = B_r(x)/r!$ be scaled Bernoulli polynomials where $B_r(x)$, $r = 0, 1, \ldots, x \in [0, 1]$ satisfy the recursive definition [132]

$$B_0(t) = 1$$

$$\frac{1}{r}\frac{d}{dt} = B_{r-1}(t), \qquad \int_0^1 B_r(u)\,\mathrm{d}u = 0, r = 1, 2, \ldots$$

From [1] the following formula holds

$$B_{2m}(x) = (-1)^{m-1} 2(2m)! \sum_{\nu=1}^{\infty} \frac{\cos 2\pi\nu x}{(2\pi\nu)^{2m}}, \qquad x \in [0, 1]$$

The Sobolev Hilbert space for periodic spline of order m is [134]

$$W_2^m(per) = \big\{f : f^{(j)}\text{are absolutely continuous}, f^{(j)}(0) = f^{(j)}(1), j = 0, \ldots, m-1, \\ \int_0^1 (f^{(m)})^2 \,\mathrm{d}x < \infty\big\}$$

Craven [28] derived that the space $W_2^m(per)$ is an RKHS with inner product

$$< f, g >= \Big(\int_0^1 f \,\mathrm{d}x\Big)\Big(\int_0^1 g \,\mathrm{d}x\Big) + \int_0^1 f^{(m)} g^{(m)} \,\mathrm{d}x \tag{3.37}$$

And $W_2^m(per)$ can be decomposed into two orthogonal subspace as [134]

$$W_2^m(per) = \mathcal{H}_0 \oplus \mathcal{H}_1 \tag{3.38}$$

where

$$\mathcal{H}_0 = span\{1\}, \qquad \mathcal{H}_1 = \big\{f \in W_2^m(per) : \int_0^1 f \,\mathrm{d}x = 0\big\} \tag{3.39}$$

$W_2^m(per)$ is an RKHS with the corresponding RK: $R(x, z) = R_0 + R_1$

$$R_0(x, z) = 1,$$

$$R_1(x, z) = (-1)^{m-1} k_{2m}(|x - z|)$$

recall $k_{2m}(\cdot)$ is a scaled Bernoulli polynomials (see above). In the experiments we chose $m = 2$, i.e., the cubic periodic spline as the building elements for f.

3.4.2 Thin-plate splines

Introduced by Duchon [45] for geometric design, thin-plate splines (TPS) were adapted [132, 65] for spatial smoothing problems due to their rotation-invariant property. Given a set of control nodes $\{D_i : i = 1, 2, \ldots, K\}$, a two dimensional TPS

$$f(x_1, x_2) = \sum_{i=1}^{K} c_i U(\|(x_1, x_2) - D_i\|) \tag{3.40}$$

where $U(r) = -r^2 \log r^2$ is the radial basis kernel function. For d dimensional m order thin plate spline, the smoothness penalty of f is [134]

$$J_m^d(f) = \sum_{\alpha_1+\cdots+\alpha_d=m} \frac{m!}{\alpha_1! \ldots \alpha_d!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{\partial^m f}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \right)^2 \prod_{j=1}^{d} \mathrm{d}x_j \tag{3.41}$$

In our experiment we chose $d = 2$ and $m = 2$, the smoothness penalty becomes

$$J_2^2(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(f_{x_1x_1}^2 + 2f_{x_1x_2}^2 + f_{x_2x_2}^2 \right) \mathrm{d}x_1 \mathrm{d}x_2$$

Define the Sobolev Hilbert space for a thin-plate spline as

$$W_2^m(\mathbb{R}^d) = \{f : J_m^d(f) < \infty\} \tag{3.42}$$

$W_2^m(\mathbb{R}^d)$ is a RKHS if and only if $2m - d > 0$ with an inner product defined as [45]

$$< f, g >= \sum_{\alpha_1+\cdots+\alpha_d=m} \frac{m!}{\alpha_1! \ldots \alpha_d!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \\ \left(\frac{\partial^m f}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}}\right)\left(\frac{\partial^m g}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}}\right) \prod_{j=1}^{d} \mathrm{d}x_j \tag{3.43}$$

The minimizer to the following PLS (Penalized Linear Smoothing) problem is a thin-plate spline function f

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - f(\mathbf{x}_i))^2 + \lambda\, J_m^d(f) \tag{3.44}$$

in $W_2^m(\mathbb{R}^d)$ where $\mathbf{x_i} = (x_1^i, x_2^i, \ldots, x_d^i)$. The null space $\mathcal{H}_0$ of the penalty functional J_m^d is of the dimensionality $M = \begin{pmatrix} d+m-1 \\ d \end{pmatrix}$ and spanned by polynomials denoted as $\phi_1, \ldots, \phi_M$ in d variables of total degree up to $m-1$ [132]. In our experiments we chose $d = 2$ and $m = 2$, then $M = 3$, and $\mathcal{H}_0$ is spanned by $\phi_1(\mathbf{x}) = 1$, $\phi_2(\mathbf{x}) = x_1$, $\phi_3(\mathbf{x}) = x_2$.

To avoid the tedious computation of the RK of $W_2^m(\mathbb{R}^d)$, Duchon [45] and Wahba [132] showed the solution to Eq. (3.44)

$$\hat{f}(\mathbf{x}) = \sum_{\nu=1}^{p} d_\nu \phi_\nu(\mathbf{x}) + \sum_{i=1}^{n} c_i \xi_i \tag{3.45}$$

can be approximated by replacing ξ_i with $E_m(\mathbf{x}, \mathbf{x}_i)$ under the condition that $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are such that least square regression on $\phi_1, \ldots, \phi_M$ is unique where E_m is the Green function for the m-iterated Laplacian $E_m(\mathbf{x}, \mathbf{z}) = E(\|\mathbf{x} - \mathbf{z}\|)$, and

$$E(u) = \begin{cases} (-1)^{\frac{d}{2}+1+m} |u|^{2m-d} \log |u|, & d \text{ even} \\ \\ |u|^{2m-d}, & d \text{ odd} \end{cases}$$

For genuine RK and the spline computation, see [65].

Let $T = \{\phi_\nu(x_i)\}_{i=1}^{n}{}_{\nu=1}^{p}$ and $K = \{E_m(x_i, x_j)\}_{i,j=1}^{n}$. Assume that T is of full column rank, the coefficients c and d are solutions to

$$\begin{aligned} (K + n\lambda\, I)c + Td &= y, \\ T^\top c &= 0 \end{aligned}$$

3.4.3 Spherical splines

Spherical splines are the fundamental tools for the surface assimilation on the closed surfaces. By combining the periodic and thin-plate splines, Wahba [131] developed the spherical splines for smoothing spline regression on the unit sphere domain. The Laplacian operator on the unit sphere $\mathcal{S}$ is defined as

$$\Delta f = \frac{1}{\cos^2 \phi} f_{\theta\theta} + \frac{1}{\cos \theta} (\cos \phi f_\phi)_\phi \tag{3.46}$$

where $\theta \in [0, 2\pi]$ is the longitude and $\phi \in [-\pi/2, \pi/2]$ is the latitude. The energy penalty of f is given by [134]

$$J(f) = \begin{cases} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\Delta^{\frac{m}{2}} f)^2 \cos\phi \, \mathrm{d}\phi d\theta, & m \, \text{is even} \\ \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left\{ \frac{(\Delta^{\frac{m-1}{2}} f)_\theta^2}{\cos^2\phi} + (\Delta^{\frac{m-1}{2}} f)_\phi^2 \right\} \cos\phi \, \mathrm{d}\phi d\theta, & m \, \text{is odd} \end{cases} \tag{3.47}$$

Let $\mathbf{x} = (\theta, \phi)$ the Sobolev Hilbert space

$$W_2^m(\mathcal{S}) = \left\{ f : \left\| \int_{\mathcal{S}} f \, \mathrm{d}\mathbf{x} \right\| < \infty, J(f) < \infty \right\} \tag{3.48}$$

$W_2^m(\mathcal{S})$ is an RKHS when $m > 1$ which can be decomposed into two orthogonal subspace as

$$W_2^m(\mathcal{S}) = \mathcal{H}_0 \oplus \mathcal{H}_1 \tag{3.49}$$

where

$$\begin{aligned} \mathcal{H}_0 =& span\{1\}, \\ \mathcal{H}_1 =& \left\{ f \in W_2^m(\mathcal{S}) : \int_{\mathcal{S}} f \, \mathrm{d}x = 0 \right\} \end{aligned}$$

The corresponding RK of $W_2^m(\mathcal{S})$ is given by [134]

$$\begin{aligned} R_0(\mathbf{x}, \mathbf{z}) &= 1, \\ R_1(\mathbf{x}, \mathbf{z}) &= \sum_{l=1}^{\infty} \frac{2l+1}{4\pi} \frac{1}{\{l(l+1)\}^m} P_l(\cos\gamma(\mathbf{x}, \mathbf{z})) \end{aligned}$$

where P_l are the Legendre polynomials and $\gamma(\mathbf{x}, \mathbf{z})$ is the angle between $\mathbf{x}$ and $\mathbf{z}$ [132].

Since the RK R_1 is the summation of an infinite series and its closed form expression only available for $m = 2$ or $m = 3$, the computation for R_1 is inconvenient. Wahba[131] proposed approximating R_1 by a closed-form RK Q_1. $\mathcal{H}_{Q_1}$ is topologically equivalent to $\mathcal{H}_{R_1}$.

3.5 Estimating the Smoothing Parameter

In our cost function (3.7), the weight λ is a crucial parameter for the overall performance. Over-smoothing and under-smoothing both leads to unsatisfactory results. As an important part of the theory of smoothing spline regression, automatic selection of the smoothing parameter is introduced in this section. Its limitation will be discussed in Sec. 3.7. Sec. 3.5.1 and Sec. 3.5.2 were adapted from [134].

3.5.1 Predictive Mean Squared Error

To estimate the smoothing parameter λ we need a criterion for fitting performance. Define the expected squared prediction error as

$$ET(\lambda) = \frac{1}{N} \sum_{i=1}^{N} (L_i f_\lambda - L_i f)^2 \tag{3.50}$$

where f_λ is the estimate of the unknown true function f. Under the smoothing spline regression model, f_λ is a function of λ indicated by the subscript. Suppose $g = (L_1 f, \ldots, L_N f)^\top$, with Eq. (3.34) we have $(L_1 f_\lambda, \ldots, L_N f_\lambda)^\top = H(\lambda) y = h(\lambda)(g + \epsilon)$, where $\epsilon \sim \mathcal{N}(0, \sigma^2 I)$. This error can be decomposed into two components:

$$\begin{aligned} ET(\lambda\,) =& \frac{1}{N}\mathrm{E}\,\|H(\lambda\,)(g+\epsilon) - g\|^2 \\ =& \frac{1}{N}\mathrm{E}\,\|(I - H(\lambda\,)g) + H(\lambda\,)\epsilon\|^2 \\ =& \frac{1}{N}\mathrm{E}\,\|(I - H(\lambda\,)g)\|^2 + \frac{2}{N}\mathrm{E}\,(I - H(\lambda\,))gH(\lambda\,)\epsilon + \frac{1}{N}\mathrm{E}\,\|H(\lambda\,)\epsilon\|^2 \\ =& \frac{1}{N}\|(I - H(\lambda\,)g)\|^2 + \frac{\sigma^2}{N}\,\mathrm{Tr}\,H^2(\lambda\,) \\ =& b^2(\lambda\,) + \sigma^2 v(\lambda\,) \end{aligned} \tag{3.51}$$

where b^2 and $\sigma^2 v$ are known as squared bias and variance, respectively [134]. With Sec. 3.3.2, we have

$$I - H(\lambda\,) = N\lambda\, Q_2(Q_2^\top(\Sigma + N\lambda\, I)Q_2)^{-1}Q_2^\top \tag{3.52}$$

Let $UDU^\top$ be the eigenvector eigenvalue decomposition of $Q_2^\top \Sigma Q_2$ and $\Gamma = Q_2 U$,

$$I - H(\lambda\,) = N\lambda\,\Gamma(D + n\lambda\, I)^{-1}\Gamma^\top \tag{3.53}$$

Further, with $h = \Gamma^\top g$, $b^2(\lambda\,)$ and $v(\lambda\,)$ can be expressed as [132]:

$$\begin{aligned} b^2(\lambda\,) =& \frac{1}{N}\sum_{\nu=1}^{N-p}\Big(\frac{N\lambda\, h_\nu}{\lambda_{\,\nu} + N\lambda}\Big)^2 \\ v(\lambda\,) =& \frac{1}{N}\left(\sum_{\nu=1}^{N-p}\Big(\frac{\lambda_{\,\nu}}{\lambda_{\,\nu} + N\lambda}\Big)^2 + p\right) \end{aligned} \tag{3.54}$$

where h_ν are the elements of h and λ_ν are the diagonal elements of D. Wahba [132] showed that $ET(\lambda)$ have at least one minimizer $\lambda^* > 0$ because $\frac{\mathrm{d}b^2(\lambda)}{\mathrm{d}\lambda}|_{\lambda=0} = 0$ and $\frac{\mathrm{d}v(\lambda)}{\mathrm{d}\lambda}|_{\lambda=0} < 0$ lead to $\frac{\mathrm{d}ET(\lambda)}{\mathrm{d}\lambda}|_{\lambda=0} < 0$.

3.5.2 Generalized Cross Validation

$b^2(\lambda)$ in Eq. (3.51) is a monotone increasing function with λ. $v(\lambda)$ is a decreasing one. Therefore λ balances the trade-off between the bias and variance. This smoothing parameter deals with the model complexity and eventually controls the under- and over- fitting [132]. The minimizer to $ET(\lambda)$ suggests a straightforward choice of λ. However, since $b^2(\lambda)$ depends on the unknown true function f (the variance $v(\lambda)$ does not), $ET(\lambda)$ depends on f. It is not practical to compute λ from $ET(\lambda)$.

Independent of f, the cross-validation provides an estimate of $ET(\lambda)$. Let $f_\lambda^{[k]}$ be the minimizer of Eq. (3.13) involving all observational data except y_k

$$\frac{1}{N}\sum_{\substack{i=1\\ i\neq k}}^{N}\Big(y_i - L_i f\Big)^2 + \lambda\,\|P_1\lambda\,\|^2 \tag{3.55}$$

The cross-validation estimate of the expected squared prediction error (3.50) is

$$CV(\lambda) = \frac{1}{N}\sum_{k=1}^{N}\Big(y_k - L_i f_\lambda^{[k]}\Big)^2 \tag{3.56}$$

Minimizing CV(λ) will give an estimate of λ. However, the computation of $f_\lambda^{[k]}$ needs expensive iteration from $i = 1, \ldots, N$. With the following theoremthe computation can be reduced. $CV(\lambda)$ can be computed using the whole observed data and the diagonal elements of the influence matrix [132].

Theorem 3.4 ([132]).

$$CV(\lambda) = \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - L_i f_\lambda)^2}{(1-h_{ii})^2} \tag{3.57}$$

where h_{ii} is the ii-th entry of the influence matrix $H(\lambda)$ given in Sec. 3.3.2.

Replacing h_{ii} by $\mu_1(\lambda) = \frac{1}{N}\sum_{i=1}^{N} h_{kk}(\lambda) = \frac{1}{N}\operatorname{Tr} H(\lambda)$, we have the generalized cross-validation (GCV)

$$\begin{aligned} GCV(\lambda) =& \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - L_i f_\lambda)^2}{(1-\mu_1(\lambda))^2} \\ =& \frac{1}{N}\|(I - H(\lambda))y\|^2 / \Big[\frac{1}{N}\operatorname{Tr}(I - H(\lambda))\Big]^2 \end{aligned} \tag{3.58}$$

Minimizing Eq. (3.58) gives GCV estimate of λ. The “weak GCV theorem” [28] shows that $GCV(\lambda)$ is a consistent and robust estimate of $ET(\lambda)$ [132].

3.5.3 Fast Algorithm for Evenly Spaced Gridded Data

Garcia [60] proposed a fast numerical method minimizing the Eq. (3.4) under the assumptions: 1) a large amount of observational data is available. 2) observational data are evenly spaced on Euclidean space $\mathbb{R}^n$. 3) f is smooth in the sense that f has up to order 2 continuous derivatives, i.e., $m = 2$ in (3.4). Though this method has not been extended to the unit space $\mathcal{S}$, it can be used in the surface assimilation for curves and tubular objects. Our curves or surfaces are generally smooth. With the re-parameterization technique, it is easy to make ready the arbitrary evenly spaced data. Therefore the assumptions above holds.

This method works on a grid of n-dimensions, $n = 1, 2, 3, \cdots$. It uses the discrete cosine transform (DCT) and the fast Fourier transform (FFT) to speed up the computation of the Laplacian regularization. It can smooth closed curves if the periodic boundary condition is satisfied, and open surfaces if the surfaces are mapped to the planar domain. The downside of this algorithm is that it is not directly applicable for closed surface mapped to the unite sphere. Bülow [13] solved the diffusion equation on spherical surface. And Healy [71] published the software package for spherical FFT. As such a fast method for spherical surface assimilation is possible.

3.6 Experiments

We have applied the surface assimilation algorithm to the synthetic and the patient data. In each experiment we measured the performance improvements before and after the use of the algorithm. The performance metrics can be divided into two categories: 1) spatial overlap metric 2) spatial distance metric.

Relative Overlap (RO) between two segmentations measures how well two binary objects are matched. It is a common metric for evaluating segmentation and registration algorithms. Given two corresponding regions of interest (ROI) A and B, RO is defined as

$$\mathrm{RO}\,(A, B) = \frac{\mathrm{volume}(A \cap B)}{\mathrm{volume}(A \cup B)} \tag{3.59}$$

where RO $= 0$ means the two ROIs are disjoint. RO $= 1$ means the two ROIs are fully overlapped. When RO is used as the evaluation metric, the segmentation or registration algorithm with the highest RO has the best performance. RO has the following two weaknesses: 1) It is biased. Given an algorithm, a large ROI tends to have higher RO than a small one. 2) It is global. It ignores the quality of local structure matching.

Dice Similarity Coefficient (DSC) is another spatial overlap metric for segmentation benchmark [42]. It is a concept in κ-statistics commonly used for reliability studies [37]. For two binary segmentations A and B, the DSC is defined as

$$\mathrm{DSC}\,(A, B) = \frac{2(\mathrm{volume}(A \cap B))}{\mathrm{volume}(A) + \mathrm{volume}(B))} \tag{3.60}$$

Similar to RO, the range of DSC is $[0, 1]$ with 1 indicating a perfect overlap. However since the distribution of DSC usually leans towards the value of 1, it is hard to apply the statistics hypothesis testing techniques.

Logit Dice Similarity Coefficient (LDSC) can be used to overcome the limitation of RO and DSC. Similar to [148], we apply the logit function to transform the distribution of DSC into a normal distribution. LDSC is defined by:

$$\mathrm{LDSC}\,(A, B) = \ln\left(\frac{\mathrm{DSC}\,(A, B)}{1 - \mathrm{DSC}\,(A, B)}\right) \tag{3.61}$$

where the logit function maps the domain of DSC from $[0, 1]$ to the range of $[-\infty, \infty]$. Note that both RO and DSC could be biased and erroneous. Compared to the small binary objects, the larger ones are prone to have higher spatial overlap. Therefore other evaluation measures are required for an objective benchmark.

Mean Absolute Distance (MAD) is a performance metric based on distance. The geometric differences, e.g., total curvature or distance, between two curves or surfaces before and after data assimilation can also be used for performance evaluation. MAD represents the global disagreement between two surfaces.

Let $A = a_1, a_2, \ldots, a_n$, $B = b_1, b_2, \ldots, b_m$, a_i, b_i representing the boundary of the surfaces S_1 and S_2, MAD is defined as:

$$\text{MAD } = \frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n} d(a_i, B) + \frac{1}{m}\sum_{j=1}^{m} d(b_j, A)\right) \tag{3.62}$$

where $d(a_i, B) = \min_j \|b_j - a_i\|$.

Hausdorff Distance (HD) is another distance-based metric that compares the local similarities between two objects [87]. It is defined as

$$\text{HD } = \max(\max_i d(a_i, B), \max_j d(b_j, A)) \tag{3.63}$$

MAD measures the average closeness, and HD emphasizes the closeness in worst case.

3.6.1 Experiment with Synthetic Curves

We designed a proof-of-concept experiment with 2D curves to demonstrate our algorithm. A smooth and closed parametric curve $\mathbf{x}(t) = (x_1(t), x_2(t))$ represents the region of interest (ROI). $\mathbf{x}(t)$ is actually an instance generated by a statistical shape

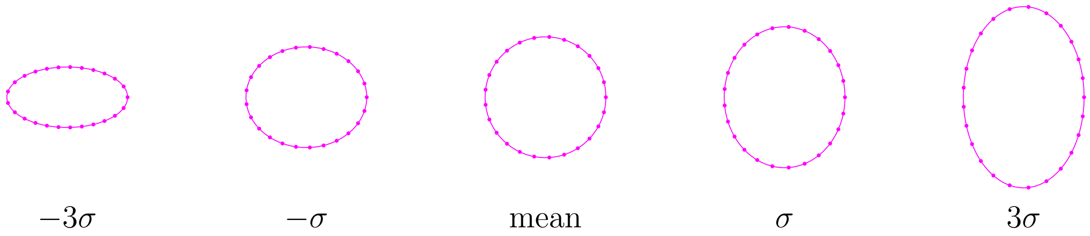


Figure 3.2: Five instances of the true ROIs with different shape parameter w. In the middle, the mean shape is obtained when $w = 0$. The shape model has only one variation mode, which along the vertical direction.

model. For simplicity, we choose the circle as the mean shape of $\mathbf{x}(t)$ and restrict $x(t)$ to have only one shape variation mode changing the height of $\mathbf{x}(t)$ along the vertical axis. Mathematically,

$$x_1(t) = \cos(t)$$

$$x_2(t) = \sin(t) + w\sin(t)$$

where $t \in [0, 2\pi)$, $(\cos(t), \sin(t))$ forms the mean shape and w is a normally distributed variable, $w \sim N(0, \sigma^2)$ and $\sigma = \frac{1}{6}$. Fig. 3.2 shows five shape samples ranging from $w = -3\sigma$ to $w = 3\sigma$.

To simulate the segmentations $\mathbf{y}(t) = (y_1(t), y_2(t))$ among the different observers, we adopt the inter-observer segmentation error (ISE) model with the following assumptions: 1) the boundary of $\mathbf{x}(t)$ at the north and south poles are assumed to be known, therefore the segmentations there are accurate, and ISE is small. 2) the boundary of $\mathbf{x}(t)$ at the equator is unclear, thus the segmentation at the equator is inaccurate leading to large ISE. For simplicity, we only simulate one mode of segmentation variation which is purposely set perpendicular to that of $\mathbf{x}(t)$. Mathematically,

$$y_1(t) = x_1(t) + p\cos(t)$$

$$y_2(t) = x_2(t)$$

where $t \in [0, 2\pi)$, $x_1(t), x_2(t)$ presents the mean shape of the segmentation, the random variable $p \sim N(0, \sigma^2)$ and $\sigma = \frac{1}{6}$ controlling the shape of observed curve.

We show a group of simulated true and observed curves in Fig. 3.3. From the top to bottom, the true curves (pink) $\mathbf{x}(t)$ with $w = -\sigma$, 0, σ are shown. From the left to right, five observed curves with the shape parameter p ranging from -3σ to 3σ.

An example of estimating the true curve from the observed one is illustrated in Fig. 3.4. The truth is colored pink and the observed is black. The equi-distant dots along the black curve show the parameterization of the curve. The same parameterization are used through out all other curves but not drawn. By minimizing Eq. (3.6), the shape model predicts the cyan curve. This curve and the black curve are fused into the blue curve by following Eq. (3.8) where we choose $\alpha_i = |1 - \sin(t_i)|$. In such a way, the position of the new curve at the poles is primarily contributed by the observations (black curve) and the position at the equator are dominantly affected by the model curve (cyan curve). The blue curve presents a unrealistic dent around the equator. That is removed by the smoothing spline regression, giving the

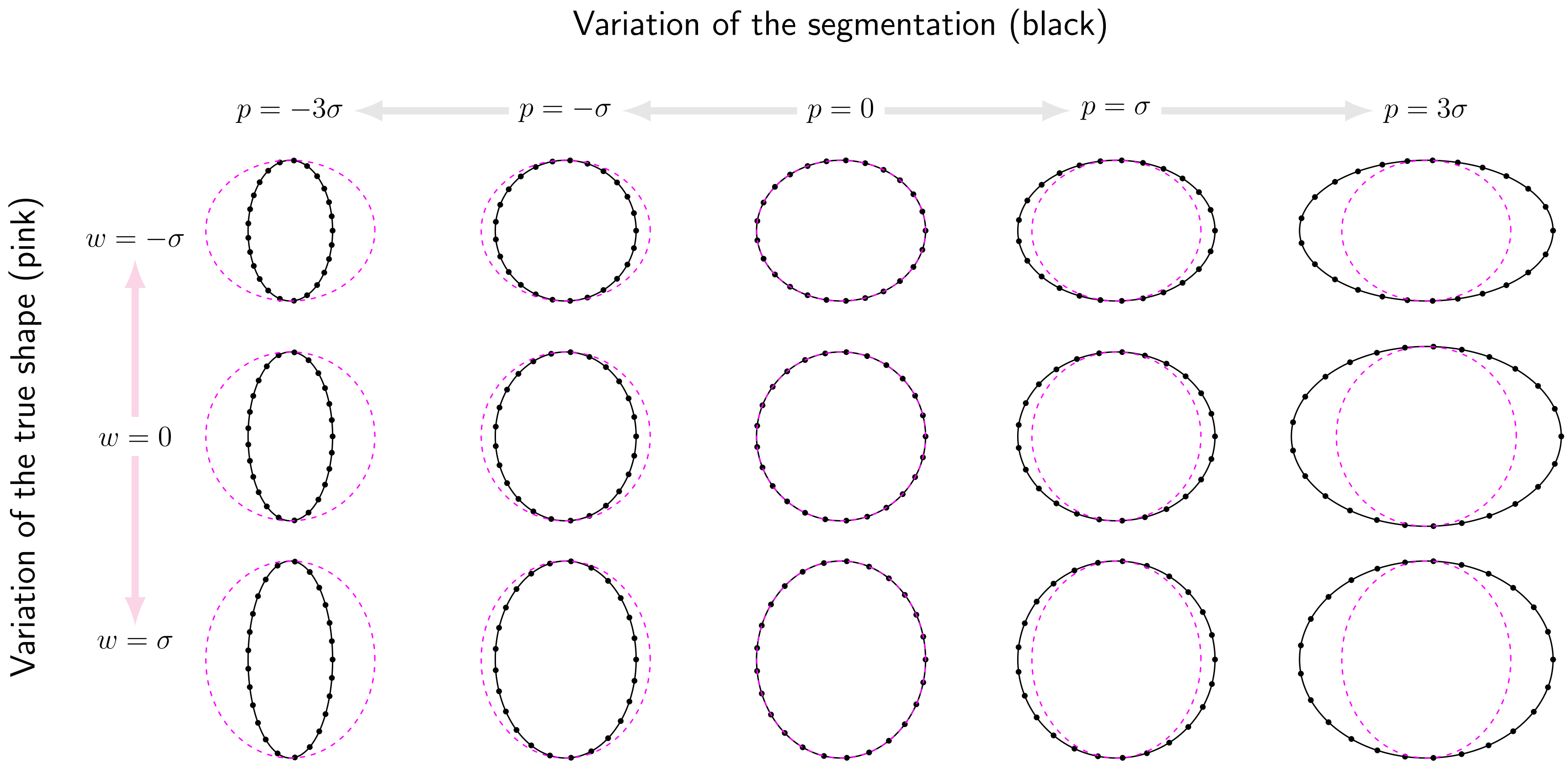


Figure 3.3: Illustration of the true and observed shapes in the 2D phantom experiment 1. The pink circles are the true. The black ones are the observed (simulated) segmentations. Each row, from left to right, shows five segmentations with different shape parameter p, overlaying the true shape having the shape parameter w. The true shape varies along the vertical direction. The segmentation varies along the horizontal direction. Therefore large inter-observer segmentation error happens at the equator.

Table 3.1: Results of 2D synthetic curve assimilation, Case 1: the shape variations of the true ROI and the segmentation are perpendicular

	RO		DSC		LDSC		MAD		HD	
	Before	After	Before	After	Before	After	Before	After	Before	After
Mean	0.877	0.942	0.932	0.970	3.041	3.798	0.067	0.031	0.134	0.051
Median	0.892	0.952	0.943	0.975	2.799	3.680	0.057	0.026	0.115	0.041
STD	0.090	0.044	0.055	0.024	1.167	0.904	0.051	0.023	0.102	0.034
IQR	0.124	0.058	0.069	0.030	1.381	1.291	0.069	0.029	0.141	0.044
Mean improvement	7.4% ($p < 0.001$)		4.0% ($p < 0.001$)		24.9% ($p < 0.001$)		53.0% ($p < 0.001$)		61.9% ($p < 0.001$)	

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance, STD: Standard Deviation, IQR: Interquartile Range

final surface assimilation result, the red curve. Comparing with the red curve with the black curve, we can see the red one reduces the segmentation error that occurs along the horizontal direction.

We simulated the true and observed curves for 2500 times. For each pair, we ran the surface assimilation algorithm to estimate the true. This estimate was compared to the observed curve with the evaluation measures mentioned in the previous section. The average evaluation results are illustrated in Fig. 3.5 and numerically summarized in Table 3.1. The mean improvements on RO, DSC, LDSC, MAD and HD are 7.4%, 4.0%, 24.9%, 53.0% and 61.9%, respectively and statistically significant ($p < 0.001$).

We also designed a counter example on 2D curves to show the limitation of our algorithm. The instances of the true curves $\mathbf{x}(t)$ are as same as those used in the previous experiment. Instead of being perpendicular to the true curve's variation direction, the observed one $\mathbf{y}(t)$ varies along the same direction of the true.

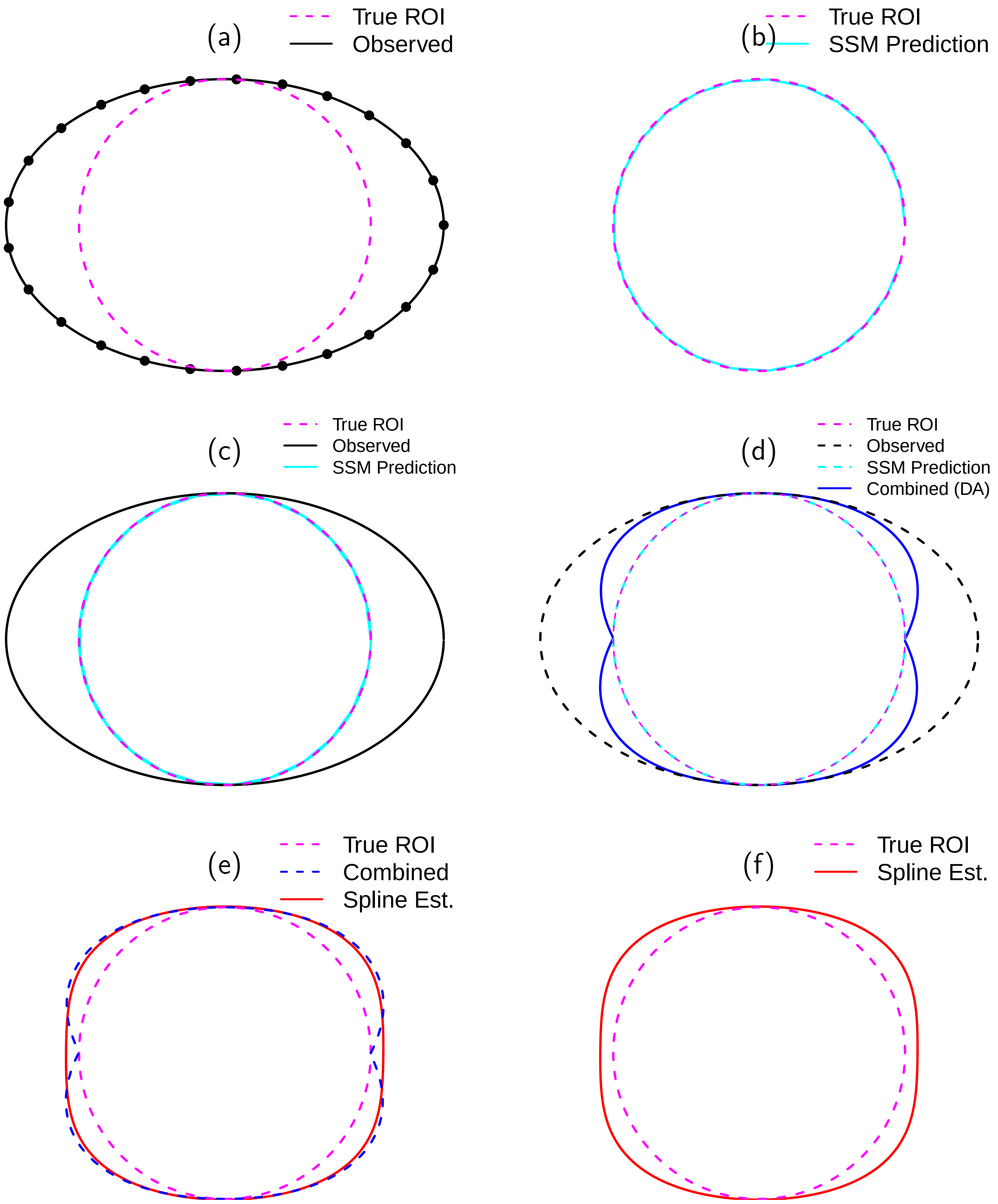
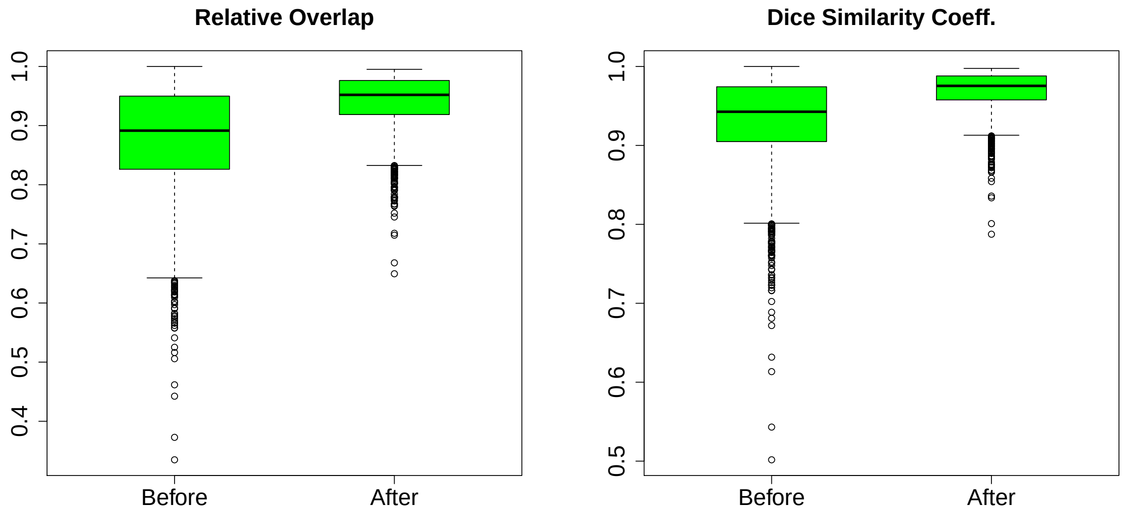
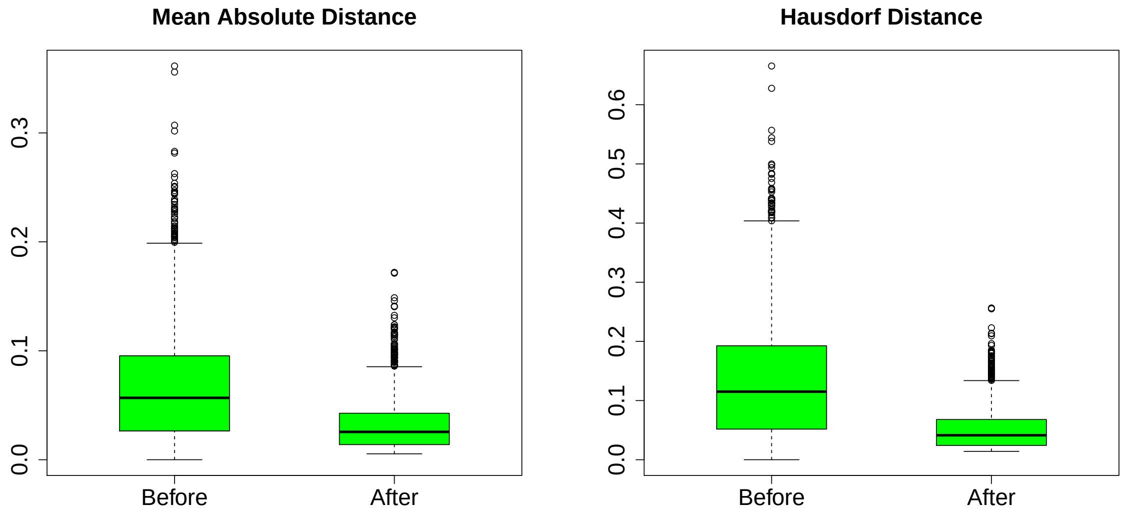


Figure 3.4: Illustration of surface assimilation algorithm in 2D phantom experiment 1. (a) shows a possible observed segmentation (black) of a true shape (pink). The statistical shape model (SSM) predicts a new segmentation (cyan) in (b). (c) shows the true, the observed and the SSM-predicted together. The latter two are combined into the blue one in (d). The estimated true shape by the spline regression is shown red in (e) and (f)



(a) RO, higher is better

(b) DSC, higher is better



(c) MAD, lower is better

(d) HD, lower is better

Figure 3.5: Results of 2D synthetic curve assimilation, Case 1: the shape variations of the true ROI and the segmentation are perpendicular. Each plot shows the results before and after the curve assimilation.

Mathematically,

$$
\begin{aligned}
y_1(t) &= x_1(t) \\
y_2(t) &= x_2(t) + p\cos(t)
\end{aligned}
$$

where $t \in [0, 2\pi)$, $x_1(t), x_2(t)$ presents the mean shape of the segmentation, the random variable $p \sim N(0, \sigma^2)$ and $\sigma = \frac{1}{6}$ controlling the shape of observed curve.

A set of simulated true and observed curves is shown in Fig. 3.6. From the top to bottom, the true curves (pink) $\mathbf{x}(t)$ with $w = -\sigma$, 0, σ are shown. From the left to right, five observed curves with the shape parameter p ranging from -3σ to 3σ. We choose $\alpha_i = |\cos(t_i)|$ to simulate the segmentation error along the curve, i.e., large error happens at the poles and small error happens at the equator.

Similarly, we simulated the true and observed curves for 2500 times and report the evaluation results in Fig. 3.8 and Table 3.2. Statistically, the curve assimilation makes no difference on RO, DSC, LDSC, MAD and HD in this case.

We use an example in Fig. 3.7 to show the limitation of the algorithm. The true ROI is a pink circle at the center. Given an observed curve (black) similar to a plausible instance of the true shape, the shape model is unable to distinguish the two and predicts a cyan curve close to the observed one. Consequently the blue one, combination of the cyan and the black, is close to the observed. So is the final assimilated curve, the red one.

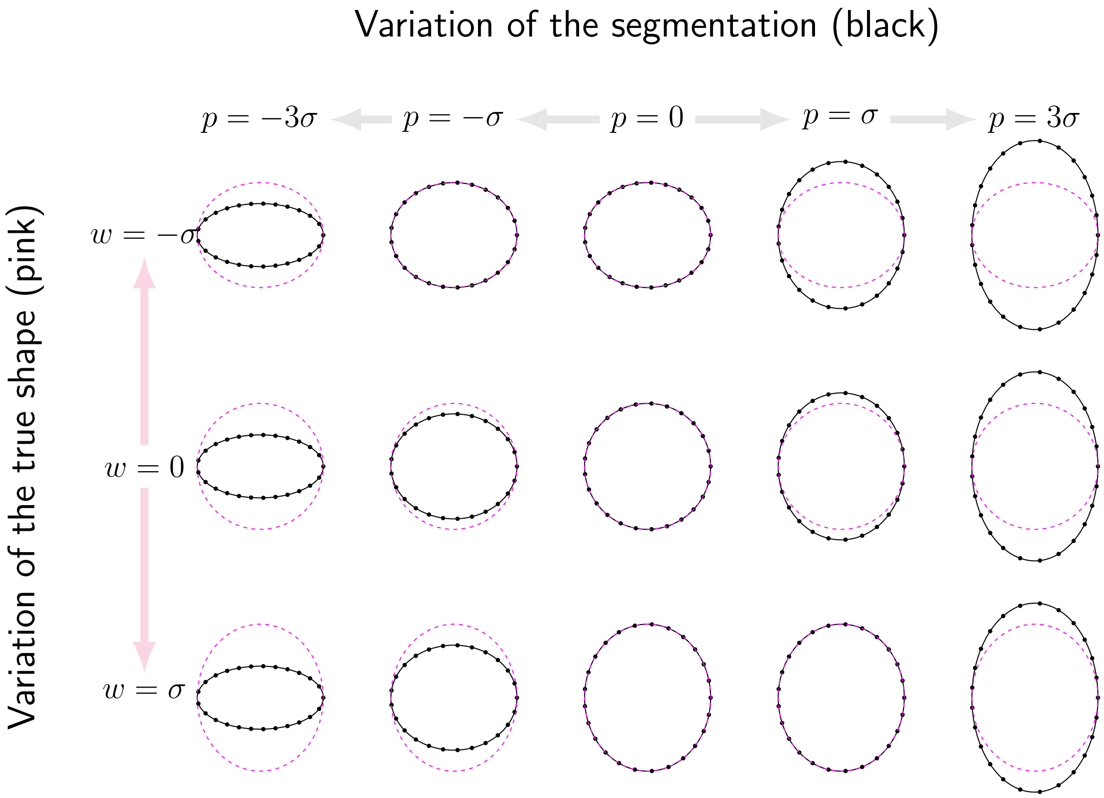


Figure 3.6: Illustration of the true and observed shapes in the 2D phantom experiment 2. The pink circles are the true. The black ones are the observed (simulated) segmentations. Each row, from left to right, shows five segmentations with different shape parameter p, overlaying the true shape having the shape parameter w. The true shape and the segmentation both vary along the vertical direction. Large inter-observer segmentation error happens at the poles. This is a counter example of surface assimilation (SA) where SA can not estimate a segmentation closer to the truth.

Table 3.2: Results of 2D synthetic curve assimilation, Case 2: the shape variations of the true ROI and the segmentation are in the same direction

	RO		DSC		LDSC		MAD		HD	
	Before	After	Before	After	Before	After	Before	After	Before	After
Mean	0.838	0.839	0.907	0.908	2.692	2.674	0.092	0.092	0.185	0.183
Median	0.861	0.862	0.925	0.924	2.519	2.527	0.075	0.075	0.150	0.150
STD	0.116	0.115	0.073	0.071	1.154	1.081	0.071	0.070	0.144	0.141
IQR	0.171	0.171	0.100	0.095	1.465	1.467	0.099	0.098	0.201	0.198
Mean improvement	0.09% ($p = 0.82$)		0.05% ($p = 0.81$)		-0.66% ($p = 0.57$)		0.23% ($p = 0.92$)		0.80% ($p = 0.71$)	

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance, STD: Standard Deviation, IQR: Interquartile Range

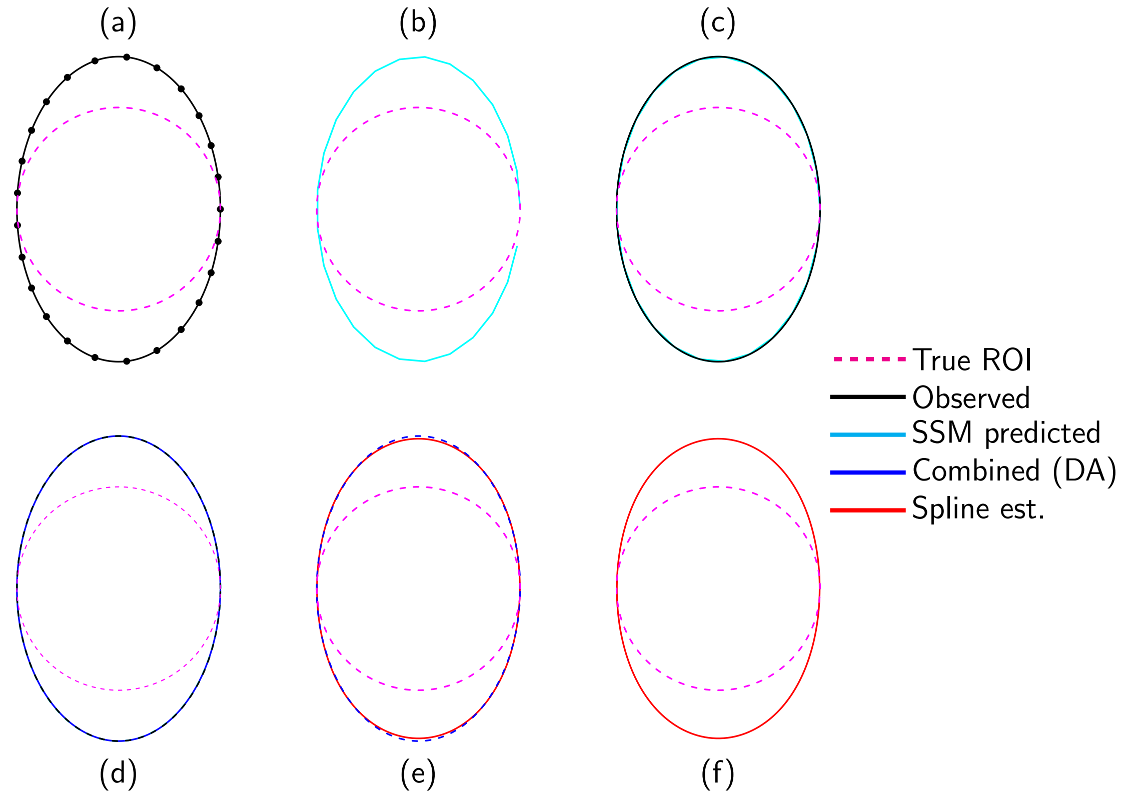
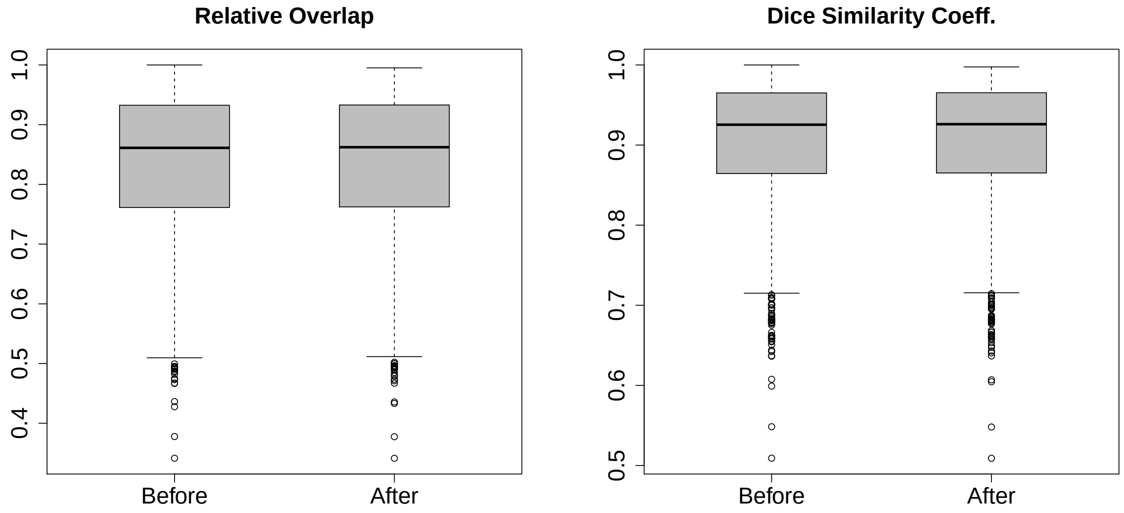
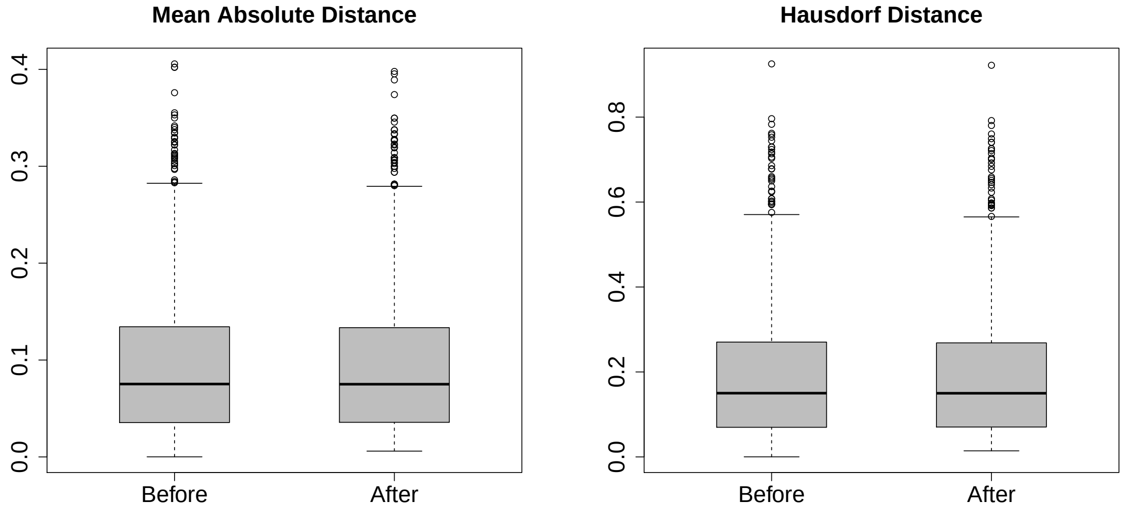


Figure 3.7: Illustration of surface assimilation algorithm in 2D phantom experiment 2. (a) shows a possible observed segmentation (black) of a true shape (pink). The statistical shape model (SSM) predicts a new segmentation (cyan) in (b). (c) shows the true, the observed and the SSM-predicted together. The latter two are combined into the blue one in (d). The estimated true shape by the spline regression is shown red in (e) and (f)



(a) RO, higher is better

(b) DSC, higher is better



(c) MAD, lower is better

(d) HD, lower is better

Figure 3.8: Results of 2D synthetic curve assimilation, Case 2: the shape variations of the true ROI and the segmentation are in the same direction. Each plot shows the results before and after the curve assimilation.

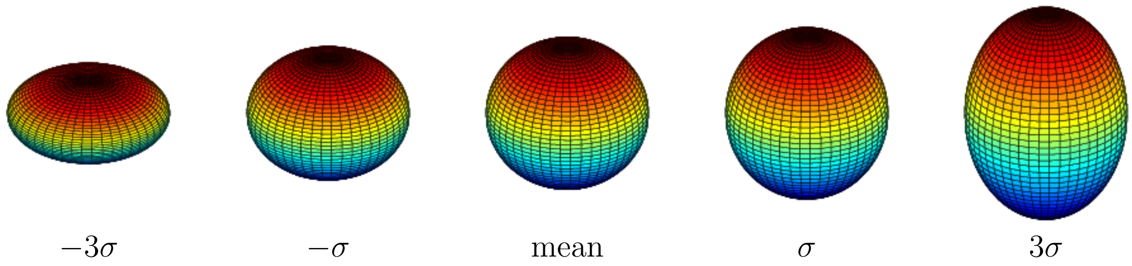


Figure 3.9: Five true ROI instances with different shape parameter w.

3.6.2 Experiment with Synthetic Surfaces

We extended the experiment of 2D curve assimilation to the 3D surfaces. A smooth and closed parametric surface $\mathbf{x}(u,v) = (x_1(u,v), x_2(u,v), x_3(u,v))$ represents the region of interest (ROI). For simplicity, we choose a sphere as the mean shape of $\mathbf{x}(u,v)$ and restrict $\mathbf{x}(u,v)$ only changes along the vertical axis. Mathematically,

$$x_1(u,v) = r\cos(u)\cos(v)$$

$$x_2(u,v) = r\cos(u)\sin(v)$$

$$x_3(u,v) = r\sin(u) + w\sin(u)$$

where $r = 24$ simulating the radius of a regular prostate, $u \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $v \in [0,\, 2\pi)$ and w is a normally distributed variable, i.e., $w \sim N(0, \sigma^2)$, $\sigma = 4$. Fig. 3.9 shows five true shape samples from the model.

We follow the similar fashion as the 2D curve experiment to simulate the 3D inter-observer segmentations $\mathbf{y}(u,v) = (y_1(u,v), y_2(u,v), y_3(u,v))$. Mathematically,

$$\begin{aligned} y_1(u,v) &= x_1(u,v) \\ y_2(u,v) &= x_2(u,v) + p\cos(u)\text{sin}(v) \\ y_3(u,v) &= x_3(u,v) \end{aligned}$$

where $(x_1(u,v), x_2(u,v), x_3(u,v))$ serves the mean shape of the segmentation, $p \sim N(0, \sigma^2)$, $\sigma = 4$ controlling the shape of observed surface along the horizontal direction. This segmentation model has only one variation mode, perpendicular to the true shape model's. In such way, The inter-observer segmentation error (ISE) is large at the left and right ends of the equator. ISE is small at the poles. The ISE was quantified using Eq. (2.27) and normalized to $[0,1]$ to help the fusion of the observed and shape-model-predicted surfaces in Eq. (3.8). A group of observed surfaces are shown in Fig. 3.10 where the true surfaces are not drawn. The color on the surfaces is for visualization of correspondence among surfaces and not for indicating the segmentation error.

We simulated the true and observed surfaces for 1000 times and computed the assimilated surfaces. The evaluation scores of each observed surface and its assimilated surface against the true curve are collected. The average performance comparison is reported in Fig. 3.11 and Table 3.3. The results show a statistically significant improvement is achieved.

Two specific examples are provided in Fig. 3.12 and Fig. 3.13 to demonstrate

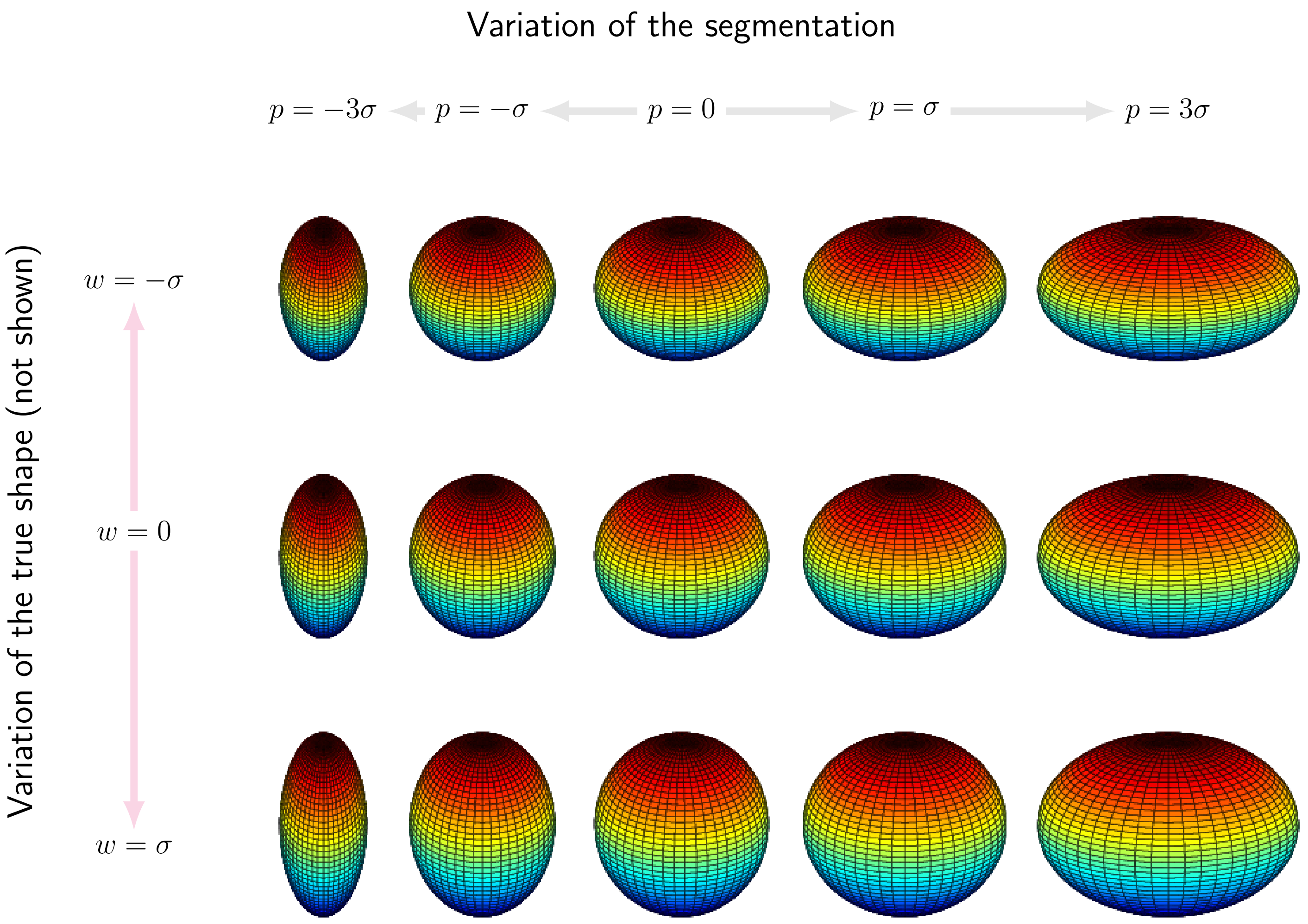
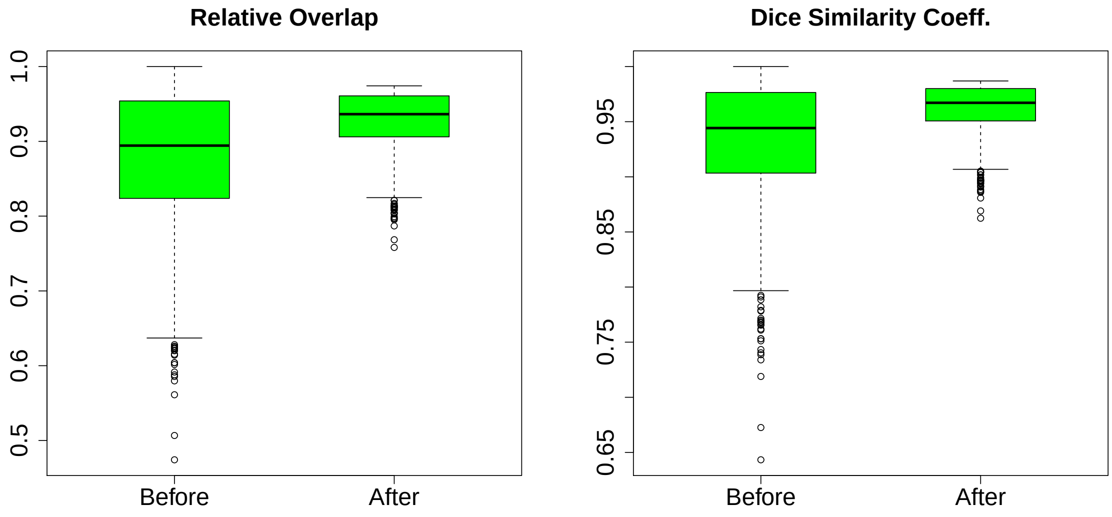
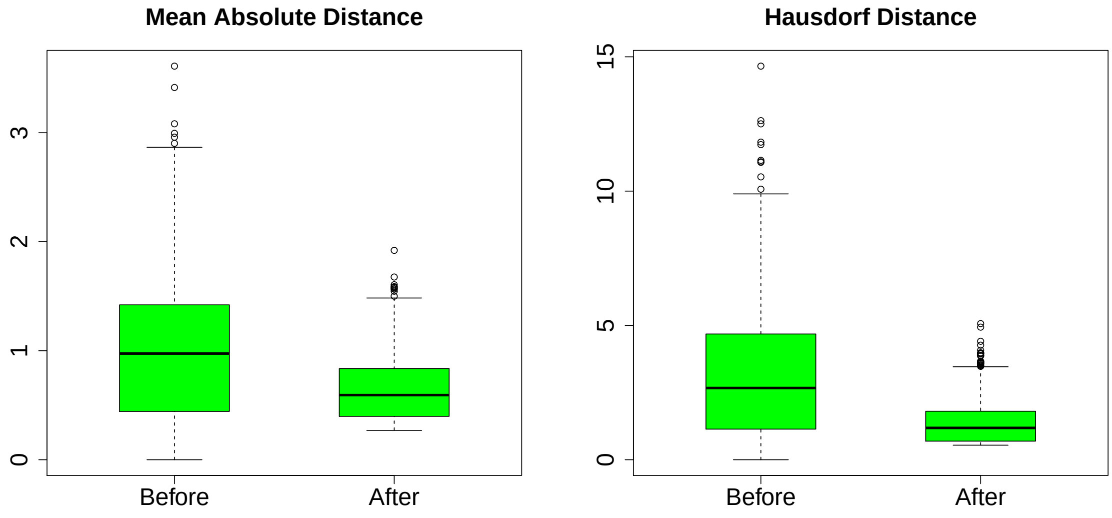


Figure 3.10: Illustration of observed 3D ROI instances with different shape parameter p. Each one is associated with a true ROI with the shape parameter w. For 2D slice view, see Fig.3.3. The color on the surfaces is for visualization of correspondence among surfaces and not for indicating the segmentation error.



(a) RO, higher is better

(b) DSC, higher is better



(c) MAD, higher is better

(d) HD, higher is better

Figure 3.11: Results of 3D synthetic surface assimilation, Case 1: the shapes of true ROI and segmentation vary in the same direction. Each plot shows the results before and after the surface assimilation.

Table 3.3: Results of 3D synthetic surface assimilation, Case 1: the shape variations of the true ROI and the segmentation are perpendicular

	RO		DSC		LDSC		MAD		HD	
	Before	After	Before	After	Before	After	Before	After	Before	After
Mean	0.879	0.927	0.933	0.962	3.086	3.382	0.997	0.654	3.181	1.366
Median	0.894	0.936	0.944	0.967	2.827	3.382	0.974	0.593	2.669	1.184
STD	0.092	0.041	0.055	0.022	1.230	0.581	0.651	0.299	2.499	0.794
IQR	0.130	0.055	0.073	0.029	1.481	0.934	0.975	0.439	3.537	1.113
Mean improvement	5.5% ($p < 0.001$)		3.1% ($p < 0.001$)		9.6% ($p < 0.001$)		34.5% ($p < 0.001$)		57.1% ($p < 0.001$)	

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance, STD: Standard Deviation, IQR: Interquartile Range

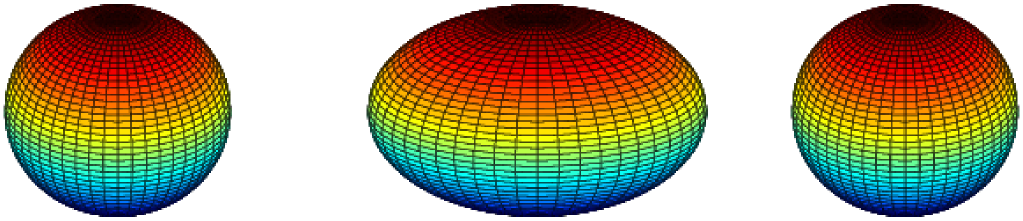
the algorithm of the surface assimilation. In Fig. 3.12 a *fat* segmentation is observed in (b). Based on that the statistical shape model (SSM) predicts a new surface in (c). This SSM-predicted surface is fused with the observed one using Eq. (3.8). The smoothing spline regression makes a final prediction of the true surface (e). Similar to the case of Fig. 3.4, the bumps in (d) are smoothed out. In Fig. 3.13, a *thin* segmentation is observed and the spline regression smoothes out the pointy edges at the equator.

To demonstrate the limitation of our method, we generated the observed surfaces by making them to change the shape vertically, as same as the shape model of the true ROI (Fig. 3.14). Mathematically,

$$y_1(u,v) = x_1(u,v)$$

$$y_2(u,v) = x_2(u,v)$$

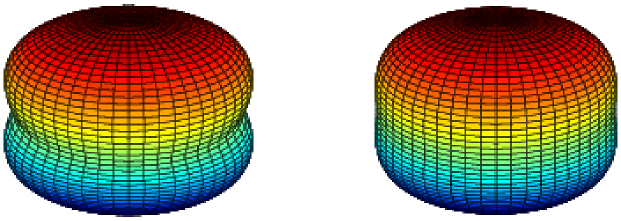
$$y_3(u,v) = x_3(u,v) + p\sin(u)$$



(a) True ROI

(b) Observed

(c) SSM predicted



(d) Combined

(e) Spline estimated

Figure 3.12: Demonstration of 3D synthetic surface algorithm, Case 1. (a) shows an instance of the true shape. (b) shows an observed surface with large outward segmentation error occurring at the left and right ends of the equator. (c) the statistical shape model (SSM) predicts a new segmentation. The surfaces of (b) and (c) are fused into one shown in (d). The estimated true shape by the smoothing spline regression (SSR) is shown in (e). The bumps in (d) are smoothed out by SSR. Colors on each surface are for visualization of the correspondence among surfaces not for indicating the segmentation error.

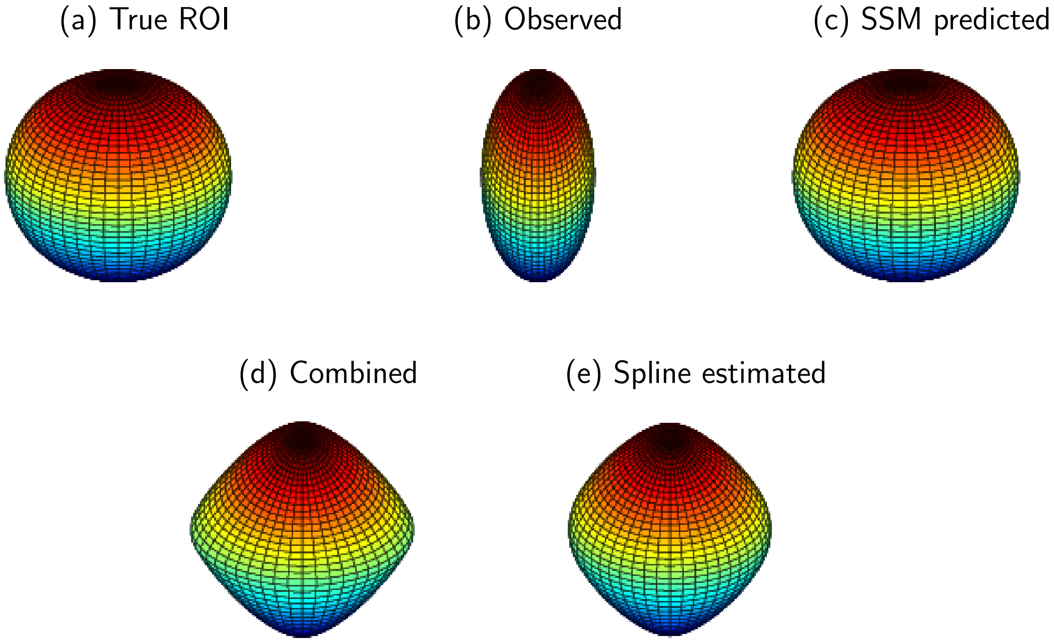
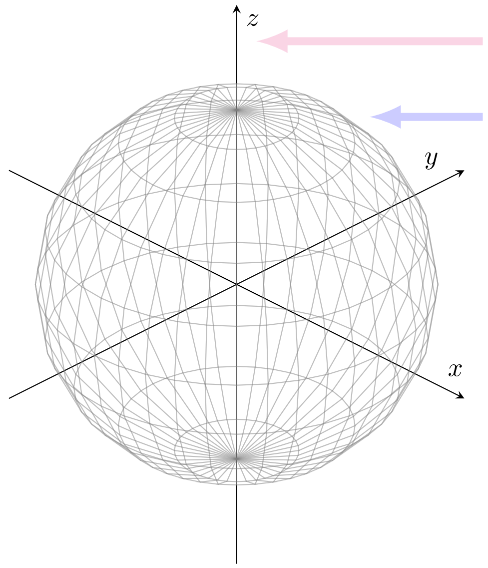


Figure 3.13: Demonstration of 3D synthetic surface algorithm, Case 2. (a) shows an instance of the true shape. (b) shows an observed surface with large inward segmentation error occurring at the left and right ends of the equator. (c) the statistical shape model (SSM) predicts a new segmentation. The surfaces of (b) and (c) are fused into one shown in (d). The estimated true shape by the smoothing spline regression (SSR) is shown in (e). The pointy boundary at the equator in (d) are smoothed out by SSR. Colors on each surface are for visualization of the correspondence among surfaces not for indicating the segmentation error.



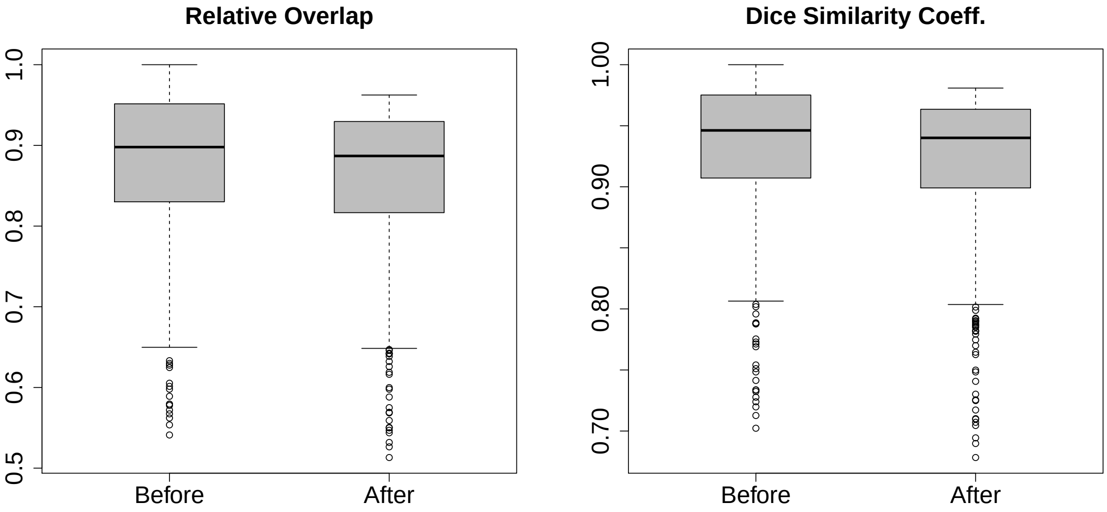
True shape symmetrically changes along the z axis

Observed shape changes along
the same direction as the true one

Figure 3.14: 3D synthetic surface assimilation, Case 2: The sphere at the center is the mean shape of the true ROI. The shape variations of the true ROI and the segmentation are in the same direction (along the z axis). For a 2D slice view, see Fig. 3.6.

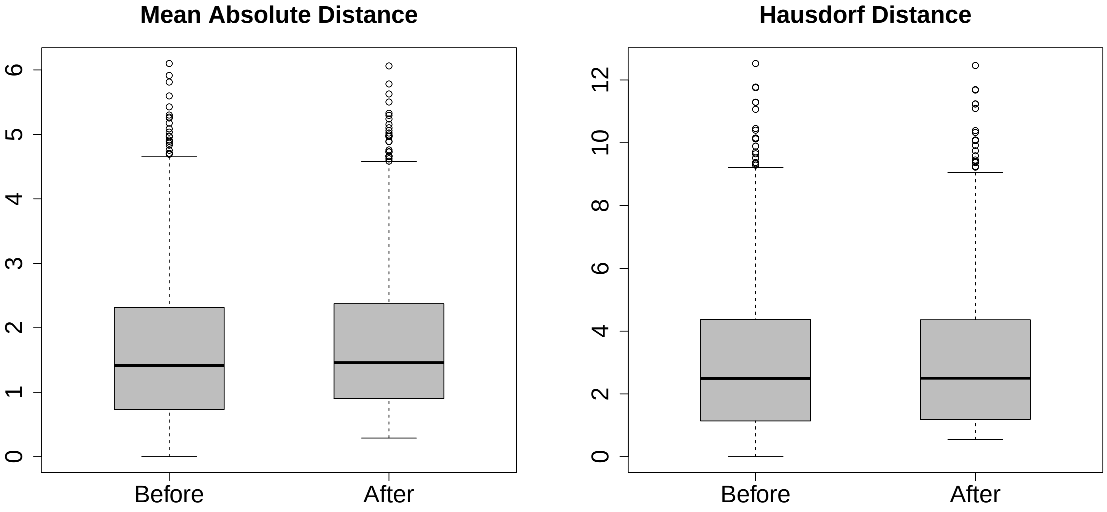
where $(x_1(u,v), x_2(u,v), x_3(u,v))$ serves the mean shape of the segmentation, $p \sim N(0, \sigma^2)$, $\sigma = 4$ controlling the shape of observed surface along the z-axis. The assimilation steps are similar to the experiment in 2D curve assimilation (Fig. 3.7). To avoid verbosity we skip the detailed description here. The results reported in Fig. 3.15 and Table 3.4 show no improvement after the surface assimilation. The results actually became worse. It is caused by the spline regression that automatically chose a slightly under-fit smoothing parameter λ which shrinks the fused surface toward the center. This limitation will be discussed in Sec. 3.7.

We designed a third experiment for 3D synthetic surface assimilation to show the algorithm strength. The true ROI $\mathbf{x}(u,v) = (x_1(u,v), x_2(u,v), x_3(u,v))$ has the sphere as the mean shape and two modes of shape variation along the x- and y- axis,



(a) RO, higher is better

(b) DSC, higher is better



(c) MAD, lower is better

(d) HD, lower is better

Figure 3.15: Results of 3D synthetic surface assimilation, Case 2: the shape variations of the true ROI and the segmentation are in the same direction

,

Table 3.4: Results of 3D synthetic surface assimilation, Case 2: the shape variations of the true ROI and the segmentation are in the same direction

	RO		DSC		LDSC		MAD		HD	
	Before	After	Before	After	Before	After	Before	After	Before	After
Mean	0.881	0.863	0.935	0.924	3.063	2.721	1.670	1.739	3.066	3.099
Median	0.898	0.887	0.946	0.940	2.865	2.753	1.415	1.460	2.493	2.498
STD	0.087	0.086	0.052	0.053	1.123	0.720	1.171	1.127	2.353	2.305
IQR	0.121	0.113	0.068	0.064	1.389	1.085	1.581	1.466	3.231	3.168
Mean improvement	-2.02% ($p < 0.001$)		-1.09% ($p < 0.001$)		-11.5% ($p < 0.001$)		-4.12% ($p = 0.180$)		-1.09% ($p = 0.748$)	

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance, STD: Standard Deviation, IQR: Interquartile Range

respectively (shown in Fig. 3.16). Mathematically,

$$
\begin{aligned}
x_1(u,v) &= r\cos(u)\cos(v) + w_1\cos(u)\cos(v) \\
x_2(u,v) &= r\cos(u)\sin(v) + w_2\cos(u)\sin(v) \\
x_3(u,v) &= r\sin(u)
\end{aligned}
$$

where $r = 24$ simulating the radius of a regular prostate, $u \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $v \in [0,\ 2\pi)$ and w_1, w_2 are normally distributed variables, i.e., $w_1 \sim N(0,\sigma^2)$, $w_2 \sim N(0,\sigma^2)$, $\sigma = 4$. The observed segmentation $\mathbf{y}$ is generated from

$$
\begin{aligned}
y_1(u,v) &= x_1(u,v) \\
y_2(u,v) &= x_2(u,v) \\
y_3(u,v) &= x_3(u,v) + p\sin(u)
\end{aligned}
$$

where $(x_1(u,v), x_2(u,v), x_3(u,v))$ serves the mean shape of the segmentation, $p \sim$

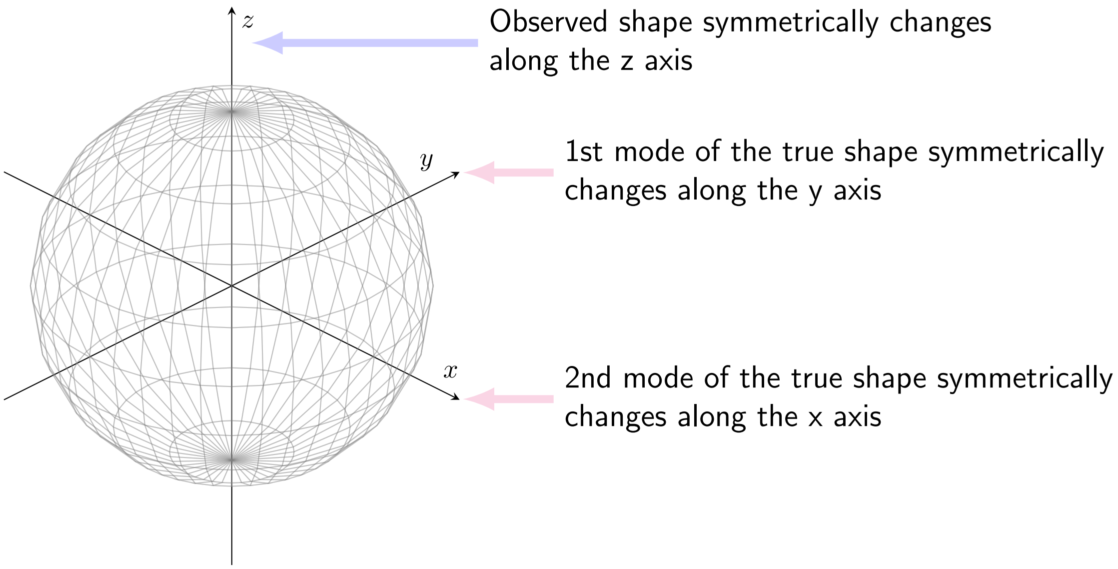
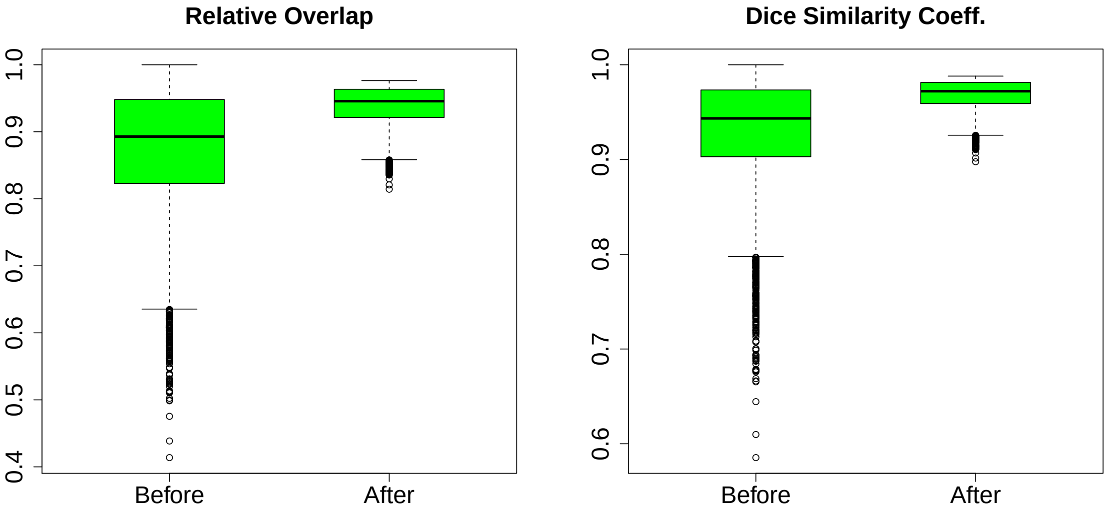


Figure 3.16: 3D synthetic surface assimilation, Case 3: the true ROI has two variation modes and both are perpendicular to that of the segmentation
,

$N(0, \sigma^2)$, $\sigma = 4$ controlling the shape of observed surface along the z-axis. The segmentation variation mode is perpendicular to those of the true ROI. Based on our theory of surface assimilation, our algorithm should output a better estimate of the true ROI. We simulated 8000 times and the results are reported in Fig. 3.17 and Table 3.5 showing a statistically significant improvement.

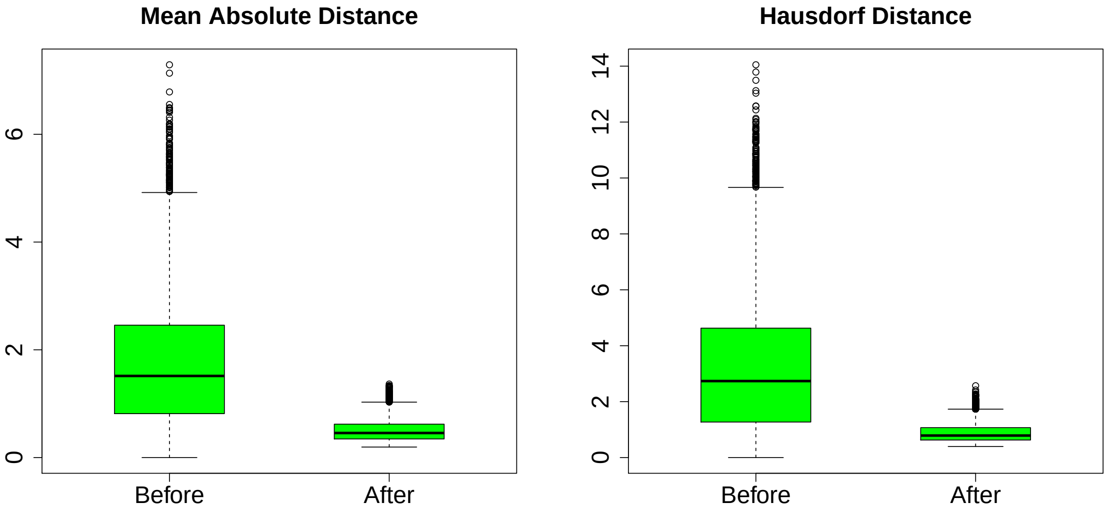
3.6.3 Experiment with Real Data

We have validated the surface assimilation algorithm with the NKI prostate database containing the FBCT pelvic images from 19 patients acquired daily through the treatment session (see [41] for detailed description). Table 2.1 provides a quick overview of the availability of the patient data. Based on the conclusion of Chapter 2, for each patient, we chose the first eight prostate surfaces as the training set for shape



(a) RO, higher is better

(b) DSC, higher is better



(c) MAD, lower is better

(d) HD, lower is better

Figure 3.17: Results of spline simulation experiment with the true shape having two variation modes

Table 3.5: Results of 3D synthetic surface assimilation, Case 3: the true ROI has two variation modes and both are perpendicular to that of the segmentation

	RO		DSC		LDSC		MAD		HD	
	Before	After	Before	After	Before	After	Before	After	Before	After
Mean	0.877	0.940	0.932	0.969	3.040	3.545	1.741	0.505	3.213	0.889
Median	0.893	0.946	0.943	0.972	2.814	3.550	1.515	0.456	2.737	0.789
STD	0.090	0.028	0.054	0.015	1.181	0.490	1.212	0.202	2.418	0.338
IQR	0.125	0.042	0.070	0.022	1.368	0.810	1.641	0.274	3.358	0.441
Mean improvement	7.2% ($p < 0.001$)		4.0% ($p < 0.001$)		16.6% ($p < 0.001$)		71.0% ($p < 0.001$)		72.3% ($p < 0.001$)	

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance, STD: Standard Deviation, IQR: Interquartile Range

model building and the remaining prostate surfaces as the truth in this validation experiment (Fig. 3.18).

Taking Patient A as an example, we describe the details of the validation experiment with Fig. 3.19. Patient A has 13 prostate segmentations from the FBCT scans. FBCT has better image resolution and grayscale contrast than CBCT. Therefore we assume the prostate surfaces on these scans are accurate. The first eight prostate surfaces are used to build the statistical shape model (SSM) of Patient A's prostate resulting the purple surface at the left of Fig. 3.19. The prostate surface of Scan 9 is used as the true for validation purpose.

To simulate the CBCT segmentations on the prostate of Scan 9, we use the inter-observed segmentation error (ISE) model obtained in Sec. 2.5. The mean shape of the ISE model is replaced by Scan 9 surface and 100 samples are generated from the model (the blue arrows in Fig. 3.19). These samples serve as the possible observed surfaces (gray) among the experts.

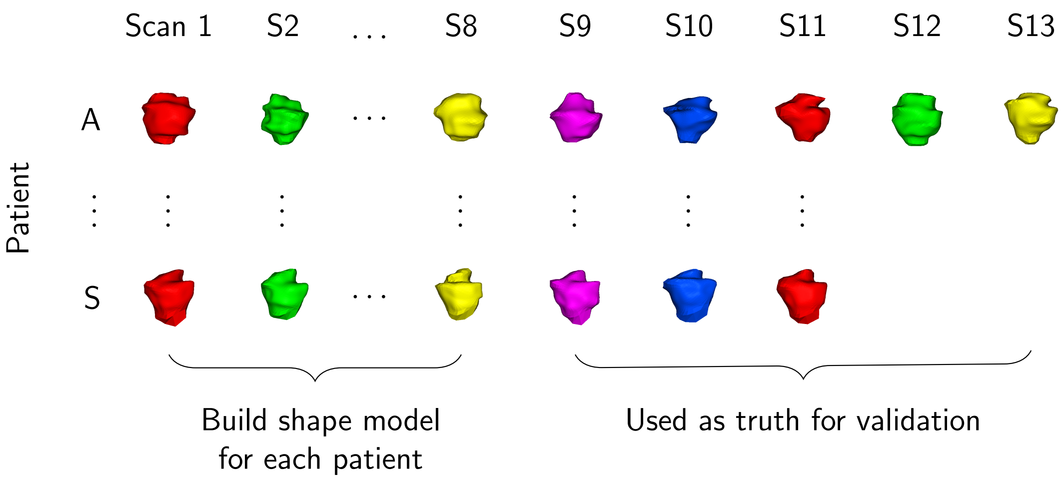


Figure 3.18: Illustration of the use of the real patient data in this experiment.

Following the red arrows from left to right in Fig. 3.19, for each observed surface, we obtain the predicted true (cyan) based on the knowledge contained in the SSM. This cyan surface is fused with the observed one using the Data Assimilation equation Eq. (3.8), resulting an intermediate surface (blue). We then apply the smoothing spline regression (SSR) to the blue surface to obtain the spline-estimated surface, the red one. Each red surface, as the final result of the surface assimilation algorithm, is compared with the true surface indicated by the dotted gray arrow. The observed surface is also compared with the true.

We chose Patient 3's CBCT ISE model in Sec. 2.5 to simulate the observed segmentation on NKI prostate surfaces. This ISE model shows a localized large segmentation error appears at the base of a prostate. This will lead to a bump or a dent in an observed surface. Accordingly, we provide two examples showing how our algorithm works on both cases. Fig. 3.20 and Fig. 3.21 illustrate the *bump* case with

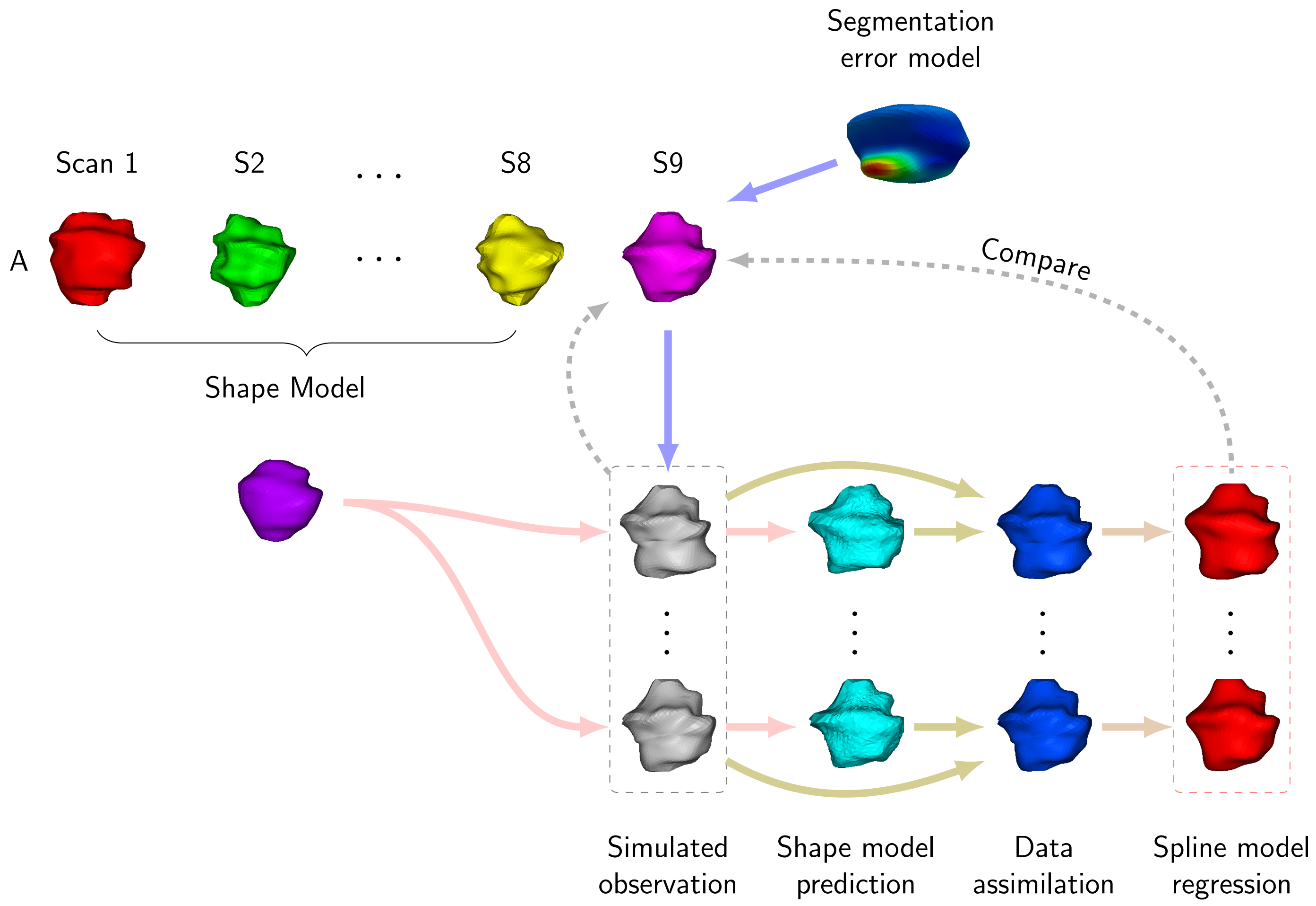


Figure 3.19: Data flow in the validation experiment on Patient A's Scan 9 prostate surface.

2D slice view and 3D view, respectively. Fig. 3.22 and Fig. 3.23 illustrate the *dent* case with 2D slice view and 3D view, respectively.

The experiment results are reported in Fig. 3.24, Fig. 3.25 and Table 3.7. Although we generated 100 observed surfaces for each patient at particular day, only part of them are topologically correct through the whole surface assimilation process (Table 3.6). The observed surface from the ISE model may not be a valid closed surface (topologically equivalent to a sphere) with the triangles on the surface cross over each other. At the data assimilation (fusion) phase, the topology of the surface can not be always guaranteed. Despite this limitation, the validation results are promising.

Table 3.6: Numbers of valid surface assimilation cases in the NKI experiment. Patient C, K, H, O and P are not included

Patient	Scan 9	Scan 10	Scan 11	Scan 12	Scan 13	Scan 14
A	53	51	36	54	38	-
B	52	45	-	-	-	-
D	57	53	67	-	-	-
E	76	84	62	84	64	-
F	62	-	-	-	-	-
G	73	68	67	72	-	-
I	67	69	71	70	-	-
J	84	86	82	-	-	-
L	68	72	61	45	60	61
M	62	48	46	-	-	-
N	58	48	62	-	-	-
Q	66	72	69	44	63	-
R	76	83	68	71	-	-
S	75	68	77	-	-	-

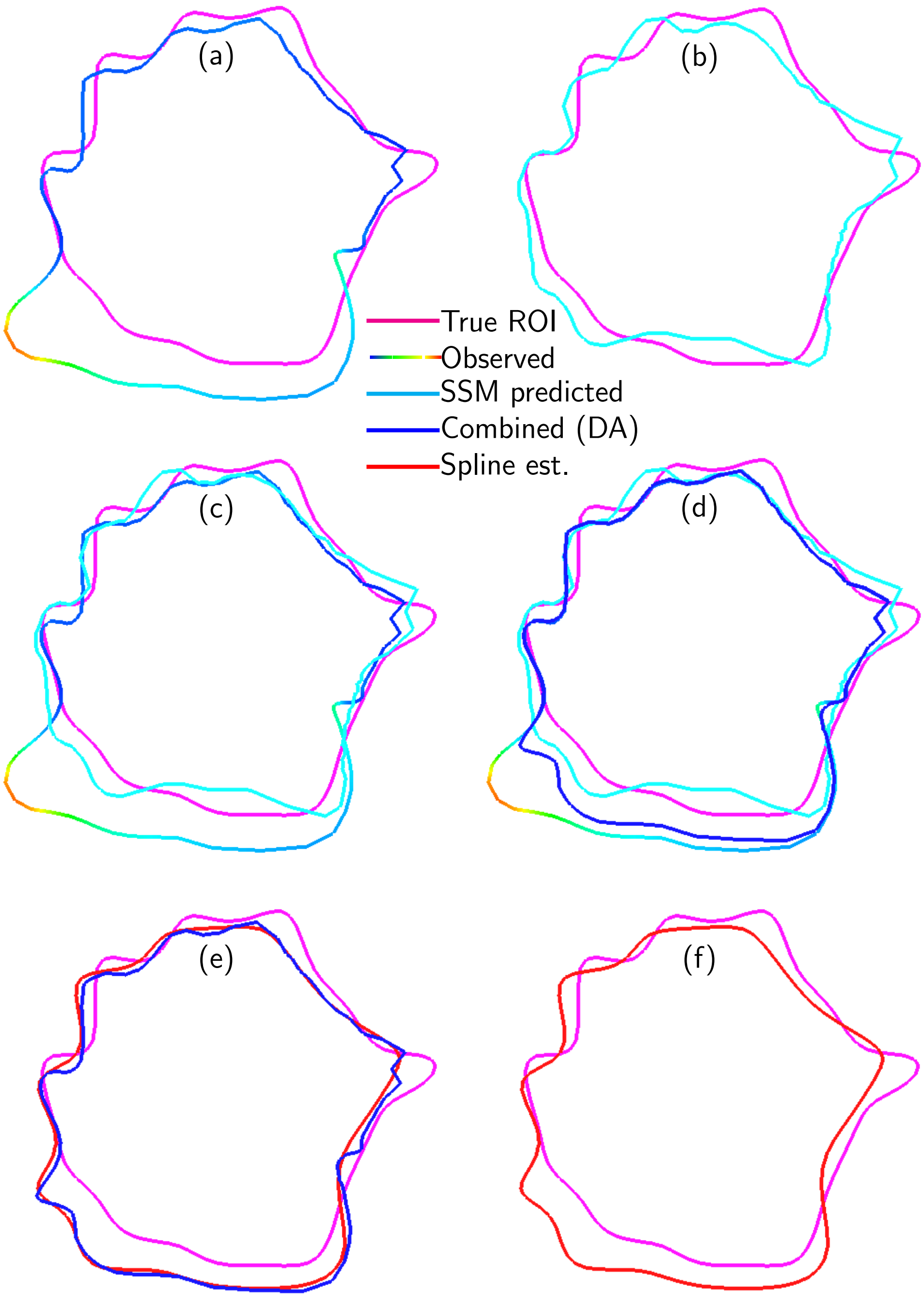


Figure 3.20: 2D slice view of surface assimilation on real patient data, *bump* case. The colors on the observed curve indicate the segmentation error. Large segmentation error at the base of the prostate forming a bump. The shape model predicts the truth (cyan). Blue curve in (d) is the result of Data Assimilation combining the observed and SSM-predicted curve. The red curve in (e) and (f) is the result of spline regression.

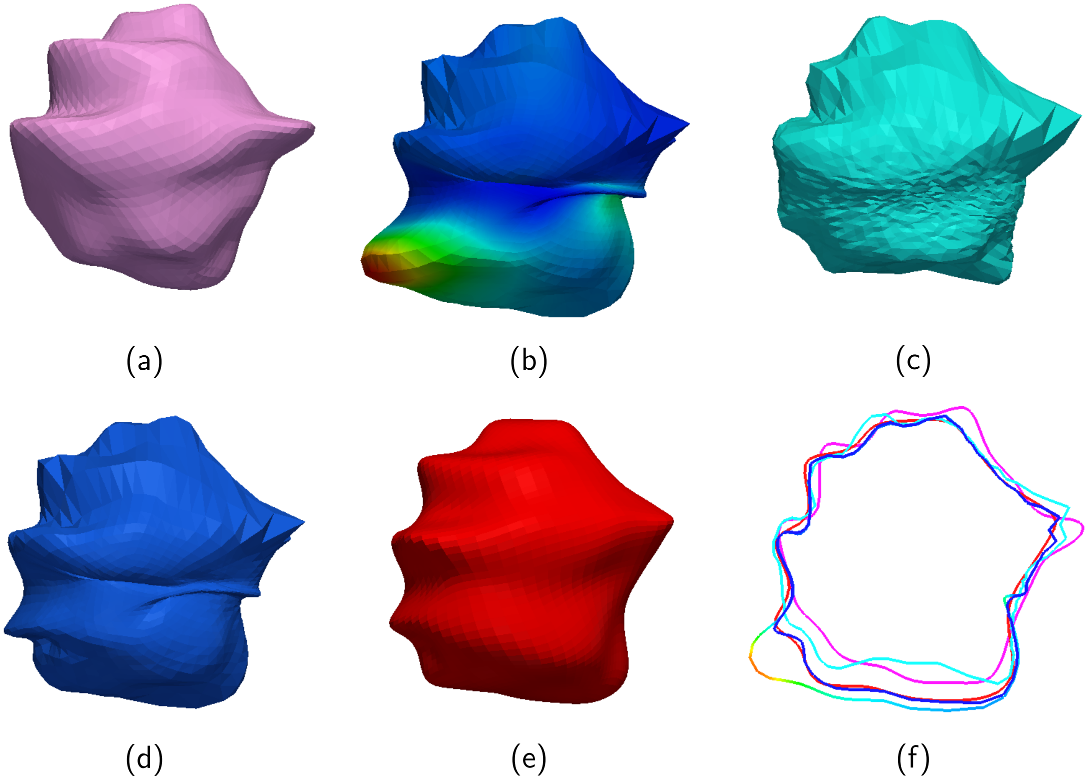


Figure 3.21: 3D view of surface assimilation on real patient data, *bump* case. From (a) to (f): (a) the true ROI, (b) the observed surface. The colors on the observed surface indicate the segmentation error. Large segmentation error at the base of the prostate forming a bump. (c) the prediction made by the statistical shape model (SSM). (d) the result of Data Assimilation combining the observed and SSM-predicted surfaces. (e) the result of smoothing spline regression. (f) 2D slice view of the surfaces from (a) to (e)

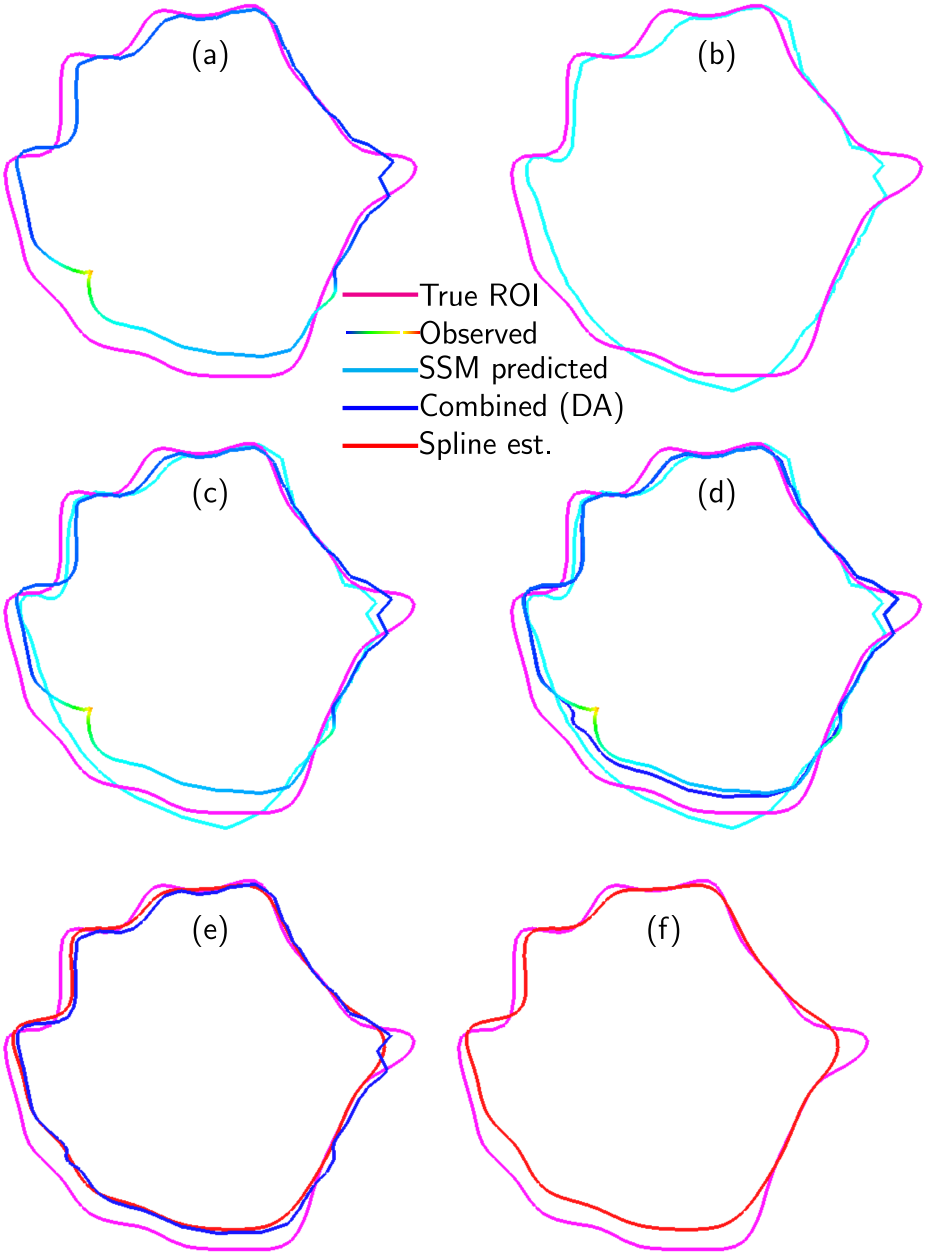


Figure 3.22: 2D slice view of surface assimilation on real patient data, *dent* case. The colors on the observed curve indicate the segmentation error. Large segmentation error at the base of the prostate forming a dent. The shape model predicts the truth (cyan). Blue curve in (d) is the result of Data Assimilation combining the observed and SSM-predicted curve. The red curve in (e) and (f) is the result of spline regression.

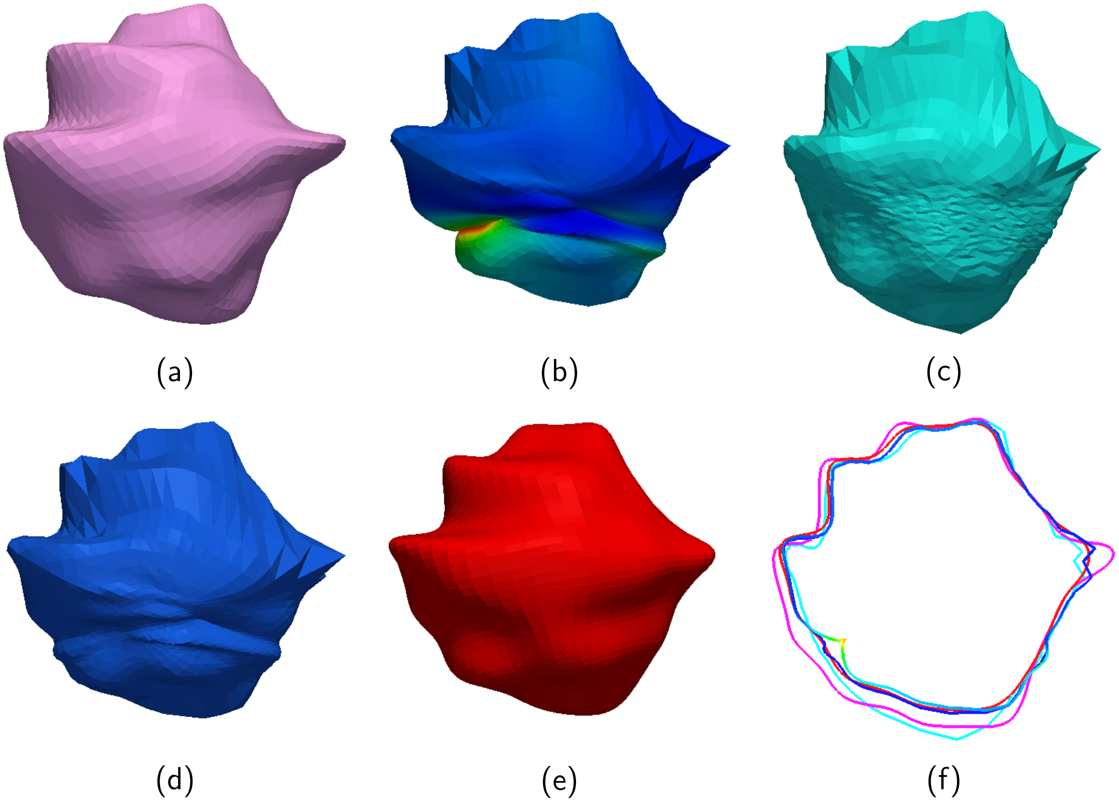
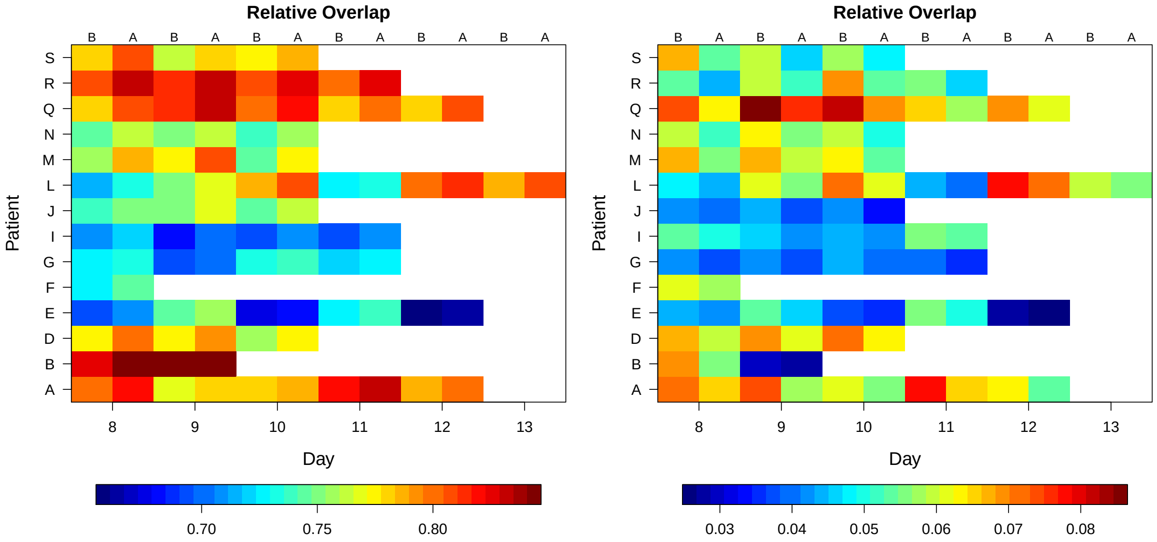
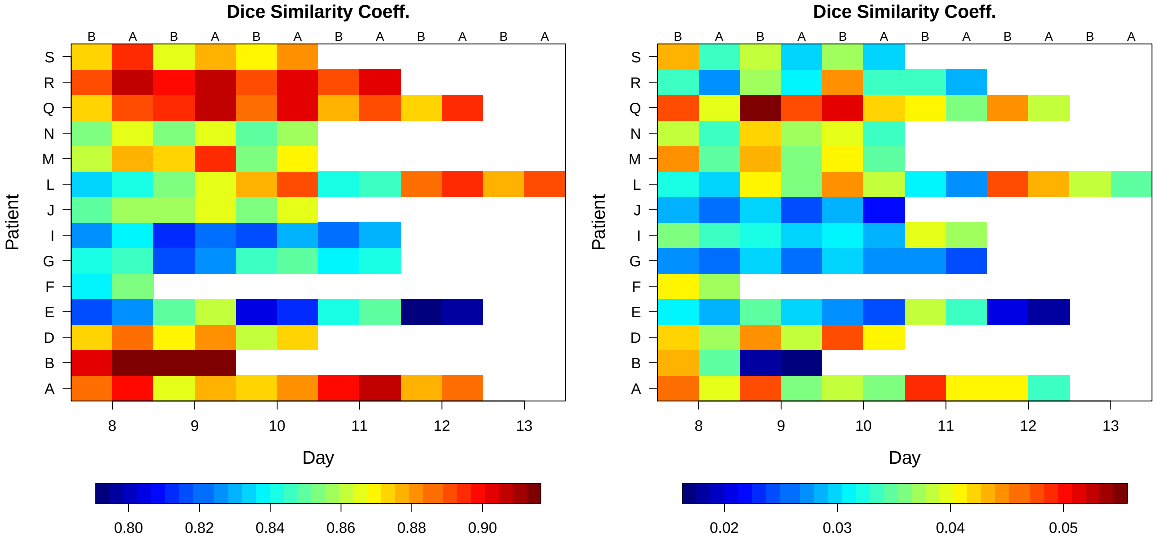


Figure 3.23: 3D view of surface assimilation on real patient data, *dent* case. From (a) to (f): (a) the true ROI, (b) the observed surface. The colors on the observed surface indicate the segmentation error. Large segmentation error at the base of the prostate forming a dent. (c) the prediction made by the statistical shape model (SSM). (d) the result of Data Assimilation combining the observed and SSM-predicted surfaces. (e) the result of smoothing spline regression. (f) 2D slice view of the surfaces from (a) to (e)

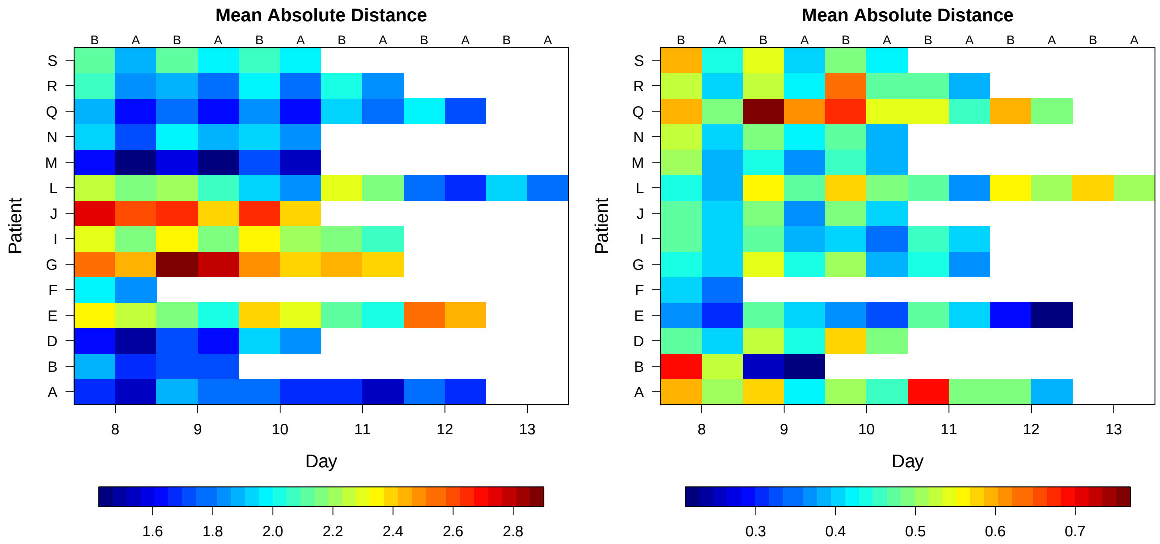


(a) RO, mean, higher is better (b) RO, stddev, lower is better



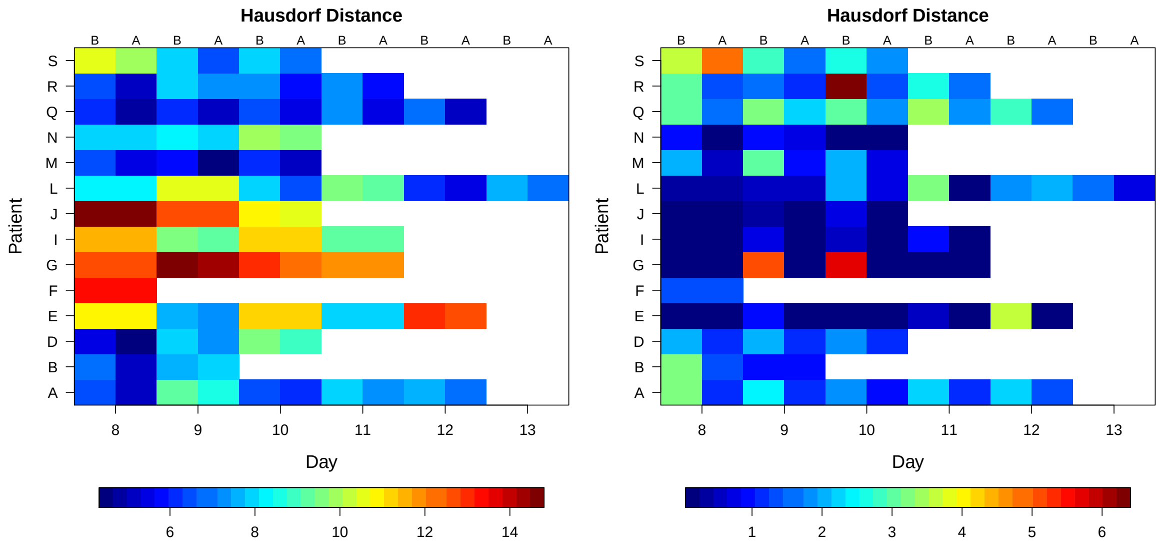
(c) DSC, mean, higher is better (d) DSC, stddev, lower is better

Figure 3.24: Results of prostate surface assimilation experiment on NKI data, RO and DSC. For each subfigure, the letters B and A under the title are for Before and After, respectively



(a) MAD, mean, lower is better

(b) MAD, stddev, lower is better



(c) HD, mean, lower is better

(d) HD, stddev, lower is better

Figure 3.25: Results of prostate surface assimilation experiment on NKI data, MAD and HD. For each subfigure, the letters B and A under the title are for Before and After, respectively

Table 3.7: Improvement in percentage after surface assimilation on NKI prostate data

Patient	RO		DSC		LDSC		MAD		HD	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
A	1.7	14.7	1.0	16.3	3.7	11.1	5.9	20.1	10.8	50.6
B	1.5	13.4	0.8	14.9	3.6	8.1	6.0	16.7	12.1	28.4
D	2.3	12.3	1.4	14.3	5.2	7.3	7.8	16.7	10.6	38.3
E	1.6	10.7	1.0	11.9	3.5	9.0	4.8	17.2	2.2	52.5
F	2.9	7.5	1.7	9.6	6.3	2.8	7.1	13.8	0.2	2.5
G	1.2	8.1	0.7	9.0	2.5	6.7	4.0	17.2	2.6	68.7
I	2.0	7.3	1.2	8.5	4.4	5.4	6.2	14.9	1.6	68.2
J	2.2	14.7	1.3	16.0	4.9	11.9	7.4	19.8	0.8	47.6
L	1.9	9.4	1.1	10.7	4.2	6.4	5.7	13.6	5.4	31.8
M	3.9	15.6	2.2	18.2	8.8	8.3	11.7	18.8	19.8	65.5
N	2.2	12.3	1.3	14.2	4.8	8.5	7.2	18.5	2.3	59.0
Q	2.9	13.0	1.7	15.4	6.8	6.0	9.6	18.4	22.1	42.0
R	2.5	17.8	1.4	20.0	5.9	11.4	8.3	21.4	16.3	51.7
S	2.7	20.9	1.6	23.0	5.9	16.0	7.2	23.0	12.3	14.8
Mean	2.3	12.7	1.3	14.4	5.0	8.5	7.1	17.9	8.5	44.4

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance, STD: Standard Deviation, IQR: Interquartile Range

3.7 Discussion

Splines are extensively used in surface fitting problem. To deal with heteroscedastic data (variance is pointwisely unequal), the ideas like robust fitting [109], weighted splines [32] and adaptive splines [88, 110] have been proposed. Besides the smoothing splines defined on the regular grid domains, the spherical splines are our major interest since they work on the closed surfaces. The extension of spherical splines include fitting scattered data on sphere-like surface [2], spherical interpolation and approximation [58] and the vector spherical splines [59].

These regression methods are data driven without resorting to the prior knowledge. When the noise is independent and uncorrelated with signal, the trend in the observational data provides enough information for these methods to recover the signal. However in many applications (like our organ segmentation problem), the noise is highly correlated with the signal and the true trend in the signal is disguised. If the history (prior knowledge) of the signal is available, it can help to separate the signal and noise. This is the basic idea of surface assimilation.

Shape priors have been used extensively in image segmentation algorithms [27]. The contribution of this study is to combine two priors, shape and segmentation error, together. As a variant of Data Assimilation, the surface assimilation can be extended in multiple ways. The 3D variance analysis is one of the themes in Data Assimilation. Due to limited resources, we only describes a crude estimation to determine the variance and correlation in the segmentation errors (see Sec. 2.5). The errors are assumed isotropic and Gaussian. Analysis on the variance-covariance structure remains interest [134] for future study.

Surface assimilation can be extended to time-dependent data with 4D variance analysis. Wang [134] generalized the smoothing spline regression to ANOVA and data can be smoothed along the time dimension. In our clinical application setting, patients are scanned consecutive days. Therefore, the time-dependent surface assimilation seems like another possible extension to this work.

Both the theory and experiments suggest that the performance of our algorithm mainly depends on the orthogonality of the true object shape space and the segmentation error space. Ideally when the segmentation noise is perpendicular to the true object shape, the proposed algorithm has the potential to fully recover the true object boundary. However the limitation is that this algorithm is insensitive to the translation, rotation and scaling difference since the Procrustes mapping in Eq. (3.6) discards these global shape changes.

The segmentation uncertainty in this study is assumed fixed and unchanged across the images (either intra-subject or inter-subject). This segmentation error model requires training process and probably is not valid in a wider scenario. A possible extension could be to determine the segmentation error from the local intensity context. For instance, if the local intensity is homogeneous, it is likely the manual contours across this region has large variation comparing those passing through the region with rich intensity changes. Various feature extraction techniques could be applied here, such as the Gabor filter used in [19] to detect the spherical boundary, or Hessian image filter [57] for detecting the sheet-like structure.

The choice of the building blocks, splines, can also be an interesting problem. For example, the splines with finite support, like B-splines, are used in [44] to match historical data. In robust spline regression, the splines with infinite support, like Thin-plate splines, are preferred since the negative effects caused by the missing data can be compensated from the other data [60].

We mainly studied the surface assimilation for single object. With the development in statistical shape modeling and segmentation error analysis for multiple objects, the surface assimilation for multiple objects will be a possible research direction for future.

We have seen that surface folding, triangle crossing, etc. happens during the fusion of the two surfaces. Therefore maintaining valid surface geometry and improving the robustness of the surface assimilation algorithm is an important future direction of this work.

3.8 Summary

We proposed a surface assimilation algorithm for predicting the true object boundary in the presence of the segmentation errors. Formulated as a smoothing spline regression problem, the surface assimilation combines the knowledge about shape and segmentation errors. Benchmark experiment was performed on phantom and real data with different configurations. Although cautions should be taken from the few examples presented, the results seem promising.

CHAPTER 4
NONRIGID IMAGE REGISTRATION WITH ASSIMILATED SURFACE CONSTRAINT

This chapter presents an application of the surface assimilation (SA) technique related to the nonrigid image registration. Instead of using the manual contoured surface, we use the surface estimated by SA (see Chapter 3) to constrain the intensity-driven nonrigid image registration to increase the registration accuracy. The proposed registration algorithm is evaluated with phantom and real clinical data.

4.1 Introduction

As we mentioned in Sec. 1.1.2, the low contrast of intensity at the object boundary in CBCT contributes to the image registration uncertainty. To reduce this uncertainty, one can identify features (e.g., point landmarks, curves and segmentations) in the images to provide additional matching guidance [77, 21, 78]. However, the uncertainty among different observers often exists in these identified features. In case of contouring the prostate boundary, studies have shown that such errors vary dramatically (from 1-4 mm) with location on the prostate surface and adjacent organs at risk [104]. If these label features are directly used as the matching constraint, large registration uncertainty can occur.

The identification errors in point landmark registration have been actively studied [51, 50]. These points are generally sparse and the errors are therefore assumed independent. A common approach to lower these errors in the registration is to weigh the landmark matching force with their estimated landmark error [108, 137].

Limited studies consider the inter-observer segmentation error (ISE) in image registration. We proposed a surface constraint that accommodates the segmentation uncertainty [141]. The experiment in [141] showed that the deformation from the intensity matching can not always recover the correct deformation at the object's boundary. Therefore, registration accuracy is not always improved.

To overcome this limitation, we propose a new surface constraint to reduce the image registration uncertainty in the presence of segmentation error. With the help of surface assimilation, we estimate the true object boundary from the manual (observed) segmentation (Sec. 3.3). The estimated surface is then used to guide the intensity-based image registration (Sec. 4.2).

4.2 Methods

Suppose the template and the target images are denoted as $T : \Omega \mapsto \mathbb{R}$ and $S : \Omega \mapsto \mathbb{R}$, respectively, where $\Omega = [0,1]^3$ is the image domain. Let $X : \mathbb{S}^2 \mapsto \Omega$, as a parametric surface, denote the true ROI in the template space. Let $Y_g : \mathbb{S}^2 \mapsto \Omega$ and $Y : \mathbb{S}^2 \mapsto \Omega$ denote the parametric surfaces of the true and observed ROI in the target spaces, respectively, and let $F : \mathbb{S}^2 \mapsto \Omega$ be the assimilated surface of Y, where $\mathbb{S}^2$ is the unit sphere. The image registration problem can be stated as: Find a dense transformation $h : \Omega \mapsto \Omega$ that maps points in the target image to the corresponding points in the template image.

4.2.1 Design of cost function

The cost function provides the driving force to deform the coordinate system of the target image into that of the template image. A cost function usually contains similarity metrics and transformation constraints. Based on the availability of the matching features, various similarity metrics can be used. When registering two grayscale images, we can use intensity difference, mutual information or correlation coefficient as a proxy for correspondence. If other matching features are available, such as landmarks, curves and surfaces, we can design additional similarity cost terms to define correspondence between the. Prior knowledge about the transformation, such as smoothness, are assumed and included in the cost function. Particularly, our algorithm uses the following metrics.

Sum of Squared Difference (SSD)

A common intensity similarity cost is the sum of squared difference (SSD), which measures the intensity difference at corresponding position between two images. Mathematically, it is defined by

$$C_{\mathrm{SSD}}(h) = \int_{\Omega} \left[T(h(x)) - S(x)\right]^2 \,\mathrm{d}x \tag{4.1}$$

where $h(x)$ is the transformation field, $T(x)$ and $S(x)$ are the template and target image, respectively. Ω denotes the image domain. When the template and target images are from the same imaging modality and the intensity profiles of two images are similar, C_{SSD} will serve as a good surrogate indicating the correspondence. C_{SSD}

will be used in our phantom experiment.

Mutual information (MI)

If the template and target images are from two different imaging modalities and the intensities between the two are not linearly related, then C_{SSD} is not a good choice for measuring similarity. We can instead use the negative mutual information (MI) [136] between two images. MI expresses the amount of information that two image share and is defined as

$$C_{\mathrm{MI}}(h) = -\sum_{a}\sum_{b} p(a,b) \log \frac{p(a,b)}{p_{T\circ h}(a)\, p_S(b)} \tag{4.2}$$

where $T \circ h$ denotes the deformed template image, a and b are the indices of the joint histogram bins. Similar to the Kullback-Leibler distance, $C_{\mathrm{MI}}(h)$ measures the distance between the joint intensity distribution of deformed template and target image $p(a,b)$, and the joint marginal distribution $p_{T\circ h}(a)$ and $p_S(b)$. It is assumed that good correspondence is achieved when MI is maximized. In practice, SSD and MI are not mutually exclusive. Under the multi-stage strategy, Han[68] and Lou[86] combined SSD and MI for faster boundary matching at initial stage in lung and prostate image registration.

Surface Similarity Metric

Analogous to landmark registration, the distance between the template and target surfaces can help registering two images. Assuming the correspondence between surfaces has been obtained using the surface registration technique discussed in Chapter 2, we define the equally-weighted assimilated-surface constraint (EWAS) as:

$$C_{\mathrm{EWAS}}(h) = \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \| h(F(u,v)) - X(u,v) \| |\sin u| \, \mathrm{d}u \, \mathrm{d}v \tag{4.3}$$

where u, v establish the parametric mesh on the surfaces of X and F, and $F(u,v)$ is the assimilated surface of the given manual surface $Y(u,v)$. A straightforward variant of C_{EWAS} is the equally-weighted manual surface constraint (EWMS):

$$C_{\mathrm{EWMS}}(h) = \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \| h(Y(u,v)) - X(u,v) \| |\sin u| \, \mathrm{d}u \, \mathrm{d}v \tag{4.4}$$

Similar to the landmark uncertainty in the landmark cost [107], we can define an uncertainty-weighted manual surface constraint (UWMS) as:

$$C_{\mathrm{UWMS}}(h) = \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{\omega(u,v)} \| h(Y(u,v)) - X(u,v) \| |\sin u| \, \mathrm{d}u \, \mathrm{d}v \tag{4.5}$$

where $\omega(u,v)$ is given in Eq. (2.27) and indicates how much agreement there is on the boundary location among the observers. It is a priori knowledge carried in the inter-observer segmentation error (ISE) model. Similarly, we have the uncertainty-weighted assimilated surface constraint (UWAS)

$$C_{\mathrm{UWAS}}(h) = \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{\omega(u,v)} \| h(Y(u,v)) - X(u,v) \| |\sin u| \, \mathrm{d}u \, \mathrm{d}v \tag{4.6}$$

The rationale is that $\frac{1}{\omega(u,v)}$ preferentially decreases the influence of those surface patches that are less accurately known. The cost function in Eq. (4.5) acts like a distance function (similar to that used for ICP) to generate a potential or gradient to draw the surfaces together. The distance function associated with a point on the surface is adapted to become flat (i.e., stop attracting) over a range corresponding to the uncertainty in the surface contour at that point. In regions where the surface is known with high precision, the distance function will attract to within a small distance of the surface; in regions that are less precisely known, the potential will stop attracting at a greater distance.

Laplacian Regularization (LAP)

As a common transformation smoothness constraint, the Laplacian operator is used to regularize the displacement fields $u(x)$ where $u(x) = h(x) - x$. This regularization term is formed as

$$C_{\mathrm{LAP}}(h) \quad = \quad \int_{\Omega} \|\nabla^2 u(x)\|^2 \, \mathrm{d}x \tag{4.7}$$

where $\nabla = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right]$ and $\nabla^2 = \nabla \cdot \nabla = \left[\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right]$.

The total cost function is a linear combination of the intensity metric ($C_{\mathrm{INTENSITY}}$), the surface metric (C_{SURFACE}) and the regularization penalty (C_{REG}):

$$C_{\mathrm{TOTAL}}(h) = C_{\mathrm{INTENSITY}} + \rho C_{\mathrm{SURFACE}} + \alpha C_{\mathrm{REG}} \tag{4.8}$$

where the balancing parameters ρ and α control the significance of the different terms,

$C_{\text{INTENSITY}}$ can be either C_{SSD} or C_{MI}. C_{SURFACE} includes C_{EWMS}, C_{UWMS}, C_{EWAS} and C_{UWAS}, C_{REG} is C_{LAP} in this work. We therefore have a suite of registration algorithms derived from Eq. (4.8) by changing these weighting factors and the cost terms. The evaluation experiment of those algorithms is reported in Sec. 4.3.

4.2.2 Transformation and Optimization

Based on the different requirements of registration problems, the transformation $h(x)$ can be represented using various forms of basis function. As a widely used fitting basis in engineering problems, we chose cubic B-splines to represent the nonrigid transformation $h(x)$. A B-splines based transformation is generally computationally efficient and its local support facilitate the capture of local deformation.

Let $\mathbf{c}_i = [c_x(\mathbf{x}_i), c_y(\mathbf{x}_i), c_z(\mathbf{x}_i)]^T$ be the coefficients of the i-th control point $\mathbf{x}_i$ on the spline lattice G along each direction. $h(x)$ can be represented as

$$\mathbf{h}(\mathbf{x}) = \mathbf{x} + \sum_{i \in G} \mathbf{c}_i \beta^{(3)}(\mathbf{x} - \mathbf{x}_i), \tag{4.9}$$

where $\beta^{(3)}(\mathbf{x}) = \beta^{(3)}(x)\beta^{(3)}(y)\beta^{(3)}(z)$ is a separable convolution kernel [15]. $\beta^{(3)}(x)$ is the uniform cubic b-spline basis function defined as

$$\beta^{(3)}(x) = \begin{cases} \frac{1}{6}(4 - 6x^2 + 3|x|^3), & 0 \leq |x| < 1 \\ \frac{1}{6}(2 - |x|)^3, & 1 \leq |x| < 2 \\ 0, & |x| \geq 2 \end{cases} \tag{4.10}$$

With the b-spline representation, solving for $h(x)$ by minimizing the cost function Eq. (4.8) is equivalent to solving for the set of control coefficients $\mathbf{c}_i$. Among many

Table 4.1: Image registration algorithms in the evaluation experiment

Algorithm	Cost function	Transformation	Optimization
SSD	$C_{\mathrm{SSD}} + \alpha C_{\mathrm{LAP}}$	B-splines	Gradient descent
EWMS	$C_{\mathrm{SSD}} + \rho C_{\mathrm{EWMS}} + \alpha C_{\mathrm{LAP}}$	B-splines	Gradient descent
UWMS	$C_{\mathrm{SSD}} + \rho_1 C_{\mathrm{UWMS}} + \alpha C_{\mathrm{LAP}}$	B-splines	Gradient descent
EWAS	$C_{\mathrm{SSD}} + \rho C_{\mathrm{EWAS}} + \alpha C_{\mathrm{LAP}}$	B-splines	Gradient descent

SSD: Intensity-only registration. EWMS: Equally-weighted manual surface constrained image registration. UWMS: Uncertainty-weighted manual surface constrained image registration. EWAS: Equally-weighted assimilated surface constrained image registration.

optimization methods we chose a standard Gradient Descent method for the phantom experiment, a limited-memory, quasi-Newton minimization with bounds (L-BFGS-B) [145] for the real data experiment. A multi-stage multi-resolution strategy is used to speed up the convergence and to avoid local minimum.

4.3 Experiments

Evaluation of image registration is an important step in algorithm design. We evaluated the proposed algorithm on phantom and real patient data sets with three algorithms: 1) intensity-only registration; 2) intensity registration with equally-weighted manual surface constraint; and 3) intensity registration with uncertainty-weighted manual surface. These algorithms are listed in Table 4.1. For a fair comparison, the balancing parameter ρ_1 in UWMS and ρ in other algorithms satisfies the relation:

$$\rho_1 \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{\omega(u,v)} |\sin u| \, \mathrm{d}u \, \mathrm{d}v = \rho \tag{4.11}$$

Due to the absence of a “gold standard” to validate a registration algorithm, indirect evaluation methods are needed to quantify the different aspects of the performance of the algorithms. As same as the evaluation metrics used in Chapter 3, our evaluation metrics include Relative Overlap (RO), Dice Similarity Coefficient (DSC), Logit Dice Similarity Coefficient (LDSC), Mean Absolute Distance (MAD) and Hausdorff Distance (HD) (see Sec. 3.6 for details).

In our experiment the true surface locations were known and therefore an objective evaluation based on the target registration error (TRE) was computed [51]. The TRE is defined as the geometric discrepancy between two true surfaces after the registration. Assuming the template surface $X(u,v)$ is accurate, we characterize the TRE of an algorithm by reporting the RO, DSC, LDSC, MAD and HD between $X(u,v)$ and the deformed true target surface $h(Y_g(u,v))$.

4.3.1 Digital phantom experiment

In this experiment all phantom surfaces were positioned at the center of a $128 \times 128 \times 128$ mm^3 coordinate system. These surfaces were generated following the same strategy in Sec. 3.6.2. In the template space, a ellipsoid surface with Gaussian distributed z-axis is chosen to simulate the possible shape changes of $X(u,v)$:

$$X(u,v) = (a\cos u \cos v,\ a\cos u \sin v,\ b \sin u) \tag{4.12}$$

where $u \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $v \in [0,\ 2\pi)$, $a = 24$ mm to simulate the radius of a regular prostate, b is a Gaussian random variable with 24 mm mean and 4 mm standard

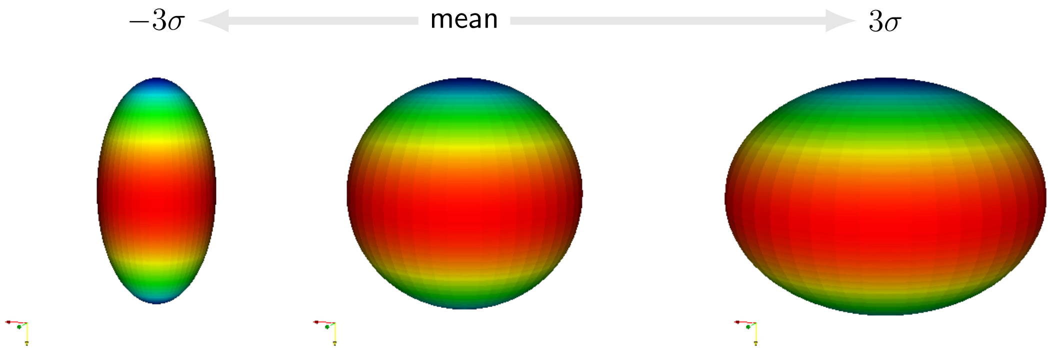


Figure 4.1: Fixing the height of the ellipsoid, the shape of $Y(u, v)$ changes along y-axis (horizontally). The colors on the surface indicate the magnitude and distribution of the segmentation error. The largest error happens at the equator, the smallest error at the poles.

deviation. The mean shape of $X(u, v)$ is a sphere with 24 mm radius. And its shape only changes along the z-axis.

To simulate the manual surface with inter-observer segmentation error, we first obtain a true surface $Y_g(u, v)$ by sampling from Eq. (4.12) and then generate $Y(u, v)$ as

$$Y(u, v) = (Y_{g1}(u, v),\, Y_{g2}(u, v) + w \cos u \sin v,\, Y_{g3}(u, v))$$

where w is a Gaussian random variable with zero mean and 4 mm standard deviation (see Fig. 3.10 for example). For a fixed $Y_g(u, v)$, the shape of $Y(u, v)$ changes only along the y-axis. The largest segmentation error appears at the equator and is illustrated in Sec. 4.3.1. The surface assimilation algorithm is applied to each instance of $Y(u, v)$ to obtain $F(u, v)$.

All surfaces are surrounded by an intensity object. In the template space,

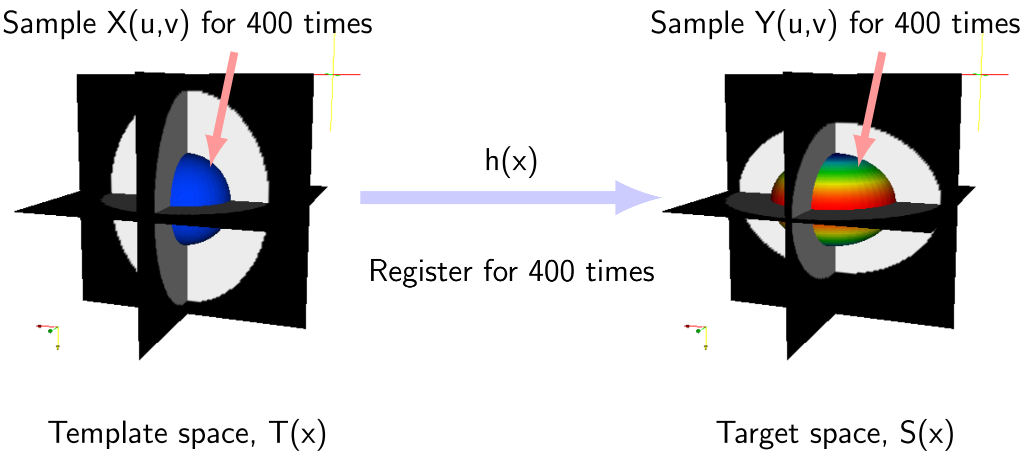


Figure 4.2: Illustration of the 3D phantom experiment. The colors on the ellipsoid surface indicate the segmentation uncertainty. The segmentation of $X(u, v)$ (blue surface) was assumed as the true surface.

$T(x)$ has a vertically elongated ellipsoid rasterized from

$$T(u, v) = (40 \cos u \cos v,\ 40 \cos u \sin v,\ 52 \sin u)$$

In the target space, $S(x)$ has a horizontally elongated ellipsoid rasterized from

$$S(u, v) = (40 \cos u \cos v,\ 52 \cos u \sin v,\ 40 \sin u)$$

where we chose 40 mm and 52 mm as the lengths of axes to guarantee no surface sample touches the boundary of the surrounding image object.

The surfaces were reused from the surface assimilation experiment in Sec. 3.6.2 where 1000 cases of the true, the observed and the assimilated surfaces were generated. We used the observed and assimilated surfaces in the first 400 cases as $Y(u, v)$ and

$F(u, v)$, respectively. The true surfaces in the next 400 cases were used as $X(u, v)$. Totally 400 pairs of images and surfaces are registered using the algorithms SSD, EWMS, UWMS and EWAS with the following balancing parameters:

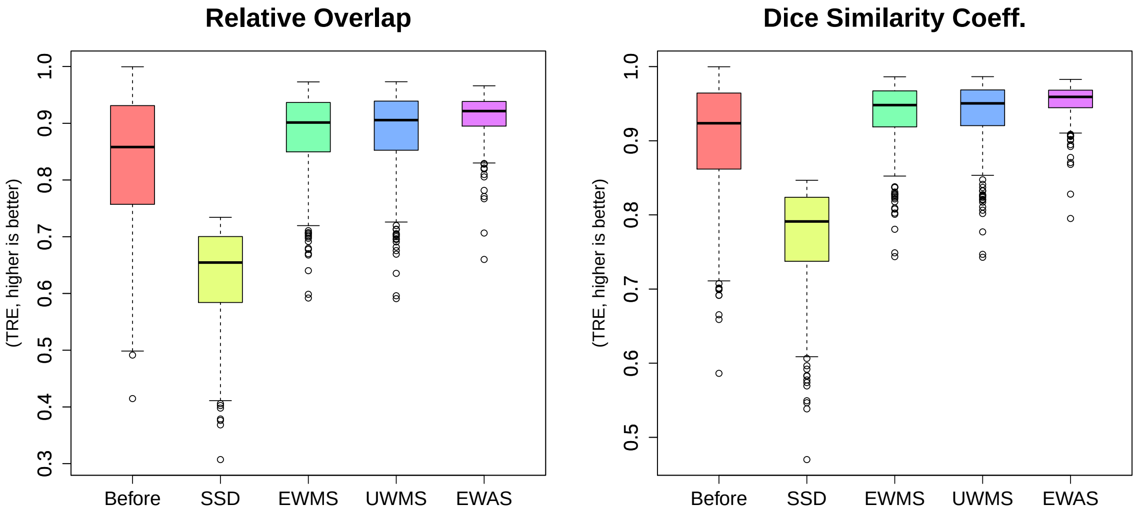
1. SSD: $\lambda = 0.005$.
2. EWMS: $\rho = 5$, $\lambda = 0.005$.
3. UWMS: $\rho_1 = 25$, $\lambda = 0.005$ and $\omega(u, v)$ defined by Eq. (2.27).
4. EWAS: $\rho = 5$, $\lambda = 0.005$.

The TRE results are shown in Fig. 4.3 and Table 4.3. SSD performs worst among all algorithms, its TRE becomes even worse than that before registration. EWMS, UWMS and EWAS all improved the TRE. EWAS achieves the best results in terms of mean and standard deviation of TRE. The performance of UWMS is slightly better than EWMS.

Table 4.2: The means of the evaluation criteria of the ellipsoid phantom experiment

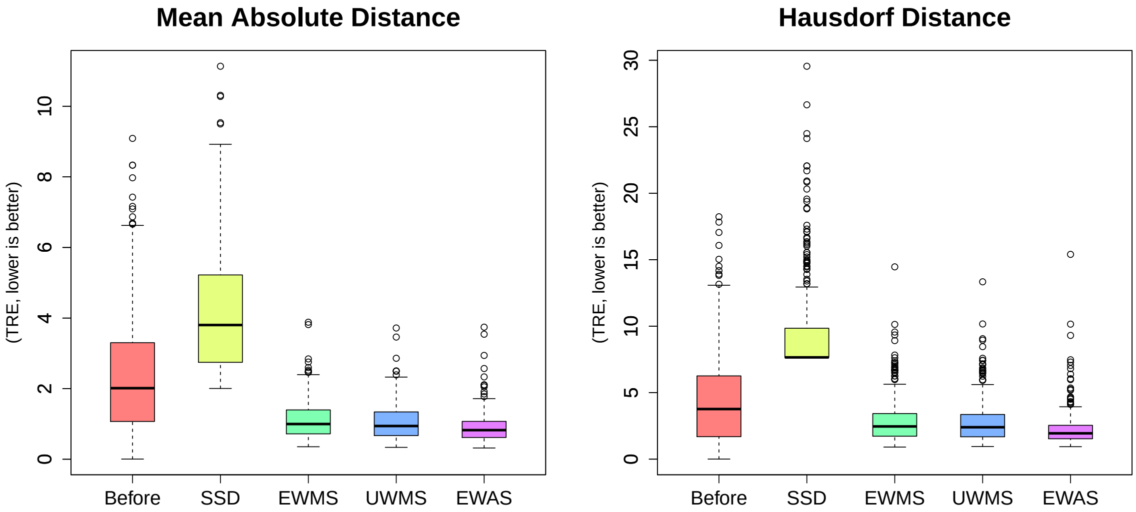
	RO	DSC	LDSC	MAD	HD
Before	0.835	0.905	2.693	2.364	4.464
SSD	0.631	0.770	1.241	4.182	9.363
EWMS	0.882	0.936	2.867	1.104	2.893
UWMS	0.886	0.938	2.914	1.052	2.811
EWAS	0.913	0.954	3.124	0.900	2.263
EWAS vs. EWMS improves	3.51% ($p < 0.001$)	1.96% ($p < 0.001$)	8.93% ($p < 0.001$)	18.5% ($p < 0.001$)	21.8% ($p < 0.001$)

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance



(a) RO, higher is better

(b) DSC, higher is better



(c) MAD, lower is better

(d) HD, lower is better

Figure 4.3: The target registration errors (TRE) in the phantom image registration. SSD: intensity-only, EWMS: equally-weighted manual surface constraint, UWMS: uncertainty-weighted manual surface constraint, EWAS: equally-weighted assimilated surface constraint.

Table 4.3: The stddev of the evaluation criteria of the ellipsoid phantom experiment

	RO	DSC	LDSC	MAD	HD
Before	0.119	0.075	1.223	1.735	3.509
SSD	0.082	0.066	0.347	1.725	3.351
EWMS	0.070	0.041	0.653	0.518	1.680
UWMS	0.070	0.042	0.670	0.501	1.635
EWAS	0.037	0.021	0.424	0.408	1.296
EWAS vs. EWMS improves	46.4% ($p < 0.001$)	48.6% ($p < 0.001$)	35.0% ($p < 0.001$)	21.3% ($p < 0.001$)	22.8% ($p < 0.001$)

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance

4.3.2 Evaluation experiment on real data

Following the real data experiment in Sec. 3.6.3, we evaluated the image registration algorithms with the NKI database. It includes 19 prostate cancer patients of with 12 (on average) repeated FBCT scans. Each scan has the bladder, prostate and rectum contoured consistently by a single expert. Table 2.1 shows the availability of the patient data. For each patient, we chose the image and the manual prostate surface at the first day as the template image $T(x)$ and surface $X(u, v)$. The image and surface at one day between Scan 9 to the last were chosen as the target image $S(x)$ and true surface $Y_g(u, v)$ (see Fig. 4.4).

The ground truth surfaces were not involved in the registration cost function. They were only used for computing the target registration error (TRE). Since all images were from the same imaging modality, the sum of squared difference C_{SSD} was used in the registration cost function.

The surfaces were reused from the surface assimilation experiment in Sec. 3.6.3.

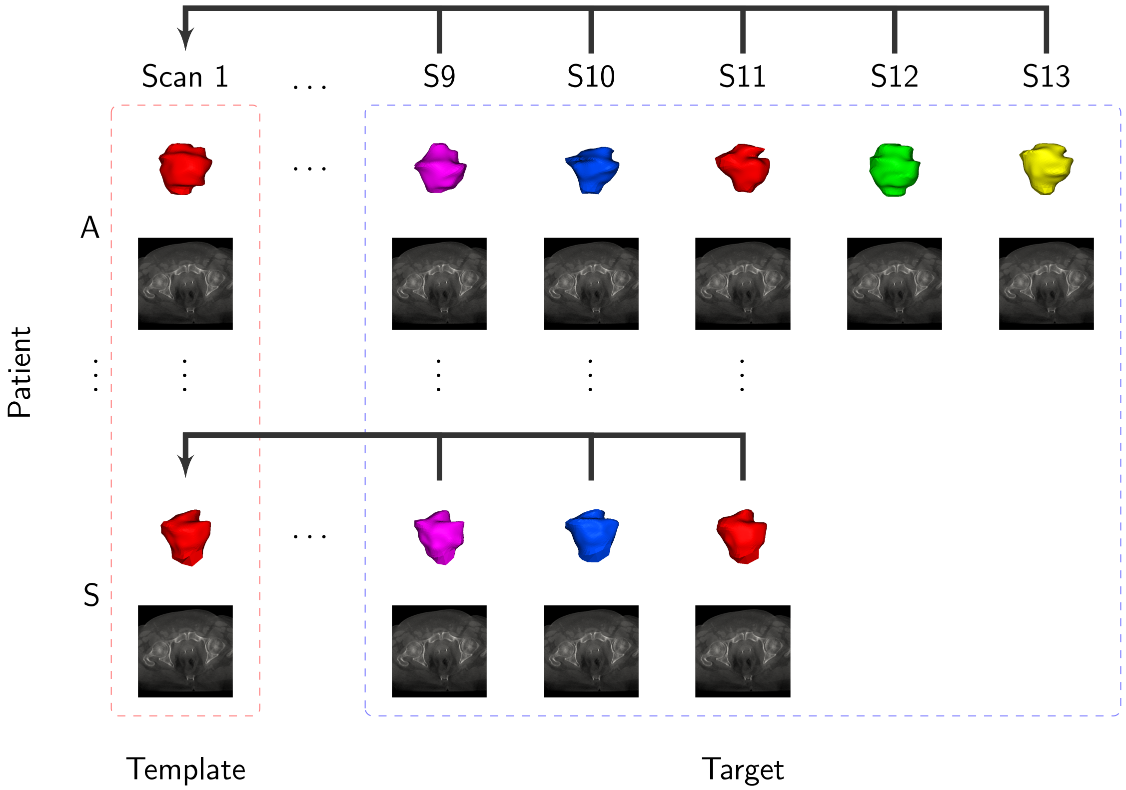


Figure 4.4: Overview of the registration experiment on the real patient data. For each patient, the initial image (Scan 1) is taken as the template image and its manual segmentation of prostate is taken as the template surface. The images and surfaces starting from Scan 9 to the end are taken as the target image and the true target surfaces. These surfaces, as the ground truth, are not involved in the registration cost function. They are used for computing the target registration error (TRE).

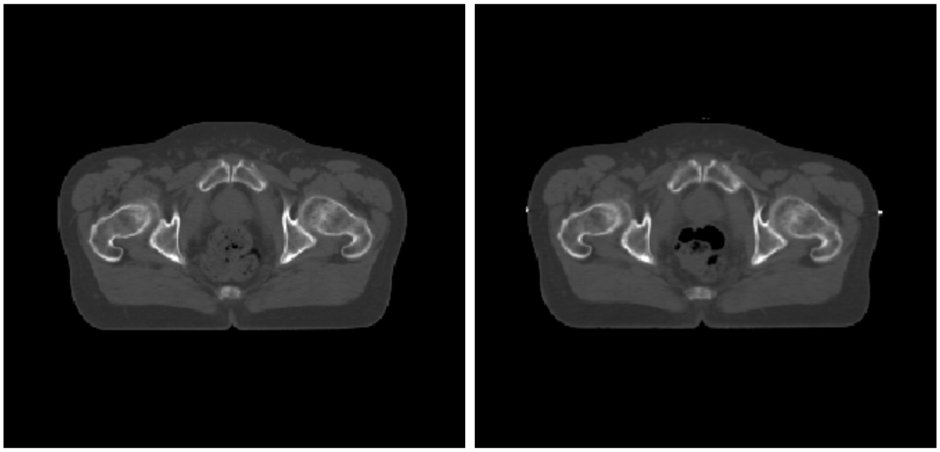
Patient A scan 9 is used as an example to explain the experiment. From Table 3.6, 53 valid surface assimilation cases were done for Scan 9. Therefore 53 pairs of observed and assimilated surfaces are available for being the target surfaces $Y(u, v)$ and $F(u, v)$ respectively. We register the images of $T(x)$ and $S(x)$ for 53 times with $X(u, v)$ and different $Y(u, v)$ and $F(u, v)$ using the algorithms: SSD, EWMS, UWMS and EWAS. The balancing parameters for each algorithm are:

1. SSD: $\lambda = 0.0005$.
2. EWMS: $\rho = 5$, $\lambda = 0.0005$.
3. UWMS: $\rho_1 = 20$, $\lambda = 0.0005$ and $\omega(u, v)$ defined by Eq. (2.27).
4. EWAS: $\rho = 5$, $\lambda = 0.0005$.

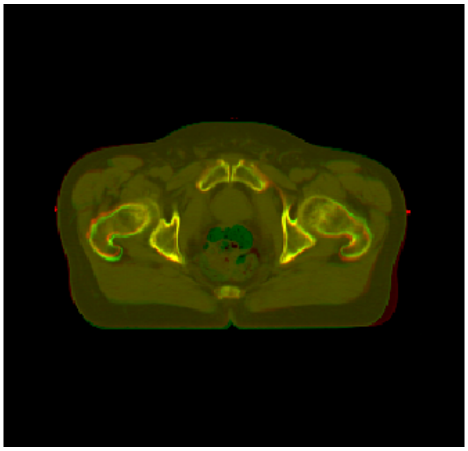
The deformed target true surface, $h(Y_g(u, v))$, was compared with $X(u, v)$ to compute the target registration error (TRE). The averaged TRE over 53 cases were reported as the points with the same x coordinate in Fig. 4.10 and Fig. 4.11.

One particular case (Fig. 4.5 and Fig. 4.6) among the mentioned 53 pairs of image registration is taken as an example demonstrating the registration results (shown in Fig. 4.7, Fig. 4.8 and Fig. 4.9).

The results of TRE over the all patients are reported in Table 4.4 and Table 4.5. As the TRE results show, EWAS achieves the best results and SSD performs worst among all algorithms. The EWAS and EWMS lines have a similar profile, both better than UWMS. Statistically, the RO and DSC improvement of EWAS over EWMS is equal. However, the LDSC, MAD and HD improvement is significantly different.



(a) Template image (b) Target image



(c) Red-Green difference

Figure 4.5: An example of the template and target images before the registration. (a) the template image of Patient A (Scan 1), (b) the target image of Patient A (Scan 9), (c) the red-green blended image showing the difference of the template and target images. The template image is tinted green and the target is tinted red. Yellowish color suggests small intensity difference.

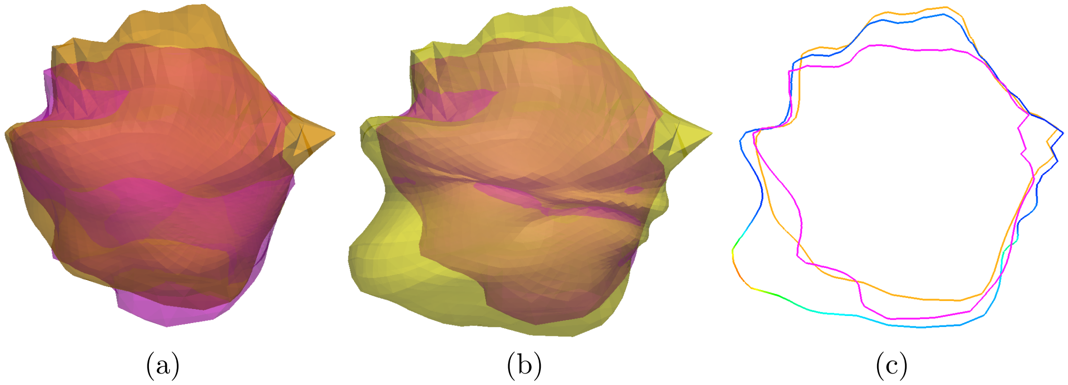


Figure 4.6: The difference among the template, the observed target and the true target surfaces of the example shown in Fig. 4.5. (a) the template (pink) and the true target (dark yellow) surfaces, (b) the template (pink) and the observed true (yellow) surfaces, (c) the slice view of the template (pink), the true target (dark yellow) and the observed target (rainbow).

4.4 Discussion

Through the phantom and patient data experiments, we have demonstrated the strength of the equally-weighted assimilated surface constraint (EWAS) in the non-rigid image registration. Comparing to the intensity-only (SSD), equally-weighted manual surface constraint (EWMS) and uncertainty-weighted manual surface constraint (UWMS), EWAS has obtained the lowest target registration error (TRE).

These results agree with the conclusion in Chapter 3 that the assimilated surface is a more accurate estimate to the truth. By matching the assimilated surface at the region lack of intensity contrast, we obtain a better result. The upper bounds of the improvement of the EWAS algorithm in the phantom and real patient experiment are governed by the improvement results in Surface Assimilation in Table 3.3 and Table 3.7, respectively. Therefore, improving the quality of surface assimilation

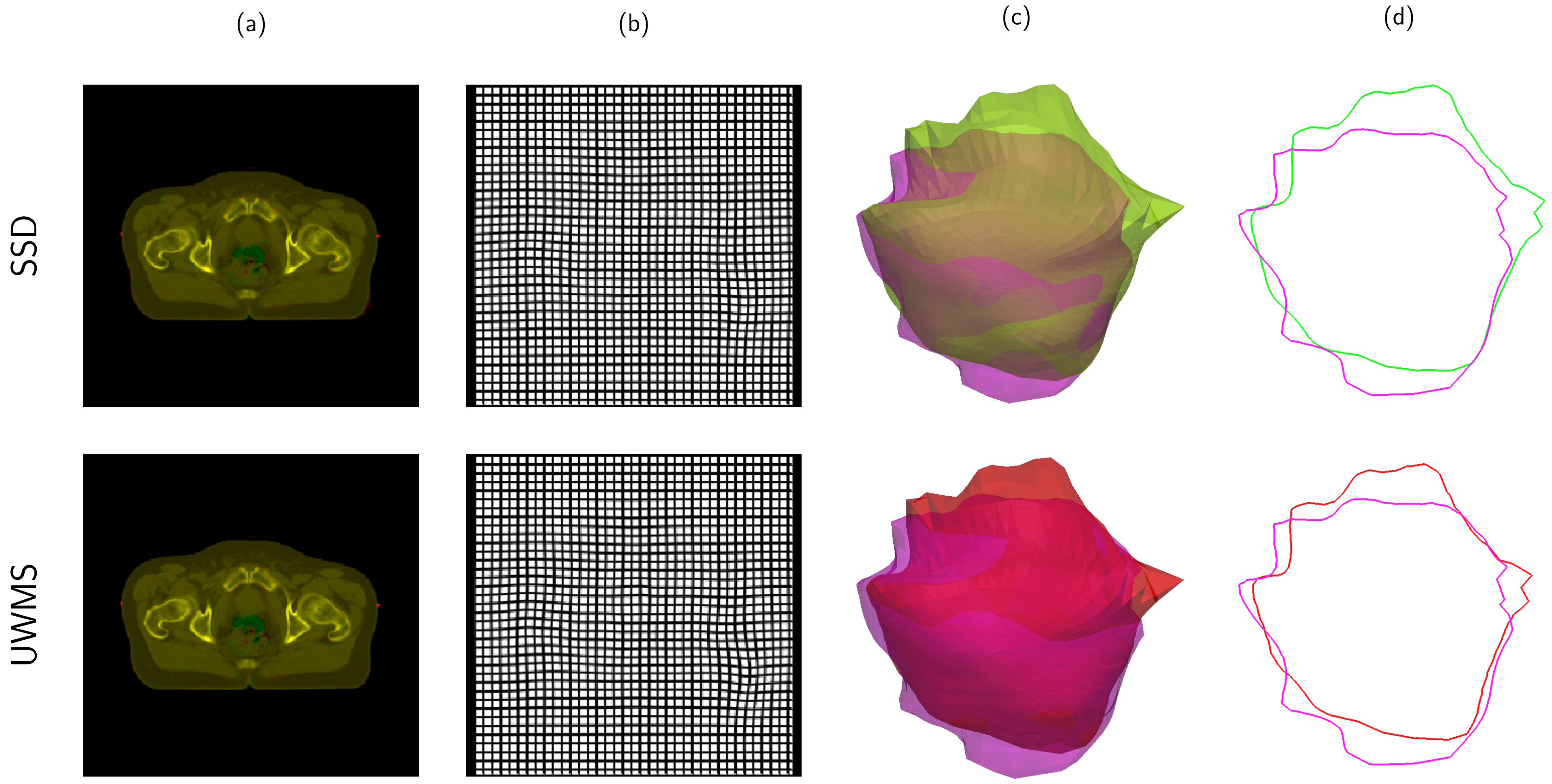


Figure 4.7: The registration results of SSD and UWMS on the patient data in the example of Fig. 4.5. (a) the red-green blended image showing the difference between the template and the deformed target images. (b) the deformed grid using the obtained transformation. (c) and (d) the 3D and 2D views of the template (pink) and the deformed true target surfaces with SSD tinted green and UWMS tinted red, respectively.

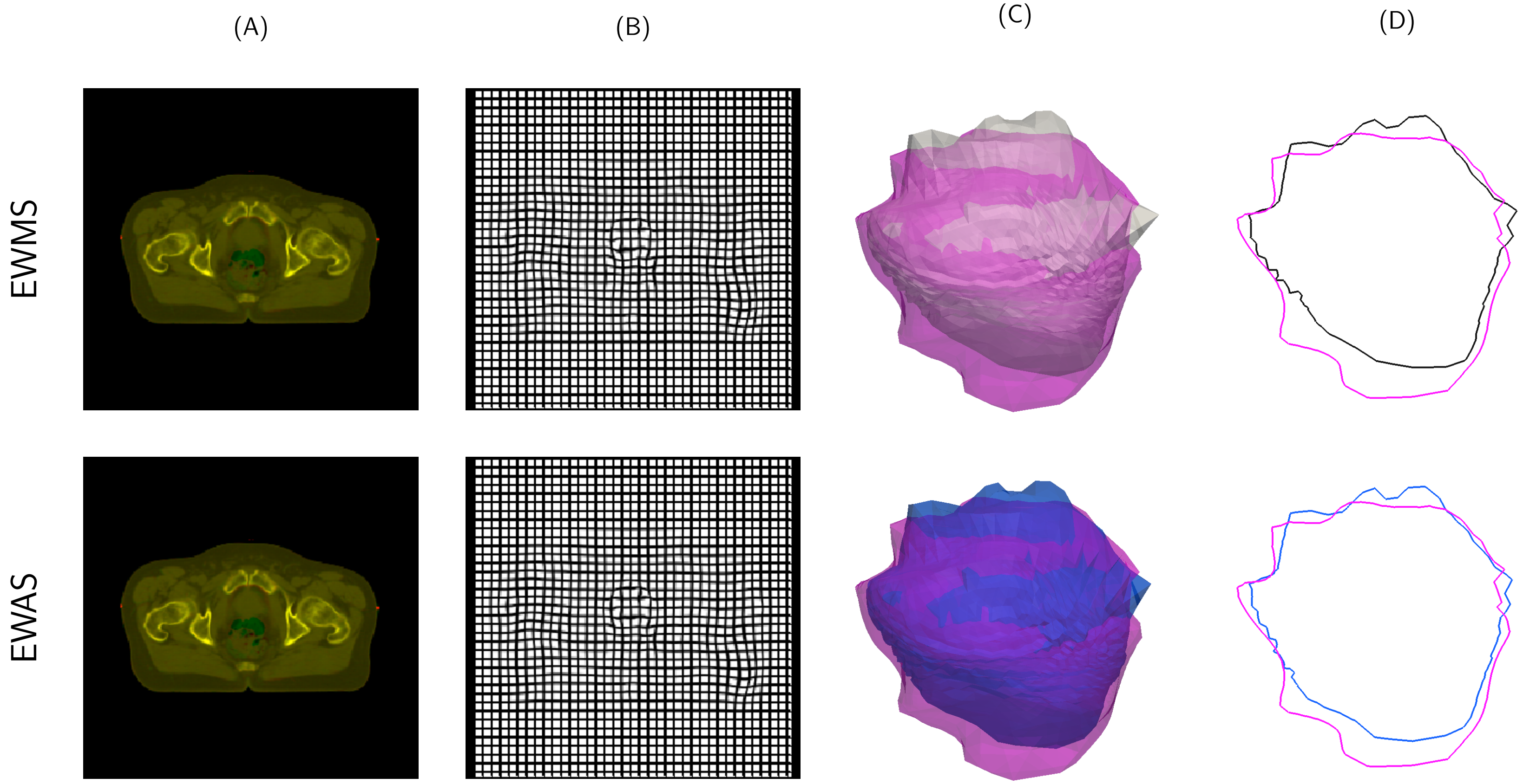


Figure 4.8: An example of the registration results of EWMS and EWAS on the patient data in the example of Fig. 4.5. (a) the red-green blended image showing the difference between the template and the deformed target images. (b) the deformed grid using the obtained transformation. (c) and (d) the 3D and 2D views of the template (pink) and the deformed true target surfaces with EWMS tinted gray and EWAS tinted blue, respectively.

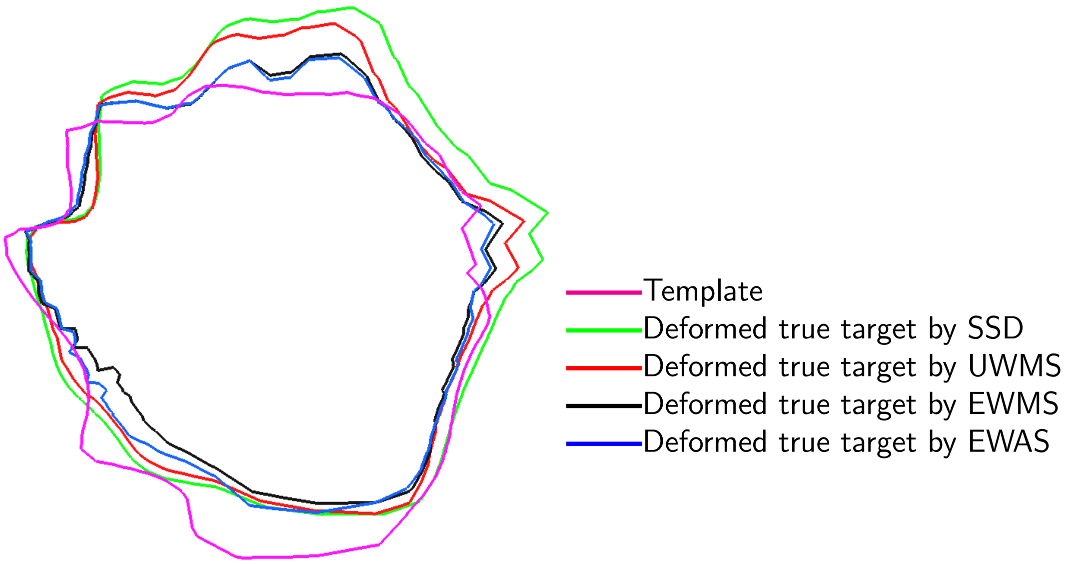


Figure 4.9: 2D slice view of the true target surfaces deformed by the different registration algorithms overlying on the template surface

algorithm remains the top priority for future work.

The performance of the uncertainty-weighted manual surface constraint, C_{UWMS}, is interesting. We initially designed this constraint by referring to a common practice in the landmark registration where the point landmarks are directly penalized by the landmark errors. Our previous study [141] shows that C_{UWMS} for the surface matching does not always help. The real patient experiment in Sec. 4.3.2 shows C_{UWMS} is worse than the equally weighted variants, C_{EWMS} and C_{EWAS}. We explain this from two aspects:

1. Different than the sparse landmarks, a surface object represent a *shape*, i.e., the points on a surface are correlated and follow a certain shape pattern. If a surface is represented as a dense point cloud in the registration, the direct penalty

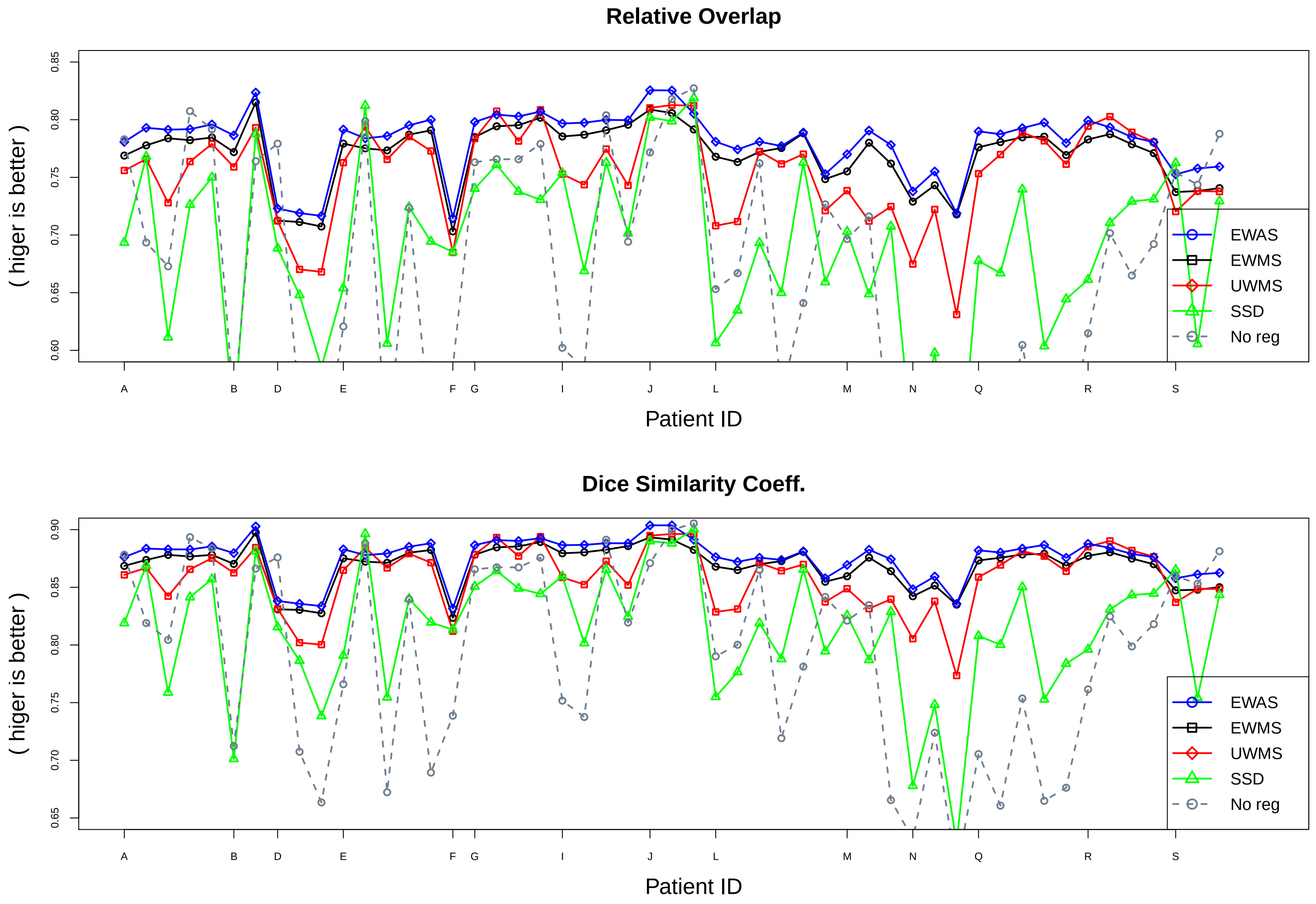


Figure 4.10: TRE results in the patient image registration (RO, DSC).

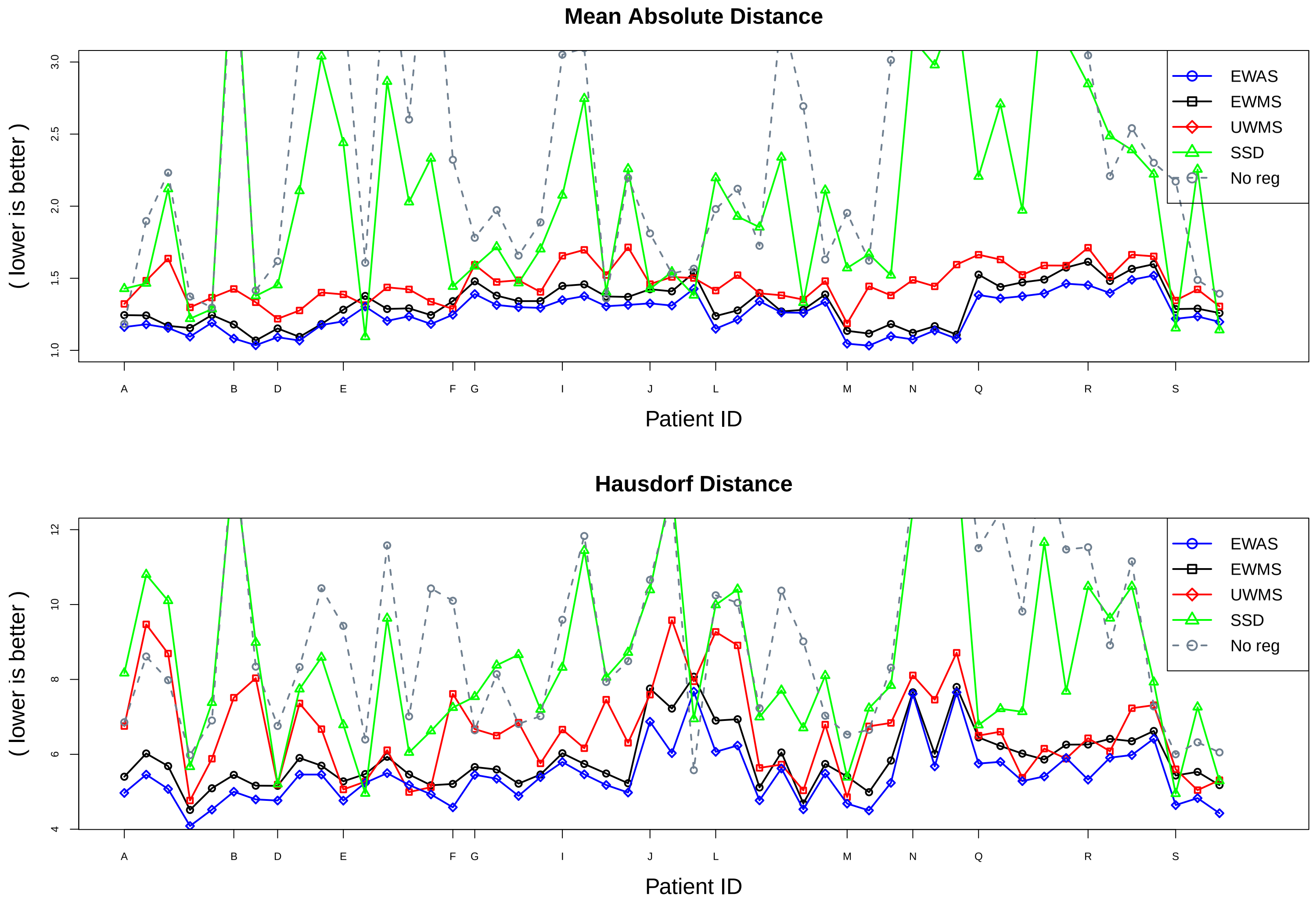


Figure 4.11: TRE results in the patient image registration (MAD, HD)

Table 4.4: The means of the evaluation criteria in the prostate registration experiment

	RO	DSC	LDSC	MAD	HD
Before	0.663	0.792	1.411	2.588	9.200
SSD	0.689	0.813	1.931	2.063	8.508
EWMS	0.771	0.869	1.931	1.322	5.879
UWMS	0.754	0.859	1.834	1.464	6.659
EWAS	0.781	0.876	1.992	1.252	5.424
EWAS vs. EWMS improves	1.30% ($p = 0.054$)	0.81% ($p = 0.055$)	3.16% ($p < 0.05$)	5.30% ($p < 0.01$)	7.74% ($p < 0.001$)

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance

ignoring this shape pattern will likely lead to unrealistic shape deformation.

2. The deformation at the region of low contrast is partially contributed by the image registration using the surrounding grayscale content. In the phantom experiment, the registration of the underlying big vertically-elongated ellipsoid object to the big horizontally-elongated one affects the deformation of the surface at the center. This underlying intensity deformation field can be either agree or disagree with the true surface deformation. C_{UWMS} only works for the former situation. If the two types of deformation compete, the equally-weighted strategy may be preferred to guard the undesirable deformation caused by the intensity matching.

As mentioned in Chapter 2, four patients' FBCT and CBCT surface segmentation errors were analyzed (see Fig. 2.13). Visually, these distribution of the segmentation errors cross patients appear inconsistent. Since the number of our training

Table 4.5: The stddev of the evaluation criteria in the prostate registration experiment

	RO	DSC	LDSC	MAD	HD
Before	0.0000	0.0000	0.0000	0.0000	0.0000
SSD	0.0000	0.0000	0.2794	0.0000	0.0000
EWMS	0.0492	0.0320	0.2794	0.3146	1.3519
UWMS	0.0343	0.0227	0.1853	0.2925	0.7982
EWAS	0.0457	0.0293	0.2693	0.2939	1.1816
EWAS vs. EWMS improves	7.1% ($p < 0.05$)	8.4% ($p < 0.05$)	3.6% ($p = 0.17$)	6.6% ($p = 0.13$)	12.6% ($p < 0.05$)

RO: Relative Overlap, DSC: Dice Similarity Coefficient, LDSC: Logit Dice Similarity Coefficient, MAD: Mean Absolute Distance, HD: Hausdorff Distance

samples is small, meaningful patterns can be missed. Additionally, the distribution of segmentation errors might depend on both human observers and also the intensity context. An alternative strategy was proposed to use the intensity context as registration constraints [118, 74].

Our original goal was to develop an image registration algorithm for registering FBCT and CBCT images. However no CBCT images were available in the NKI dataset, our whole algorithm evaluation was done using FBCT images. We can simulate the CBCT noise in the way described in [97]. This is coarse simulation. For a sophisticated CBCT noise simulation, such as considering quantum noise, detector blurring and additive system noise, see [124]. This method simulates the whole process of image projection and reconstruction, therefore requires more computational resources and parallel programming (cluster or GPU) is often needed.

4.5 Summary

We described a surface-constrained nonrigid image registration algorithm in the presence of inter-observer segmentation error (ISE). This registration algorithm used the surface assimilation method introduced in Chapter 3. In our approach, a noisy object segmentation is corrected (assimilated) using the a priori shape model and the segmentation labelling error. The assimilated segmentation surface was shown to be closer to the truth and therefore matching this surface to the template one can improve the registration accuracy. We evaluated the algorithm with the intensity-only registration, equally and uncertainty weighted manual surface registration on the phantom and the real patient data. Benchmark results show the proposed algorithm achieved the lowest target registration error over the competing algorithms.

CHAPTER 5
CONCLUSIONS

One source of error in image registration comes from the large inter-observer organ segmentation variability. Trying to reduce this error is the practical purpose of this work. It is relevant and important to the ongoing development of Image-guided Adaptive Radiotherapy (IGART) aiming for a fast and accurate cancer treatment option. Among many possible ways to estimate the segmentation uncertainty, this thesis designed one method based on the point distribution model (PDM) and showed how to use the segmentation uncertainty model to improve the image registration. Although our experiment database is limited to the prostate images, the proposed methods, from shape analysis to surface-constrained image registration, can be applied to other cancer sites, such as lung, head-and-neck and cervix. The theory of the surface assimilation is general for a wide range of image segmentation, registration and analysis problems. In summary, this dissertation makes the following contributions:

- Evaluated the group-wise surface registration algorithms for pelvic organs, built the statistical shape models and quantified the interobserver segmentation errors on the prostate.
- Developed a surface assimilation algorithm for estimating the true boundary from the observed surface.
- Designed a surface-constrained nonrigid image registration algorithm between two contoured images with the segmentation error.

A statistical shape model (SSM) captures the shape knowledge of an organ object from the training dataset. By assuming the shapes can be modeled as a linear Gaussian probability distribution, we used a Point Distribution Model (PDM) to incorporate a priori shape knowledge. The quality of a PDM largely depends on that of groupwise correspondence among the surfaces. In Chapter 2 we reviewed and evaluated different groupwise surface registration algorithms. Based on the evaluation results, the Minimum Description Length (MDL) registration algorithm was selected as the choice for shape modeling. One drawback of MDL method is the long computation time.

Inter-observer segmentation error (ISE) is a good indicator for regions that need the correction from the shape model. In regions with low ISE, the observed boundary is assumed more accurate than the shape model prediction and should be kept. In regions with high ISE, the observed boundary is assumed less accurate than the shape model prediction and should be corrected. Therefore, the quantification of the ISE was done after the organ shape modeling.

Similar to the shape model, the ISE was also modeled as linear Gaussian probability distribution and the groupwise surface registration was also used for estimating the correspondence among the training samples. The results show the inter-observer segmentation error on FBCT is smaller that on CBCT image. Therefore in the later patient experiment, we assumed that the segmentation on FBCT is accurate.

Estimating the true object boundary from the observed surface corrupted by the segmentation error is one goal of this study. In Chapter 3, we discussed the recoverability of the SSM. The orthogonality between the segmentation error and the shape model controls the accuracy of the SSM prediction. We formulated the fusion of the observed surface and the SSM-prediction as a Data Assimilation problem and solved it with the smoothing spline regression. We refer this process as surface assimilation (SA).

We introduced the theory of smoothing spline and described its application on the 1D periodic domain and the 2D spherical domain. We demonstrated the strength and weakness of the surface assimilation method using the synthetic 2D curves and 3D synthetic surfaces. We also applied this method on the real patient data. The results from the real data experiment show, on average, that the logit Dice similarity coefficient (LDSC), the mean absolute distance (MAD) and Hausdorff distance are improved by 5.0%, 7.1% and 8.5%, respectively.

Improving the registration accuracy between the FBCT and CBCT images is the primary goal of this study. In Chapter 4, we designed a B-spline image registration with the equally-weighted assimilated surface constraint (EWAS). We evaluated four algorithms on the phantom images and real clinical images: 1) the intensity-only registration (SSD); 2) the intensity and equally-weighted manual surface registration (EWMS); 3) the intensity and uncertainty-weighted manual surface registration (UWMS) and 4) the intensity and equally-weighted assimilated surface registration (EWAS). The results from the real data experiment show, on average, that the target registration errors described by logit Dice similarity coefficient (LDSC), the mean absolute distance (MAD) and Hausdorff distance are improved by 3.2%, 5.3% and

7.7%, respectively.

The prostate image registration is prone to fail due to the air bubbles in the rectum. The appearance of these irregular black holes in the CT images often caused the registration to fail (e.g. the Jacobian goes negative). Pre-processing is often needed, either to paint them out or mask out the whole rectum.

It is an interesting problem to study the registration uncertainty over the image domain. Risholm [106] proposed using Markov chain Monte Carlo (MCMC) method and Bayesian inference [105] to study the registration uncertainty. We can follow this idea to analyze the influence of our registration algorithm on dose delivery in radiotherapy. It may also help with tuning the weighting parameters in Eq. (4.8).

One limitation of the registration algorithm developed in Chapter 3 is that it requires the template and target images have been contoured before the registration. This assumption may not always be true due to the time and resources required to do the contouring, thus only the image on the planning day is contoured and the CBCT images at the treatment day are generally uncontoured. Efforts have been made on the automatic FBCT/CBCT registration problems [14]. Zhou *et al.* [144] proposed registering two images after an automatic prostate segmentation based on Active Shape Model (ASM). Even though fast, this method only segment the prostate and the segmentation and registration are two separate part. Lu *et al.* [87] proposed a joint segmentation and registration algorithm in which the bladder, prostate and rectum are segmented by level-set method. The computation time of the method reportedly needed several hours prohibiting the clinical use.

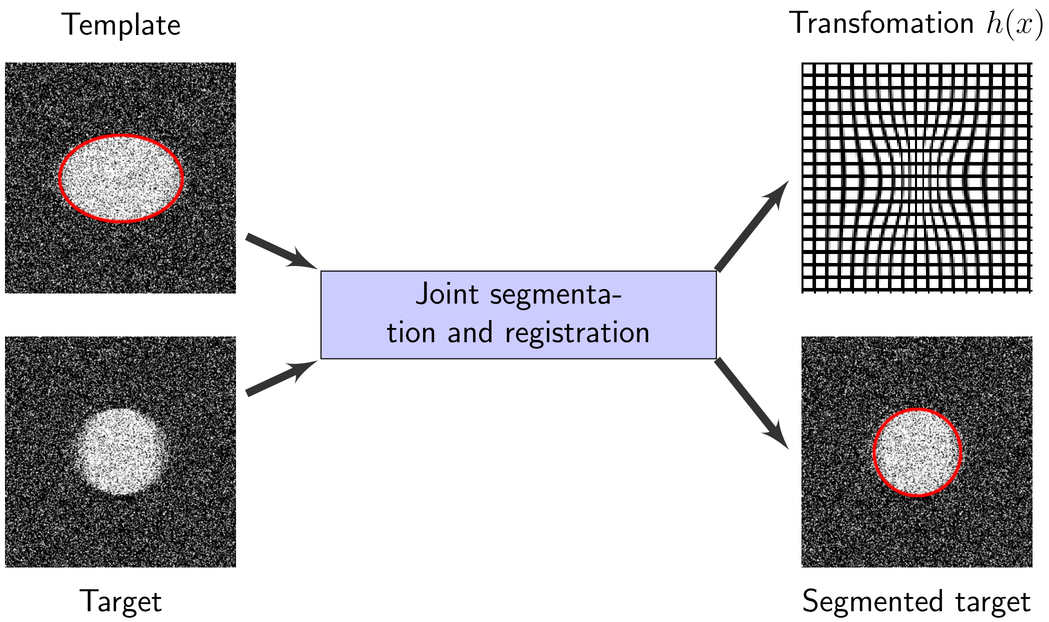


Figure 5.1: Illustration of joint registration and segmentation.

This motivates us to propose an automatic registration algorithm as future work to meet the clinical criteria such as accuracy, robustness and speed. We can combine the image registration algorithm in Sec. 4.2 with an image segmentation process. The contours the FBCT image are mapped to the CBCT image, evolves with the segmentation process and converges a good segmentation of the pelvic organs in CBCT (see Fig. 5.1 for illustration). The segmentation uncertainty (can be inferred from the intensity context) together with the shape model to penalize the unrealistic shape deformation during the local boundary finding. It will be similar to Lu's method [87] in the sense that the segmentation and registration are alternatively used. With the surface assimilation method, the time consuming Level-set segmentation is unnecessary and simpler and faster deformable segmentation methods would be sufficient. Thus, we speculate that the new algorithm would converge faster and the whole algorithm would need less time.

REFERENCES

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* US Government Printing Office, 1966.

[2] P. Alfeld, M. Neamtu, and L. L. Schumaker. "Fitting scattered data on sphere-like surfaces using spherical splines". In: *Journal of Computational and Applied Mathematics* 73.12 (1996), pp. 5 –43.

[3] N. Altman. "Kernel smoothing of data with correlated errors". In: *Journal of the American Statistical Association* 85.411 (1990), pp. 749–759.

[4] D. W. Andrews. "Asymptotic optimality of generalized CL, cross-validation, and generalized cross-validation in regression with heteroskedastic errors". In: *Journal of Econometrics* 47.2 (1991), pp. 359–377.

[5] M. Audette, F. Ferrie, and T. Peters. "An algorithmic overview of surface registration techniques for medical imaging". In: *Medical Image Analysis* 4.3 (2000), pp. 201–217.

[6] S. Balik, E. Weiss, N. Jan, et al. "Evaluation of 4-dimensional Computed Tomography to 4-dimensional Cone-Beam Computed Tomography Deformable Image Registration for Lung Cancer Adaptive Radiation Therapy". In: *International Journal of Radiation Oncology*Biology*Physics* 86.2 (2013), pp. 372–379.

[7] S. M. Berry, R. J. Carroll, and D. Ruppert. "Bayesian Smoothing and Regression Splines for Measurement Error Problems". In: *Journal of the American Statistical Association* 97 (2001), pp. 160–169.

[8] P. Besl and H. McKay. "A method for registration of 3-D shapes". In: *IEEE Transactions on pattern analysis and machine intelligence* 14.2 (1992), pp. 239–256.

[9] K. Bhatia, J. Hajnal, B. Puri, et al. "Consistent groupwise non-rigid registration for atlas construction". In: *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on.* Vol. 1. 2004, pp. 908–911.

[10] F. Bookstein. "Principal warps: Thin-plate splines and the decomposition of deformations". In: *IEEE Transactions on pattern analysis and machine intelligence* 11.6 (1989), pp. 567–585.

[11] R. Bowden. “Learning non-linear Models of Shape and Motion”. PhD thesis. Brunel University, 2000.

[12] C. Brechbühler, G. Gerig, and O. Kübler. “Parametrization of Closed Surfaces for 3-D Shape Description”. In: *Computer Vision and Image Understanding* 61.2 (1995), pp. 154–170.

[13] T. Bülow. “Spherical diffusion for 3D surface smoothing”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26.12 (2004), pp. 1650–1654.

[14] M. Cabezas, A. Oliver, X. Lladó, et al. “A review of atlas-based segmentation for magnetic resonance brain images”. In: *Computer Methods and Programs in Biomedicine* 104.3 (2011), e158 –e177.

[15] K. Cao. “Mechanical analysis of lung CT images using nonrigid registration”. PhD thesis. University of Iowa, 2012.

[16] J. Cates, P. Fletcher, M. Styner, et al. “Particle-Based Shape Analysis of Multi-Object Complexes”. In: *Proceedings of the 11th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI ’08).* 2008, pp. 477–485.

[17] J. Cates, M. Meyer, T. Fletcher, et al. “Entropy-Based Particle Systems for Shape Correspondence”. In: *1st MICCAI Workshop on Mathematical Foundations of Computational Anatomy: Geometrical, Statistical and Registration Methods for Modeling Biological Shape Variability.* 2006, pp. 90–99.

[18] M. Chao, Y. Xie, and L. Xing. “Auto-propagation of contours for adaptive prostate radiation therapy”. In: *Physics in Medicine and Biology* 53.17 (2008), p. 4533.

[19] B. Chiu, A. Krasinski, J. D. Spence, et al. “Three-Dimensional Carotid Ultrasound Segmentation Variability Dependence on Signal Difference and Boundary Orientation”. In: *Ultrasound in Medicine & Biology* 36.1 (2010), pp. 95–110.

[20] G. E. Christensen and H. J. Johnson. “Invertibility and transitivity analysis for nonrigid image registration”. In: *Journal of Electronic Imaging* 12.1 (2003), pp. 106–117.

[21] G. Christensen, B. Carlson, K. Chao, et al. “Image-based dose planning of intracavitary brachytherapy: registration of serial-imaging studies using deformable anatomic templates”. In: *International Journal of Radiation Oncology Biology Physics* 51.1 (2001), pp. 227–243.

[22] G. Christensen and H. Johnson. “Consistent image registration”. In: *Medical Imaging, IEEE Transactions on* 20.7 (2001), pp. 568–582.

[23] G. Christensen, S. Joshi, and M. Miller. “Volumetric transformation of brain anatomy”. In: *Medical Imaging, IEEE Transactions on* 16.6 (1997), pp. 864–877.

[24] G. Christensen, R. Rabbitt, and M. Miller. “Deformable templates using large deformation kinematics”. In: *Image Processing, IEEE Transactions on* 5.10 (1996), pp. 1435–1447.

[25] H. Chui and A. Rangarajan. “A new point matching algorithm for non-rigid registration”. In: *Computer Vision and Image Understanding* 89.2-3 (2003), pp. 114–141.

[26] T. Cootes and C. Taylor. “A mixture model for representing shape variation”. In: *Image and Vision Computing* 17.8 (1999), pp. 567 –573.

[27] T. Cootes, C. Taylor, D. Cooper, et al. “Active shape models-their training and application”. In: *Computer vision and image understanding* 61.1 (1995), pp. 38–59.

[28] P. Craven and G. Wahba. “Smoothing noisy data with spline functions”. In: *Numerische Mathematik* 31.4 (1978), pp. 377–403.

[29] P. Dalal, B. Munsell, S. Wang, et al. “A Fast 3D Correspondence Method for Statistical Shape Modeling”. In: *Computer Vision and Pattern Recognition, 2007. CVPR ’07. IEEE Conference on.* 2007, pp. 1–8.

[30] R. Daley. *Atmospheric data analysis.* Vol. 2. Cambridge university press, 1993.

[31] H. Daumé III. *From zero to reproducing kernel hilbert spaces in twelve pages or less.* www.umiacs.umd.edu/ hal/docs/daume04rkhs.pdf. 2004.

[32] P. Davies and M. Meise. “Approximating data with weighted smoothing splines”. In: *Journal of Nonparametric Statistics* 20.3 (2008), pp. 207–228.

[33] R. Davies, C. Twining, T. Cootes, et al. “A minimum description length approach to statistical shape modeling”. In: *Medical Imaging, IEEE Transactions on* 21.5 (2002), pp. 525–537.

[34] R. Davies, C. Twining, T. Cootes, et al. “Building 3-D Statistical Shape Models by Direct Optimization”. In: *Medical Imaging, IEEE Transactions on* 29.4 (2010), pp. 961–981.

[35] R. Davies, C. Twining, T. Cootes, et al. “3D Statistical Shape Models Using Direct Optimisation of Description Length”. In: *Computer Vision - ECCV 2002*. Vol. 2352. 2002, pp. 1–17.

[36] R. H. Davies. “Learning shape: optimal models for analysing natural variability”. PhD thesis. University of Manchester, 2002.

[37] B. C. Davis. “Medical Image Analysis via Fréchet Means of Diffeomorphisms”. PhD thesis. University of North Carolina, 2008.

[38] M. Davis, A. Khotanzad, D. Flamig, et al. “A physics-based coordinate transformation for 3-D image matching”. In: *Medical Imaging, IEEE Transactions on* 16.3 (1997), pp. 317–328.

[39] L. A. Dawson and M. B. Sharpe. “Image-guided radiotherapy: rationale, benefits, and limitations”. In: *The Lancet Oncology* 7.10 (2006), pp. 848–858.

[40] K. De Brabanter, J. De Brabanter, J. A. Suykens, et al. “Kernel Regression in the Presence of Correlated Errors.” In: *Journal of Machine Learning Research* 12 (2011), pp. 1955–1976.

[41] K. E. Deurloo, R. J. Steenbakkers, L. J. Zijp, et al. “Quantification of shape variation of prostate and seminal vesicles during external beam radiotherapy”. In: *International Journal of Radiation Oncology*Biology*Physics* 61.1 (2005), pp. 228–238.

[42] L. Dice. “Measures of the amount of ecologic association between species”. In: *Ecology* 26.3 (1945), pp. 297–302.

[43] H. Drury and D. Essen. “Computerized mappings of the Cerebral cortex: A multiresolution flattening method and a surface-based coordinate system”. In: *J. Cogn. Neurosci* 8 (1996), pp. 1–28.

[44] L. Duan, C. Farmer, I. Hoteit, et al. “Data assimilation using Bayesian filters and B-spline geological models”. In: *Journal of Physics: Conference Series* 290.1 (2011), p. 012004.

[45] J. Duchon. “Splines minimizing rotation-invariant semi-norms in Sobolev spaces”. In: *Constructive theory of functions of several variables.* 1977, pp. 85–100.

[46] D. C. V. Essen, H. A. Drury, S. Joshi, et al. “Functional and structural mapping of human cerebral cortex: Solutions are in the surfaces”. In: *Proc Natl Acad Sci U S A* 95.3 (1998), pp. 788–795.

[47] J. Feldmar, G. Malandain, J. Declerck, et al. "Extension of the ICP algorithm to non-rigid intensity-based registration of 3D volumes". In: *Mathematical Methods in Biomedical Image Analysis, 1996., Proceedings of the Workshop on.* 1996, pp. 84–93.

[48] B. Fischl, M. I. Sereno, and A. M. Dale. "Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System". In: *NeuroImage* 9.2 (1999), pp. 195–207.

[49] B. Fischl, M. I. Sereno, R. B. Tootell, et al. "High-resolution intersubject averaging and a coordinate system for the cortical surface". In: *Human Brain Mapping* 8.4 (1999), pp. 272–284.

[50] J. Fitzpatrick and J. West. "The distribution of target registration error in rigid-body point-based registration". In: *Medical Imaging, IEEE Transactions on* 20.9 (2001), pp. 917–927.

[51] J. Fitzpatrick, J. West, and J. Maurer C.R. "Predicting error in rigid-body point-based registration". In: *Medical Imaging, IEEE Transactions on* 17.5 (1998), pp. 694–702.

[52] P. Fletcher, C. Lu, S. Pizer, et al. "Principal geodesic analysis for the study of nonlinear statistics of shape". In: *Medical Imaging, IEEE Transactions on* 23.8 (2004), pp. 995–1005.

[53] M. S. Floater and K. Hormann. "Surface Parameterization: a Tutorial and Survey". In: *Advances in Multiresolution for Geometric Modelling.* 2005, pp. 157–186.

[54] M. Foskey, B. Davis, L. Goyal, et al. "Large deformation three-dimensional image registration in image-guided radiation therapy". In: *Physics in Medicine and Biology* 50.24 (2005), p. 5869.

[55] M. Francisco-Fernandez and J. D. Opsomer. "Smoothing parameter selection methods for nonparametric regression with spatially correlated errors". In: *Canadian Journal of Statistics* 33.2 (2005), pp. 279–295.

[56] A. Frangi, D. Rueckert, J. Schnabel, et al. "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling". In: *IEEE Transactions on Medical Imaging* 21.9 (2002), pp. 1151–1166.

[57] A. Frangi, W. Niessen, K. Vincken, et al. "Multiscale vessel enhancement filtering". In: *Medical Image Computing and Computer-Assisted Interventation - MICCAI'98.* Vol. 1496. 1998, pp. 130–137.

[58] W. Freeden and W. Törnig. “On spherical spline interpolation and approximation”. In: *Mathematical Methods in the Applied Sciences* 3.1 (1981), pp. 551–575.

[59] W. Freeden and T. Gervens. “Vector spherical spline interpolationbasic theory and computational aspects”. In: *Mathematical Methods in the Applied Sciences* 16.3 (1993), pp. 151–183.

[60] D. Garcia. “Robust smoothing of gridded data in one and higher dimensions with missing values”. In: *Computational Statistics & Data Analysis* 54.4 (2010), pp. 1167–1178.

[61] X. Geng. “Transitive Inverse-consistent Image Registration And Evaluation”. PhD thesis. University of Iowa, 2007.

[62] X. Geng, D. Kumar, and G. Christensen. “Transitive Inverse-Consistent Manifold Registration”. In: *Information Processing in Medical Imaging.* Vol. 3565. 2005, pp. 3–30.

[63] J. Glaunès, A. Trouve, and L. Younes. “Diffeomorphic Matching of Distributions: A New Approach for Unlabelled Point-Sets and Sub-Manifolds Matching”. In: *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on* 2 (2004), pp. 712–718.

[64] W. Greene, S. Chelikani, K. Purushothaman, et al. “Constrained non-rigid registration for use in image-guided adaptive radiotherapy”. In: *Medical Image Analysis* 13.5 (2009), pp. 809–817.

[65] C. Gu. *Smoothing spline ANOVA models.* Springer, 2002.

[66] X. Gu, Y. Wang, T. Chan, et al. “Genus zero surface conformal mapping and its application to brain surface mapping”. In: *Medical Imaging, IEEE Transactions on* 23.8 (2004), pp. 949–958.

[67] S. Haker, S. Angenent, A. Tannenbaum, et al. “Conformal surface parameterization for texture mapping”. In: *Visualization and Computer Graphics, IEEE Transactions on* 6.2 (2000), pp. 181–189.

[68] X. Han. *Feature-constrained Nonlinear Registration of Lung CT Images.* MICCAI EMPIRE. 2010.

[69] X. Han, L. Hibbard, and V. Willcut. “An Efficient Inverse-Consistent Diffeomorphic Image Registration Method for Prostate Adaptive Radiotherapy”. In: *Prostate Cancer Imaging. Computer-Aided Diagnosis, Prognosis, and Intervention.* Vol. 6367. 2010, pp. 34–41.

[70] T. Hartkens, D. Hill, A. Castellano-Smith, et al. "Using Points and Surfaces to Improve Voxel-Based Non-rigid Registration". In: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002*. Vol. 2489. 2002, pp. 565–572.

[71] D. Healy Jr, D. N. Rockmore, P. J. Kostelec, et al. "FFTs for the 2-sphere-improvements and variations". In: *Journal of Fourier Analysis and Applications* 9.4 (2003), pp. 341–385.

[72] T Heimann, I Oguz, I Wolf, et al. *Implementing the Automatic Generation of 3D Statistical Shape Models with ITK*. Tech. rep. The Insight Journal, 2006.

[73] P. Horkaew and G. Yang. "Optimal Deformable Surface Models for 3D Medical Image Analysis". In: *Information Processing in Medical Imaging*. Vol. 2732. 2003, pp. 13–24.

[74] Y. Hu, H. U. Ahmed, Z. Taylor, et al. "MR to ultrasound registration for image-guided prostate interventions". In: *Medical Image Analysis* 16.3 (2012), pp. 687–703.

[75] E. Huot, H. Yahia, I. Cohen, et al. "Surface Matching with Large Deformations and Arbitrary Topology: A Geodesic Distance Evolution Scheme on a 3-Manifold". In: *Computer Vision - ECCV 2000*. Vol. 1842. 2000, pp. 769–783.

[76] T. Huysmans, J. Sijbers, and V. Brigitte. "Automatic Construction of Correspondences for Tubular Surfaces". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 32.4 (2010), pp. 636–651.

[77] H. Johnson and G. Christensen. "Consistent Landmark and Intensity-based Image Registration". In: *IEEE Transactions on Medical Imaging* 21.5 (2002), pp. 450–461.

[78] A. Joshi, D. Shattuck, P. Thompson, et al. "Surface-Constrained Volumetric Brain Registration Using Harmonic Mappings". In: *Medical Imaging, IEEE Transactions on* 26.12 (2007), pp. 1657–1669.

[79] E. Kalnay. *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press, 2003.

[80] A. Kelemen, G. Szekely, and G. Gerig. "Elastic model-based segmentation of 3-D neuroradiological data sets". In: *Medical Imaging, IEEE Transactions on* 18.10 (1999), pp. 828–839.

[81] G. Kimeldorf and G. Wahba. “Some results on Tchebycheffian Spline Functions”. In: *J. Mathematical Analysis and Applications* 33.1 (1971), pp. 82–95.

[82] S. Kurtek, E. Klassen, Z. Ding, et al. “A novel riemannian framework for shape analysis of 3D objects”. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on.* 2010, pp. 1625–1632.

[83] S. Kurtek, E. Klassen, Z. Ding, et al. “Parameterization-Invariant Shape Comparisons of Anatomical Surfaces”. In: *Medical Imaging, IEEE Transactions on* 30.3 (2011), pp. 849–858.

[84] S. Kurtek, E. Klassen, Z. Ding, et al. “Parameterization-Invariant Shape Statistics and Probabilistic Classification of Anatomical Surfaces”. In: *Information Processing in Medical Imaging.* Vol. 6801. 2011, pp. 147–158.

[85] B. Lévy, S. Petitjean, N. Ray, et al. “Least squares conformal maps for automatic texture atlas generation”. In: *ACM Trans. Graph.* 21.3 (2002), pp. 362–371.

[86] Y. Lou, T. Niu, X. Jia, et al. “Joint CT/CBCT deformable registration and {CBCT} enhancement for cancer radiotherapy”. In: *Medical Image Analysis* 17.3 (2013), pp. 387–400.

[87] C. Lu, S. Chelikani, X. Papademetris, et al. “An integrated approach to segmentation and nonrigid registration for application in image-guided pelvic radiotherapy”. In: *Medical Image Analysis* 15.5 (2011), pp. 772–785.

[88] Z. Luo and G. Wahba. “Hybrid adaptive splines”. In: *Journal of the American Statistical Association* 92.437 (1997), pp. 107–116.

[89] C. Lütgendorf-Caucig, I. Fotina, M. Stock, et al. “Feasibility of CBCT-based target and normal structure delineation in prostate cancer radiotherapy: Multi-observer and image multi-modality study”. In: *Radiotherapy and Oncology* 98.2 (2011), pp. 154–161.

[90] F. Maes, A. Collignon, D. Vandermeulen, et al. “Multimodality image registration by maximization of mutual information”. In: *Medical Imaging, IEEE Transactions on* 16.2 (1997), pp. 187–198.

[91] T. A. Marcel Lüthi and T. Vetter. “Curvature Guided Surface Registration using Level Sets”. In: *21st Congress on Computer Assisted Radiology and Surgery.* 2007.

[92] T. Masuda, K. Sakaue, and N. Yokoya. "Registration and integration of multiple range images for 3-D model construction". In: *Pattern Recognition, 1996., Proceedings of the 13th International Conference on.* Vol. 1. 1996, 879–883 vol.1.

[93] T. McInerney and D. Terzopoulos. "Deformable models in medical image analysis: a survey". In: *Medical Image Analysis* 1.2 (1996), pp. 91–108.

[94] G. J. Meijer, J. de Klerk, K. Bzdusek, et al. "What CTV-to-PTV Margins Should Be Applied for Prostate Irradiation? Four-Dimensional Quantitative Assessment Using Model-Based Deformable Image Registration Techniques". In: *International Journal of Radiation Oncology*Biology*Physics* 72.5 (2008), pp. 1416–1425.

[95] J. Montagnat, H. Delingette, and N. Ayache. "A review of deformable surfaces: topology, geometry and deformation". In: *Image and Vision Computing* 19.14 (2001), pp. 1023–1040.

[96] B. Munsell, P. Dalal, and S. Wang. "Evaluating Shape Correspondence for Statistical Shape Analysis: A Benchmark Study". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 30.11 (2008), pp. 2023–2039.

[97] M. J. Murphy, Z. Wei, M. Fatyga, et al. "How does CT image noise affect 3D deformable image registration for image-guided radiotherapy planning?" In: *Medical Physics* 35.3 (2008), pp. 1145–1153.

[98] J. Opsomer, Y. Wang, and Y. Yang. "Nonparametric Regressin with Correlated Errors". In: *Statistical Science* 16.2 (2001), pp. 134–153.

[99] B. Paniagua, L. Bompard, J. Cates, et al. "Combined SPHARM-PDM and entropy-based particle systems shape analysis framework". In: *SPIE Medical Imaging.* International Society for Optics and Photonics. 2012, pp. 83170L–83170L.

[100] S. Pizer, G. Gerig, S. Joshi, et al. "Multiscale medial shape-based analysis of image objects". In: *Proceedings of the IEEE* 91.10 (2003), pp. 1670–1679.

[101] J. Pluim, J. Maintz, and M. Viergever. "Mutual-information-based registration of medical images: a survey". In: *Medical Imaging, IEEE Transactions on* 22.8 (2003), pp. 986–1004.

[102] G. Postelnicu, L. Zollei, and B. Fischl. "Combined Volumetric and Surface Registration". In: *Medical Imaging, IEEE Transactions on* 28.4 (2009), pp. 508–522.

[103] A. Potosky, J. Legler, P. Albertsen, et al. “Health outcomes after prostatectomy or radiotherapy for prostate cancer: results from the Prostate Cancer Outcomes Study”. In: *Journal of the National Cancer Institute* 92.19 (2000), p. 1582.

[104] P. Remeijer, C. Rasch, J. V. Lebesque, et al. “A general methodology for three-dimensional analysis of variation in target volume delineation”. In: *Medical Physics* 26.6 (1999), pp. 931–940.

[105] P. Risholm, F. Janoos, I. Norton, et al. “Bayesian Characterization of Uncertainty in Intra-subject Non-rigid Registration”. In: *Med Image Anal* 17.5 (2013), pp. 538–555.

[106] P. Risholm, J. Balter, and W. Wells. “Estimation of Delivered Dose in Radiotherapy: The Influence of Registration Uncertainty”. In: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011.* Vol. 6891. 2011, pp. 548–555.

[107] K. Rohr, H. Stiehl, R. Sprengel, et al. “Point-based elastic registration of medical image data using approximating thin-plate splines”. In: *Visualization in Biomedical Computing.* Vol. 1131. 1996, pp. 297–306.

[108] K. Rohr, H. Stiehl, R. Sprengel, et al. “Landmark-based elastic registration using approximating thin-plate splines”. In: *Medical Imaging, IEEE Transactions on* 20.6 (2001), pp. 526–534.

[109] D. Ruppert. “Selecting the Number of Knots for Penalized Splines”. In: *Journal of Computational and Graphical Statistics* 11.4 (2002), pp. 735–757.

[110] D. Ruppert and R. J. Carroll. “Theory & Methods: Spatially-adaptive Penalties for Spline Fitting”. In: *Australian & New Zealand Journal of Statistics* 42.2 (2000), pp. 205–223.

[111] S. Rusinkiewicz and M. Levoy. “Efficient variants of the ICP algorithm”. In: *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on.* 2001, pp. 145–152.

[112] G. Sharp, S. Lee, and D. Wehe. “ICP registration using invariant features”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24.1 (2002), pp. 90–102.

[113] G. Sharp, M. Peroni, R. Li, et al. “Evaluation of plastimatch B-Spline registration on the EMPIRE10 data set”. In: *Medical Image Analysis for the Clinic: A Grand Challenge* (2010), pp. 99–108.

[114] C. Studholme. “Simultaneous Population Based Image Alignment for Template Free Spatial Normalisation of Brain Anatomy”. In: *Biomedical Image Registration*. Vol. 2717. 2003, pp. 81–90.

[115] C. Studholme, D. L. G. Hill, and D. J. Hawkes. “An overlap invariant entropy measure of 3D medical image alignment”. In: *Pattern Recognition* 32.1 (1999), pp. 71–86.

[116] M. Styner, I. Oguz, S. Xu, et al. *Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM.* Tech. rep. The Insight Journal, 2006.

[117] M. Styner, K. Rajamani, L.-P. Nolte, et al. “Evaluation of 3D Correspondence Methods for Model Building”. In: *Information Processing in Medical Imaging.* Vol. 2732. 2003, pp. 63–75.

[118] L. Tang, G. Hamarneh, and R. Abugharbieh. “Reliability-Driven, Spatially-Adaptive Regularization for Deformable Registration”. In: *Biomedical Image Registration*. Vol. 6204. 2010, pp. 173–185.

[119] X. Tao, J. Prince, and C. Davatzikos. “Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain”. In: *Medical Imaging, IEEE Transactions on* 21.5 (2002), pp. 513–524.

[120] M. Taron, N. Paragios, and M.-P. Jolly. “Registration with Uncertainties and Statistical Modeling of Shapes with Variable Metric Kernels”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31.1 (2009), pp. 99–113.

[121] J.-P. Thirion. “Image matching as a diffusion process: an analogy with Maxwell’s demons”. In: *Medical Image Analysis* 2.3 (1998), pp. 243–260.

[122] H. Thodberg. “Minimum Description Length Shape and Appearance Models”. In: *Information Processing in Medical Imaging.* Vol. 2732. 2003, pp. 51–62.

[123] D. Tosun and J. Prince. “Cortical Surface Alignment Using Geometry Driven Multispectral Optical Flow”. In: *Information Processing in Medical Imaging.* Vol. 3565. 2005, pp. 519–531.

[124] S.-J. Tu, C. C. Shaw, and L. Chen. “Noise simulation in cone beam CT imaging with parallel computing”. In: *Physics in Medicine and Biology* 51.5 (2006), p. 1283.

[125] C. Twining, T. Cootes, S. Marsland, et al. "A Unified Information-Theoretic Approach to Groupwise Non-rigid Registration and Model Building". In: *Information Processing in Medical Imaging*. Vol. 3565. 2005, pp. 167–198.

[126] C. Twining, R. Davies, and C. Taylor. "Non-parametric Surface-Based Regularisation for Building Statistical Shape Models". In: *Information Processing in Medical Imaging*. Vol. 4584. 2007, pp. 738–750.

[127] F Utreras. "Cross-validation techniques for smoothing spline functions in one or two dimensions". In: *Smoothing techniques for curve estimation*. 1979, pp. 196–232.

[128] M. Vaillant and J. Glaunès. "Surface matching via currents". In: *Proceedings of Information Processing in Medical Imaging (IPMI 2005)* 3565 (2005), pp. 381–392.

[129] M. Vaillant, A. Qiu, J. Glaunès, et al. "Diffeomorphic metric surface mapping in subregion of the superior temporal gyrus". In: *NeuroImage* 34.3 (2007), pp. 1149–1159.

[130] T. Vercauteren, X. Pennec, A. Perchant, et al. "Non-parametric diffeomorphic image registration with the demons algorithm". In: *Medical Image Computing and Computer-Assisted Intervention*. 2007, pp. 319–326.

[131] G. Wahba. "Spline Interpolation and Smoothing on the Sphere". In: *SIAM Journal on Scientific and Statistical Computing* 2.1 (1981), pp. 5–16.

[132] G. Wahba. *Spline models for observational data*. Vol. 59. Society for Industrial Mathematics, 1990.

[133] Y. Wang, B. Peterson, and L. Staib. "Shape-based 3D surface correspondence using geodesics and local geometry". In: *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*. Vol. 2. 2000, pp. 644–651.

[134] Y. Wang. *Smoothing splines: methods and applications*. Taylor & Francis US, 2011.

[135] E. Weiss, J. Wu, W. Sleeman, et al. "Clinical Evaluation of Soft Tissue Organ Boundary Visualization on Cone-Beam Computed Tomographic Imaging". In: *International Journal of Radiation Oncology*Biology*Physics* 78.3 (2010), pp. 929–936.

[136] W. M. Wells, P. Viola, H. Atsumi, et al. "Multi-modal volume registration by maximization of mutual information". In: *Medical Image Analysis* 1.1 (1996), pp. 35–51.

[137] S. Wörz and K. Rohr. "Physics-based elastic registration using non-radial basis functions and including landmark localization uncertainties". In: *Computer Vision and Image Understanding* 111.3 (2008), pp. 263–274.

[138] J. Wu, M. J. Murphy, E. Weiss, et al. "Development of a population-based model of surface segmentation uncertainties for uncertainty-weighted deformable image registrations". In: *Medical Physics* 37.2 (2010), pp. 607–614.

[139] H. Yahia, E. Huot, I. Herlin, et al. "Geodesic distance evolution of surfaces: a new method for matching surfaces". In: *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on.* Vol. 1. 2000, 663–668 vol.1.

[140] B. Yeo, M. Sabuncu, T. Vercauteren, et al. "Spherical Demons: Fast Diffeomorphic Landmark-Free Surface Registration". In: *Medical Imaging, IEEE Transactions on* 29.3 (2010), pp. 650–668.

[141] C. Zhang, G. Christensen, S. Kurtek, et al. "SUPIR: Surface Uncertainty-Penalized, Non-rigid Image Registration for Pelvic CT Imaging". In: *Biomedical Image Registration.* Vol. 7359. 2012, pp. 236–245.

[142] D. Zhang and M. Hebert. "Harmonic maps and their applications in surface matching". In: *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* Vol. 2. 1999, pp. 637–663.

[143] J. Zhang and A. Rangarajan. "Multimodality Image Registration Using an Extensible Information Metric and High Dimensional Histogramming". In: *Information Processing in Medical Imaging.* Vol. 3565. 2005, pp. 7–16.

[144] J. Zhou, S. Kim, S. Jabbour, et al. "A 3D global-to-local deformable mesh model based registration and anatomy-constrained segmentation method for image guided prostate radiotherapy". In: *Medical Physics* 37.3 (2010), pp. 1298–1308.

[145] C. Zhu, R. H. Byrd, P. Lu, et al. "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization". In: *ACM Transactions on Mathematical Software (TOMS)* 23.4 (1997), pp. 550–560.

[146] B. Zitová and J. Flusser. "Image registration methods: a survey". In: *Image and Vision Computing* 21.11 (2003), pp. 977–1000.

[147] L. Zöllei, E. Learned-Miller, E. Grimson, et al. "Efficient Population Registration of 3D Data". In: *Computer Vision for Biomedical Image Applications.* Vol. 3765. 2005, pp. 291–301.

[148] K. Zou, S. Warfield, A. Bharatha, et al. "Statistical validation of image segmentation quality based on a spatial overlap index". In: *Academic radiology* 11.2 (2004), pp. 178–189.